Single Technology Appraisal

Belzutifan for treating tumours associated with von Hippel-Lindau disease [ID3932]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Belzutifan for treating tumours associated with von Hippel-Lindau disease [ID3932]

Contents:

The following documents are made available to stakeholders:

Access the **final scope** and **final stakeholder list** on the NICE website.

1. **Company submission from** Merck Sharp & Dohme**:**
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**

3. **Patient group, professional group, and NHS organisation submissions** from:
   a. Action Kidney Cancer
   b. UK Kidney Association
   c. VHL UK

4. **External Assessment Report** prepared by Kleijnen Systematic Reviews Ltd

5. **External Assessment Report – factual accuracy check**

6. **Technical engagement response from company**

7. **Technical engagement responses and statements from experts:**
   a. Clinical expert, nominated by UK Kidney association
   b. Patient expert, nominated by VHL UK/Ireland
   c. Patient expert, nominated by VHL UK/Ireland

8. **External Assessment Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews Ltd

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932)

Document B

Company evidence submission

June 2023

File name	Version	Contains confidential information	Date
NICE ID3932 STA Submission – Document B v3.0	3.0	Yes	02-JUN-2023

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE health technology evaluation guidance development manual.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables and figures

Abbreviations

Abbreviation/acronym	Definition
AE	adverse event
AIC	Akaike information criterion
ALT	alanine transaminase
APaT	all patients as treated
ASCO	American Society of Oncology
AST	aspartate amino transferase
BIC	Bayesian information criterion
BID	twice a day
BMI	body mass index
BNF	British National Formulary
BOR	best overall response
BSA	body surface area
CE	cost effectiveness
CI	confidence interval
CKD	chronic kidney disease
CNS	central nervous system
CR	complete response
CTCAE	Common Terminology Criteria for Adverse Events
DCR	disease control rate
DFS	disease free survival
DOR	duration of response
DSA	deterministic sensitivity analysis
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDSS	Expanded Disability Status Scale
EMA	European Medicine Agency
eMIT	electronic market information tool
EMR	electronic medical record
EQ-5D	EuroQoL-5 Dimension
ESMO	European Society for Medical Oncology
ESRD	end stage renal disease
EU	European Union
FDA	Food and Drug Administration
FOLFIRI	folinic acid, fluorouracil, and irinotecan
FOLFOX	folinic acid, fluorouracil, and oxaliplatin
GB	Great Britain
GBP	Great British Pounds
GI	gastrointestinal
HCRU	Health Care Resource Utilisation
HIF	hypoxia inducible factor
HR	hazard ratio
HRG	health care resource group
HRQoL	health-related quality of life
HST	Highly Specialised Technology
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
ILAP	Innovative Licensing and Access Pathway
IRC	independent review committee

Abbreviation/acronym	Definition
ITC	indirect treatment comparison
IV	intravenous
KM	Kaplan-Meier
LY	life year
LYG	life years gained
MA	marketing authorisation
MAIC	matching-adjusted indirect comparison
MDT	multidisciplinary team
MHRA	Medicine and Healthcare products Regulatory Agency
MRI	magnetic resonance imaging
MSD	Merck Sharp & Dohme
MU	million units
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NDI	National Death Index
NE	not evaluable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NMA	network meta-analysis
NR	not reported
NSCLC	non-small cell lung cancer
ORR	overall response rate
OS	overall survival
OWSA	one way sensitivity analysis
PAS	patient access scheme
PD	progressive disease
PFS	progression-free survival
PK	pharmacokinetics
pNET	pancreatic neuroendocrine tumour
PR	partial response
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life year
QD	once a day
QW	once a week
RCC	renal cell carcinoma
RCT	randomised controlled trial
RDI	relative dose intensity
RECIST	response evaluation criteria in solid tumors
RWE	real world evidence
SACT	systemic anti-cancer therapy
SAE	serious adverse event
SC	subcutaneous
ScHARR	School of Health and Related Research
SD	stable disease
SE	standard error
SLR	systematic literature review
SmPC	summary of product characteristics
SOC	standard of care

Abbreviation/acronym	Definition
STA	single technology appraisal
TA	technology appraisal
TNM	tumour, node, metastasis
TP	transition probability
TRAE	treatment-related adverse event
TSOP	Topic Selection Oversight Panel
TTE	time to event
TTM	time to metastasis
TTR	time to response
TTS	time to surgery
UK	United Kingdom
ULN	upper limit of normal
UOB	Urology Oncology Branch
USA	United States of America
VEGF	vascular endothelial growth factor
VHL	Von Hippel Lindau
WHO	World Health Organization
WPAI	Work Productivity and Activity Impairment

Definitions and descriptions of key terms used in the submission

The definitions of key terms in the indication wording of the Summary of Product Characteristics and decision problem addressed in the company submission are shown in Table 1. The GB marketing authorisation for belzutifan (Welireg) is: Welireg is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable.

Table 1 Definition of key terms in the indication for the product detailed in the submission

Term	Definition relevant in the context of this submission
who require therapy	Patients whose VHL associated tumour(s) have reached a point where in routine clinical practice the tumour requires active intervention, usually a localised procedure (localised procedure as defined later in this table). These are: • For VHL associated RCC, when the tumour reaches 3 cm in size. • For VHL associated CNS hemangioblastoma, when the tumour has grown to a size where it is causing symptomatic disease. • For VHL associated pNET, when the tumour reaches 2 cm in size and is continuing to grow and is likely to metastasise despite treatment with a somatostatin analogue.

Term	Definition relevant in the context of this submission
localised procedures	This encompasses all non-systemic (i.e. non-pharmacological) interventions that are used in routine clinical practice to remove (wholly or partially) the VHL-tumour affected organ. This includes radiotherapy, radiofrequency ablation, thermo-ablation, cryoablation, microwave ablation, irreversible electroporation, and any other image-guided ablation targeted at these tumour(s) (1), and all surgical procedures with the objective of removing or reducing the size of the tumour. This includes procedures that are both partial- and whole-organ.
unsuitable or undesirable	Circumstances where localised procedures (as defined above) are unsuitable or undesirable are: • Where the relevant VHL associated tumour(s) have reached the threshold where localised procedures would usually be employed, but they should not/cannot be employed due to reasons including: o The localised procedures would effectively result in important loss of organ function e.g.: ▪ In VHL-associated RCC, when the localised procedure would render the patient renal replacement therapy-dependent. ▪ In VHL-associated CNS hemangioblastoma, when the localised procedure would involve removal of tissue that would lead to severe neurological or neuromuscular deficits due to removal of functionally important neurological tissue equating to severe permanent disability. ▪ In VHL-associated pNET, when the localised procedure would lead to loss of pancreatic and splenic function leading to life-long pancreatogenic diabetes and immune-compromisation such that the patient will require life-long insulin therapy, pancreatic digestive enzyme therapy, and antibiotic therapy. o Employment of available localised procedures in the specific circumstance of the patient would lead to problematic adverse events of the procedure itself in the patient and so are effectively contraindicated. ▪ E.g. radiotherapy to treat a VHL associated CNS hemangioblastoma that causes inflammation of the surrounding tissues that would result in problematic neurological adverse events and disability. o Localised procedures are not possible due to a surgical intervention being required/most appropriate but the location of the tumour being not safely accessible for surgery (i.e. effectively surgically inaccessible). ▪ E.g. a VHL associated CNS hemangioblastoma located in a place in the upper spinal cord or brainstem that cannot be safely accessed for surgery.

The GB marketing authorisation and the MK-6482-004 study in belzutifan that forms the clinical evidence base for this submission (described later in this document), namely are slightly different:

- The marketing authorisation is for patients with VHL associated “renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or

pancreatic neuroendocrine tumours (pNET)" i.e. could have any one of these tumours, whereas all patients in the MK-6482-004 study had RCC, and the subgroups of patients with CNS hemangioblastoma and pNET for which study results are available necessarily also had RCC.

- The marketing authorisation is for patients "who require therapy" for the VHL-associated tumours, whereas the MK-6482-004 study's participant eligibility criteria specified that patients who had an immediate need for surgical intervention for tumour treatment were excluded.

- The marketing authorisation is for patients "for whom localised procedures are unsuitable or undesirable" whereas this was not part of the participant eligibility criteria for the MK-6482-004 study.

The Medicines and Healthcare products Regulatory Agency (MHRA) granted this marketing authorisation in this indication based on the MK-6482-004 study and its results, i.e. based on which patients would benefit from treatment with belzutifan was demonstrated by the evidence from this very study. Such misalignments between marketing authorisation wording and supporting clinical trial patient population characteristics are not unusual for rare and highly specialised indications such as this one.

B.1 Decision problem, description of the technology and clinical care pathway

Summary of the decision problem, technology, and clinical care pathway

- This submission covers belzutifan's (Welireg) full marketing authorisation: Welireg is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable.

- Belzutifan is an orally administered medicine that selectively targets a protein called hypoxia inducible factor (HIF) - 2α. HIF-2α levels are raised in people with VHL, which can lead to the growth of both benign and malignant tumours. By blocking the activity of HIF-2α, belzutifan slows down worsening of VHL and improves symptoms.

- Von Hippel-Lindau disease is a long-term debilitating and life-limiting genetic disease caused by a mutation (fault) in the VHL gene. It is a disease characterised by growth of tumours in many organs of the body. VHL disease is different in every patient, even within the same family. The trajectory of VHL disease in patients is variable and unpredictable. Some patients may only develop a few or a single tumour in a single organ while other patients may develop multiple tumours across multiple organs throughout their life. In the worst affected patients tumours may arise repeatedly in the same organ resulting in organ function impairment. Tumours may arise in locations where surgical resection of the tumour would lead to organ function impairment or loss. In the case of RCC this can result in loss of kidney function leading to the requirement for lifelong dialysis, in pNETs this can result in Whipple's procedures or pancreatectomy leading to lifelong pancreatogenic diabetes, in CNS hemangioblastomas this can results in neurological deficits, blindness, or problematic brain injury leading to severe disability requiring 24-hour care.

- Currently there are no systemic therapeutic interventions authorised or funded in the UK for the treatment of VHL-associated cancer. In patients with pNETs, a somatostatin analogue is used as 1st line therapy to treat these tumours when it reaches 1 cm in diameter, however this is the case regardless of whether the pNET is associated with VHL disease (3, 4)

- The objective of care in patients with VHL-associated tumours is to prevent tumours from metastasising while maintaining functioning of the affected organs. At the same time clinicians attempt to minimise symptom burden and maintain patients' quality of life. A delicate balancing act, with decisions unique to individual patients on whether they would prefer to face the deeply undesirably sequelae following surgeries that cause loss of organ function (such as dialysis, pancreatogenic diabetes, or severe neurological disability, for the rest of their life), or the risk of letting tumours grow.

- The indication wording for belzutifan has a degree of ambiguity worth highlighting from the beginning [bold emphasis added by MSD]: *Welireg is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease* ***who require therapy*** *for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for* ***whom localised procedures are unsuitable or undesirable.*** This results in a subgroup of patients with VHL-associated tumours that must meet criteria:

    - First, they *require therapy*. In the absence of any approved medical treatments for VHL we interpret this to mean a surgical or related procedure for tumours >3cm in diameter for RCC, or pNETs >2cm in diameter, or CNS tumours causing symptoms that require an intervention.

    - Second, *localised procedures are unsuitable or undesirable*. It is important to note that the medical definition of localised procedures (described in Table 1) includes for example minor, partial nephrectomies and full nephrectomies resulting in organ loss and the

significant associated consequences. In extensive consultation with clinicians there is a clear medical understanding of undesirable and unsuitable (see below for further explanation). It does *not* mean a patient would rather not have a minor surgery. It is easiest to consider this criterion by its opposite: if they can have successful localised procedures they should have localised procedures.

- The current alternative to belzutifan in routine clinical practice in the UK at this stage for these patients is a complex mix of interventions.
- The current clinical standard of care (SoC) for patients with VHL RCC or pNETs for whom localised procedures are unsuitable or undesirable are still localised procedures. However, in these patients these localised procedures would result in loss of organ function with sequelae that would not allow patients to live a healthy life. In current UK clinical practice, patients undergo these procedures because they are preferable to the dire consequence of letting the tumour continue to grow.
- No equity or equality considerations are anticipated, although the inherited nature of the disease means some families are disproportionately impacted over multiple generations. There are inequities in the type of service available to VHL patients depending on which centre leads their care

B.1.1 Decision problem

The submission covers the technology's full marketing authorisation for this indication.

Table 2 The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults who require therapy for renal cell carcinoma, central nervous system hemangioblastomas, or pancreatic neuroendocrine tumours caused by von Hippel-Lindau disease, for whom localised procedures are unsuitable or undesirable	Adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable.	N/A.
Intervention	Belzutifan	Belzutifan	N/A
Comparator(s)	Renal cell carcinoma: • Standard of care without belzutifan • For advanced or metastatic disease, monotherapy or combination therapy with immunotherapies or kinase inhibitors Central nervous system hemangioblastomas: • Standard of care with belzutifan Pancreatic neuroendocrine tumours: • Standard of care without belzutifan • For unresectable or metastatic disease, monotherapy with lutetium (177Lu) oxodotreotide or combination therapy with everolimus and sunitinib	For VHL associated RCC, pNET, and CNS hemangioblastomas: • Current standard of care (SoC) ***without*** belzutifan. There are no medical treatment options approved or funded in the UK at the point in which belzutifan is indicated. Localised procedures are used, though they should be considered “last resort” interventions.	The relevant comparators are: • Primary tumour RCC or pNET: surgery resulting in loss of organ function • Primary tumour CNS hemangioblastoma: surgery with risk of problematic brain injury, or do nothing and risk problematic brain injury No treatments for advanced or metastatic disease are relevant as comparators as these would be used after treatment with belzutifan. The purpose of belzutifan is to **prevent** tumours reaching the advanced or metastatic stage. Treatments for metastatic disease are included as subsequent treatments in the economic model.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • tumour size reduction • reduction in number of surgical interventions • adverse effects of treatment • health-related quality of life.	The following outcomes were collected as part of the MK-6482-004 study: • response rates • reduction in number of surgical interventions • adverse effects of treatment • progression-free survival • tumour size reduction	Overall survival was not a designated predefined outcome in the MK-6482-004 trial. Health-related quality of life data were also not collected as part of the MK-6482-004 study. Overall survival and HRQoL are considered in the cost-effectiveness analyses, derived from sources other than the MK-6482-004 study.

Patients had to have sufficient organ function (as described in the MK-6482-004 study participant eligibility criteria described in section B.2.3 later in this document) to receive belzutifan.

B.1.2 Description of the technology being evaluated

Table 3 Technology being evaluated

UK approved name and brand name	Belzutifan (WELIREG®)
Mechanism of action	Belzutifan selectively targets a protein called hypoxia inducible factor (HIF) - 2α. HIF-2α levels are raised in people with VHL, which can lead to the growth of both benign and malignant tumours. By blocking the activity of HIF-2α, belzutifan slows down worsening of VHL and improves symptoms. Belzutifan is administered orally.
Marketing authorisation/CE mark status	Belzutifan has a GB marketing authorisation for the indication in this submission that was first granted on 31-MAY-2022 with Medicines & Healthcare products Regulatory Agency (MHRA) marketing authorisation number PL 53095/0087 (5).
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Welireg is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable (5).
Method of administration and dosage	The recommended dose of Welireg is 120 mg (three 40 mg tablets) administered orally once daily, with or without food. Tablets should be swallowed whole. Treatment should continue until disease progression or unacceptable toxicity occurs (5).
Additional tests or investigations	N/A
List price and average cost of a course of treatment	The list price of belzutifan is £11,936.70 for a 90 tablet pack of Belzutifan 40mg. Based on a median time on treatment (ToT) of [redacted], the average cost of treatment is [redacted].
Patient access scheme (if applicable)	None.

B.1.3 Health condition and position of the technology in the treatment pathway

This section describes VHL disease, the burden of the disease, its aetiology and epidemiology, followed by a description of the current UK treatment pathway for this disease, and where treatment with belzutifan would fit into it.

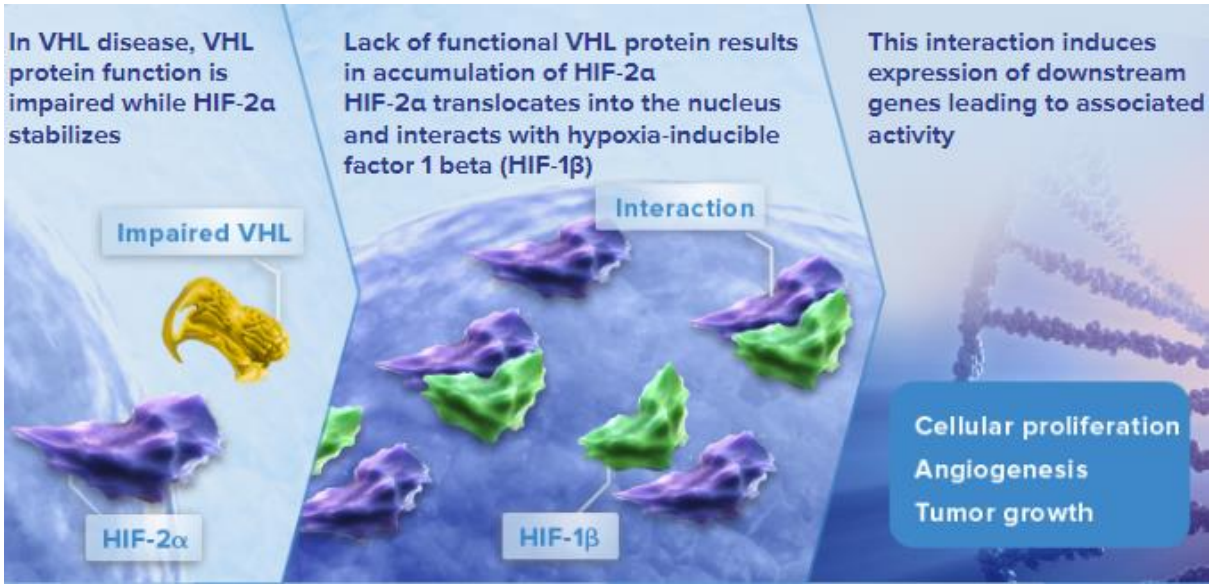
Health condition

Von Hippel-Lindau disease

Von Hippel-Lindau (VHL) disease is a genetic disorder characterised by the growth of tumours in many organs of the body, including in the kidneys, pancreas, adrenal glands and inner ear, as well as abnormal growth of blood vessels in the eye, brain and spinal cord. It is a long-term, debilitating and life-threatening disease due to the complications caused by the tumours that has wide-ranging effects on the body (6). The types and locations of tumours that can arise from VHL disease are illustrated in Figure 2.

The disease is caused by a defect in the VHL gene which is responsible for the production of a protein that regulates cell growth. pVHL is a protein encoded by the VHL gene that is critical for the regulation of hypoxia-inducible factors (HIF, Figure 1). Normally, this gene helps to keep cell growth and proliferation in check. The loss of pVHL function results in oncogenic stimulation that leads to angiogenesis, cell proliferation, cell survival, erythropoiesis and, ultimately, the growth of cysts and benign and malignant tumours (7-10). Consequently, in patients with VHL disease, cells in as many as 10 parts of the body do not die as they normally would to be replaced by healthy cells. Instead, there is a proliferation of cells that results in the formation of cysts, benign tumours and malignant tumours.

Figure 1 Molecular basis of VHL disease



HIF-2α: hypoxia-inducible factor alpha; VHL: Von Hippel-Lindau
Source: Mechanism of action of belzutifan (11)

Unlike in cancer, where a proliferation of cells is likely to be spontaneous and initially restricted to a single anatomical location, a person with VHL may experience cell proliferation and tumour growth on repeated occasions in multiple organs at different times or at the same time. Most VHL tumours begin as benign tumours. Benign tumours are those that stay in their primary location without invading other sites of the body. They do not spread to local structures or to distant parts of the body. Benign tumours tend to grow slowly and have distinct borders (12).

Benign tumours in this patient population are not problem-free and have serious negative consequences for the patient as they grow. These tumours cause an increased pressure on the structure around them creating symptoms including severe pain and/or disability.

While tumours may be multi-system, i.e., in multiple locations at the same time, patients often have a “primary tumour” that drives treatment decisions (unlike other oncology therapy areas, the primary tumour we refer to in VHL for this submission is not necessarily the first tumour). This is reflected in belzutifan’s marketing authorisation that identifies RCC, pNET and CNS hemangioblastoma “primary tumours”. Some unlucky patients have multiple significant tumours and the treatment plan needs to optimise the least worst approach.

Whilst VHL is a rare genetic disease the phenotypes are highly heterogenous and patients' experiences are very different. All diagnosed patients are kept under surveillance at regional genetic centres (13). Some patients require few, if any, surgeries in their lifetime, though being a carrier of a genetic mutation impacts life decisions such as whether to have children. At the other end of the spectrum people have a torrid experience. The most severely affected undergo many surgeries in their lifetime including surgery where organ function will be significantly impaired or completely cease, where there will be significant risk to neurological function due to CNS lesions, that require them to require medication and/or dialysis for the rest of their lives. There are patients on the spectrum between these two extremes. It is our understanding based on the regulatory process that the intent of the MHRA indication are patients that cannot be well managed by surgery.

Figure 2 Tumours that arise from VHL disease

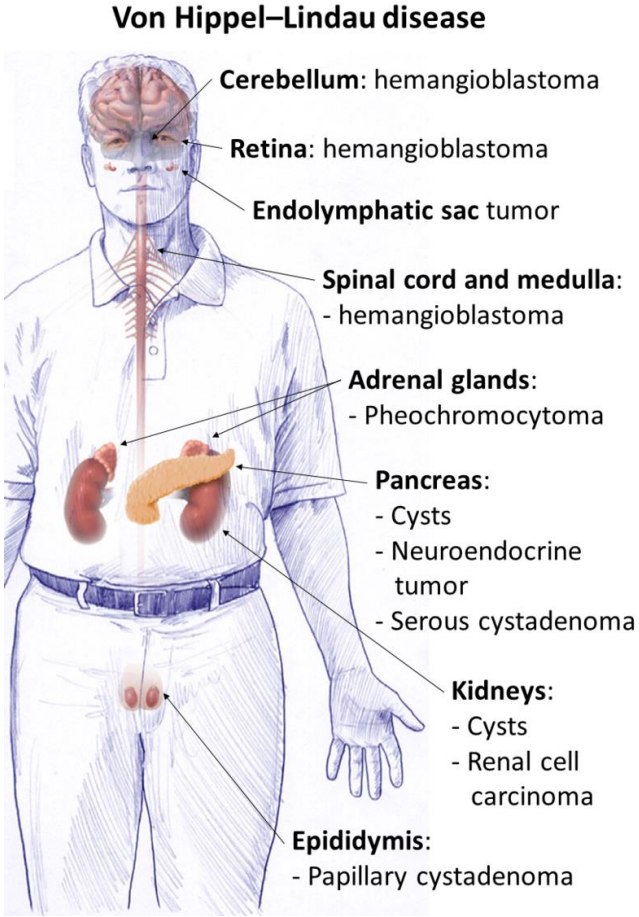


Image source: Figure 1, Schunemann et al. 2016 (14), adapted from Lonser et al. 2003 (15).

Burden of VHL disease

VHL disease significantly shortens life and severely impairs quality of life. Data from the North West Regional Genetic Register Service and North West Regional Cancer Intelligence Service in England shows that VHL disease reduces median life expectancy by nearly 19 years in men and by nearly 34 years in women, with 73% of people with VHL disease having death recorded as having a VHL-disease related cause (16).

VHL severely impairs patients' quality of life, it often requires multiple surgeries or other local procedure such as radiotherapy, thermo-ablation, or cryotherapy over a patient's lifetime, contributing to medical anxiety and fatigue (17). RCC and pNET tumours can and do metastasize. The objective of treatment for patients with these tumours is to balance preventing tumours metastasising (through surgical removal) and maintenance of organ function. It is not uncommon for those diagnosed with RCC to have multiple surgeries leading to reduced kidney function or organ loss equivalent to end stage renal disease leading to kidney dialysis. As is well documented, patients on dialysis are at increased risk of serious cardiovascular events. The relative risk in patients on dialysis compared to the general population is >5 fold for myocardial infarction, ~6 fold for ischaemic stroke, and ~2.2 fold for venous thromboembolism, resulting in an 10x higher age-adjusted cardiovascular mortality (18-21). Approximately 60% of patients with end-stage kidney disease on haemodialysis will experience cardiovascular death, and >50% of older patients will die within the first year of starting dialysis (22, 23). Kidney dialysis patients have a reduced life expectancy. VHL patients on dialysis with no remaining kidney function may develop tumours in other organs or blood vessels.

The presence of pNETs also creates an undesirable set of circumstances. Patients who undergo a Whipple surgery, and have only part of their pancreas removed, become reliant on life-long medication for diabetes and/or pancreatic insufficiency. Whipple surgery patients may develop insulin-dependent diabetes. pNET patients may also have a full pancreatectomy leading to pancreatogenic or Type 3c diabetes, while similar to Type 1 diabetes mellitus, pancreatogenic diabetes is unique in that the individual is deprived of both the exocrine and the endocrine functions of the pancreas: insulin and glucagon, as well as other pancreatic hormones (24).

Pancreatogenic diabetes exposes people who have undergone such surgery to life-threatening complications that arise from the wide, fast, unpredictable and inexplicable swings in blood glucose concentration this causes, often resulting in ketoacidosis or hypoglycaemic coma (24). Pancreatogenic diabetes is still treated primarily with insulin therapy (25, 26). Treatment of type 3c diabetes also should include treatment of pancreatic exocrine insufficiency with pancreatic enzyme replacement to improve absorption of fats and fat-soluble vitamins and prevent malabsorption of nutrients (26).

A Whipple surgery involves removal of the head of the pancreas, a part of the duodenum, some of the bile duct and the gall bladder. The stomach, bile duct and remainder of the pancreas will then be rejoined to the small bowel. The operation usually takes 4-6 hours. Long-term effects of this procedure include diabetes, pancreatic insufficiency, and change in bowel habit (27).

The fear of disability or death from surgery weighs heavily on minds of patients. Particularly where their tumours are asymptomatic but require intervention. i.e. asymptomatic pancreatic tumours but they have progressed necessitating the Whipple procedure which then leaves patients feeling worse than before the surgery. Patients talked about "scanxiety" ahead of every appointment where they prepare to be told that they need surgery.

While CNS hemangioblastomas do not metastasise, they can be some of the most difficult to manage. A patient that has had multiple CNS surgeries or has tumours in a location where surgery would carry too great a risk of serious complications are not suitable for surgery. There have been cases where patients have been left paralysed due to VHL tumours or their surgical treatment; with significant nerve injury or on a permanent tracheostomy, who now require 24/7 social care and increased surveillance. Patients who have had CNS manifestations need extensive support in day to day living (3). Surgery can be associated with considerable morbidity, including facial palsy, neurological damage, paresis, blindness, meningitis, and death (3).

The above describes each primary tumour site: RCC, pNET, and CNS hemangioblastomas in isolation. VHL patients may have multiple tumours in multiple

sites e.g. patients may have both RCC tumours and CNS hemangioblastomas at the same time, requiring clinical management of several tumours at the same time. Due to the nature of VHL disease, new tumours can appear in the same organ (as well as in other locations) after a tumour has been removed.

Patients at all stages of VHL disease have to plan their lives to revolve around constant scans, surgeries, recovery, and the long-term consequences of reduced function resulting from their surgeries. Many experience frustration arising from misrepresentation or misunderstanding of the disease due its complexity and different forms of manifestation making it difficult for many to fully understand. All of this has a severely negative effect on patients' quality of life. While quantitative data are scarce in the scientific literature on the impact of VHL disease and specifically VHL-associated cancer on patients' quality of life (28), feedback from patients consistently demonstrates a considerable negative impact on the quality of life and levels of distress of patients as well as their family members (29). Additionally, there is the impact of VHL-related appointments on patients education and careers (including not easily being able to relocate to progress their career, or having to delay surgery because they've had too many sick days to be eligible for sick pay that year), the impact of lifestyle changes they have to make around the need for monitoring appointments and surgeries (including having to pay for hotel accomodation to go to appointments with specialists), the impact on family reproductive rights, impact of multiple individuals within one family being affected, and the enormous mental health impact of living with an aggressive manifestation of VHL disease.

Statements from patients with VHL disease are provided below that illustrate the psychological, emotional, and quality life impact of VHL disease and dealing with the tumours that arise from it. These are sourced from the VHL UK/Ireland charity website (https://vhl-uk-ireland.org/stories/) and the 2022 patient/carer survey (17, 30).

Not every patient has such severe presentation, these patients are not eligible for chronic systemic treatment with belzutifan according to the GB marketing authorisation.

Since being diagnosed with VHL 13 years ago I have had 3 cerebellar haemangioblastomas (brain tumours) removed, a full Whipples procedure & my latest operation (of which I am currently at home recovering) a Laparoscopic Adrenalectomy – the removal of my left adrenal gland and its attached tumour.

It's only in the last 2 years, since going through 3 operations within the space of 20 months, also the fact that I'm 'growing up' (not sure that I ever really will!), that I've started to understand the full impact VHL has not only on yourself but your family and friends too. The pain you see in their eyes when you get the results of various MRI scans, blood tests almost hurts more than the physical results themselves.

…

My biggest operation, both physically and mentally, was the Whipple's Procedure that I underwent January 2017. I was not prepared for just how poorly I would feel afterwards, it was an extremely tough time. I don't remember too much but I do remember the pain I was in straight after the operation and my family telling me all about how they couldn't believe how many lines and drains I had, which stayed in for at least a couple of weeks. Since then I've lost a total of 2 and a half stone, this was due to it taking months for me to get my food intake back on track. I used to eat one mouthful and be full, I couldn't see the discomfort and pain getting any better but it does. Even now, over a year down the line, I don't feel as good as I felt before that operation.

- VHL UK/Ireland website - Sep 8, 2019

With me, the cancer was close to the major blood vessels in the kidney. […] I spent hours every day for a week after my surgery praying to God to stop the pain. I cried. I'm not ashamed to say it. I'm a teenage boy and I cried in front of my Aunt and Uncle, in front of nurses, and even in front of one of the most beautiful doctors I have ever seen in my life. […] I spent hours praying, but there was nothing. I even pleaded to God to "take me away". Being in so much pain that you are literally praying for death is not a place I ever want to be in again.[…]

There's something about pain. It's a strange thing. It can make you want to throw your whole life away just for relief. […] That was the worst pain I had ever felt. That's what I thought, until December 4th 2015. That was the day my incredible Mam, the woman who raised me and my two big brothers by herself while fighting this awful disease, passed away.

- VHL UK/Ireland website - Sep 8, 2019

[redacted]

[redacted]

[redacted]

Epidemiology and aetiology

Von Hippel-Lindau disease (VHL) is caused by a mutation (fault) in the VHL gene. Based on UK clinical expert feedback, ≈80% of patients with VHL disease will have inherited disease and be known to genetics services and so will be identified reasonably early. Feedback from clinical experts suggest ≈20% of patients will present with a *de novo* VHL mutation (i.e. no family history and so not already known to genetics services) (3, 13).

VHL is rare with a prevalence lower than 1:50,000 in England. A national audit of VHL disease in the UK published in 2022 found 842 individuals had a clinical and/or molecular diagnosis of the disease in the 22 UK regional genetics centres surveyed over the 2012-2017 audit period, and estimated that the prevalence of VHL disease in the UK is likely to be between 1 in 91,111 to 1 in 68,493 with approximately 842 patients in the UK (13, 31). Taking into account the specific group of patients with VHL disease that belzutifan is indicated for (as described in Table 2), approximately 69 people in England and Wales are eligible for treatment with belzutifan (details provided in the company budget impact analysis submission document).

VHL disease affects males and females and all ethnic groups equally (32). In patients with VHL disease in the UK, the mean age at detection of detection of relevant tumour types is 39.4 years (SD ± 12.7) for RCC and 37.04 years (9-66 years) for CNS hemangioblastomas, the earliest recorded age at diagnosis for a pNET was 18 years (13).

Treatment pathway

Current clinical practice

Patients are diagnosed with VHL disease in two ways, either they are identified through genetic counselling due to having family member already diagnosed with VHL disease (this is the case in 80% of patients), or de novo when they are tested

for VHL disease upon presentation of a tumour (this is the case in 20% of patients) (13).

Since it is impossible to predict exactly how and when the disease will present for each person, it is necessary to check regularly for possible VHL manifestations throughout a person's lifetime. All diagnosed patients receive active surveillance usually led by the clinical genetics service, though sometime by the endocrinology service (13). Most of the surveillance clinics led by a clinical genetics service include other specialties such as ophthalmology, endocrinology, urology paediatrics, neurology, nephrology, and radiology (13). The surveillance recommendation for RCC and pNETs is usually "magnetic resonance imaging (MRI) or ultrasound examinations of the abdomen every 12 months, beginning from the age of 16 years" (13). The surveillance recommendation for CNS hemangioblastomas is usually "MRI scans of the head for every 12–36 months, beginning in adolescence". The first CNS MRI scan is generally performed at 14 to 16 years of age and for those centres that performed regular MRI scans the most common interval was 36 months for both brain and spine (13).

Care of VHL patients is fragmented and inconsistent across the UK. VHL is not listed in the Highly Specialised Commissioning Policy therefore centres do not receive funding for VHL clinics. Most patients are managed by via a genetic service and received regular and frequent scans as part of "active surveillance". We are aware of some MDTs managed VHL experts with a more joined up set of monitoring and appointments, but in other centres, once multiple manifestations in different parts of the body occur there can be a lack of coordination (e.g. a patient may have a MRI on their abdomen and then require three separate appointment with consultants to discuss tumours on the kidney, pancreas, and spleen individually).

Surveillance of diagnosed patients will identify the point at which VHL-associated tumours need treatment; following scans patients will either remain under active surveillance if their tumour(s) have not reached the threshold at which treatment is required or receive treatment if the threshold is reached (the relevant thresholds for these are described in Table 1).

Treatment of VHL in the UK is variable. The objective of care in patients with VHL disease-associated tumours in the UK is to prevent tumour growth within an organ to prevent permanent damage and prevent tumours from metastasising while maintaining function of affected organs, minimising symptom burden, and preserving patients' quality of life, a delicate balancing act. Clinicians treat patients as well as they can, this might include high-risk procedures that are not desirable and/or that are known to have very serious sequelae.

In patients with pNETs, a somatstatin analogue is used as 1st line therapy to treat these tumours when it reaches 1 cm in diameter, surgery is considered when these tumours grow to 2 cm in diameter and are progressing or likely to metastasise, this is the case regardless of whether the pNET is associated with VHL disease (3, 4). In patients for whom surgery is appropriate, these may sometimes be delayed in patients with VHL-associated tumours in more than one location, e.g. surgery for one tumour may be delayed to prioritise or recover from surgery on a more urgent tumour elsewhere.

Once a patient has undergone localised therapy to resect a tumour that had reached the threshold for treatment, the cycle of active surveillance resumes until the treatment decision point relevant to this appraisal whereby an intervention is needed (i.e. "require therapy" as stated in the GB marketing authorisation wording) but localised procedures are unsuitable or undesirable. In patients for whom surgery is not suitable or desirable the standard of care (SoC) at this point is a highly varied sequence of interventions and not a single treatment strategy (for either the entire VHL disease population or for any tumour site-defined subgroup considered separately) that can appropriately be described as the "best alternative care". The treatment aim of all modalities of current management is to preserve organ function and prevent tumours becoming advanced or metastatic, while maintaining patient quality of life.

In patients for whom surgery is not suitable or desirable the standard of care (SoC) at this point is a highly varied sequence of interventions and not a single treatment strategy (for either the entire VHL disease population or for any tumour site-defined subgroup considered separately) that can appropriately be described as the "best

alternative care". The treatment aim of all modalities of current management is to preserve organ function and prevent tumours becoming advanced or metastatic, while maintaining patient quality of life.

In some patients whose VHL disease-associated RCC, CNS hemangioblastoma, or pNET have grown to an extent where localised procedures would otherwise be used in current SoC, available localised procedures may no longer be suitable nor desirable due to an elevated risk of loss of organ function or adverse effects of the procedure. The circumstances that would make localised procedures unsuitable or undesirable for the tumour(s) in question, or the patient as a whole, are manifold and include (but are not limited to):

- In VHL-associated RCC, when the localised procedure would render the patient renal replacement therapy-dependent.
- In VHL-associated pNET, when the localised procedure would lead to loss of pancreatic function leading to lifelong pancreatogenic diabetes and being immune-compromised such that the patient will require lifelong insulin therapy, antibiotic therapy and/or pancreatic enzyme insufficiency impacting digestion.
- In VHL-associated CNS hemangioblastoma, when the localised procedure could lead to severe neurological or neuromuscular deficits equating to severe permanent disability. This most often arises with tumours located in the brainstem where they are difficult to access or operate on without damaging important nearby tissues, potentially leading to significant morbidity and death.

Patients with VHL RCC or pNETs for whom localised procedures are unsuitable or undesirable may still have a localised procedure. However, the localised procedure would not preserve organ function and result in significantly burdensome sequelae. In current UK clinical practice, patients undergo surgery where organ function will be significantly impaired or completely cease, or where there will be significant risk to neurological function for CNS lesions, as they are the only treatment option available to keep patients alive, or prevent symptomatic disease progressing to the point where the severe sequelae of such procedures are on-balance preferable, or prevent

the patient developing advanced or metastatic disease. Such procedures in no way constitute satisfactory treatment options.

- For RCC tumours, the localised procedures that are no longer capable of preserving organ function are radical (i.e. full) bilateral nephrectomies.

- For pNETs, such localised procedures that are no longer capable of preserving organ function are Whipple procedures/ pancreatectomies and splenectomies (A Whipple surgery involves removal of the head of the pancreas, a part of the duodenum, some of the bile duct and the gall bladder. The stomach, bile duct and remainder of the pancreas will then be rejoined to the small bowel. The operation usually takes 4-6 hours. Long-term effects of this procedure include but are not limited to diabetes, pancreatic insufficiency, and change in bowel habit).

- For CNS hemangioblastomas, where surgery would result in significant neuro-functional loss and/or high risk of mortality, including those close to the brain stem and in the spinal cord.

There are patients in whom tumour resection surgeries are contraindicated. This may be due to characteristics of the patient such as frailty, comorbidities or characteristics of the tumour such as CNS hemangioblastoma located at a physically inaccessible site. Currently these patients can receive symptom management, until their tumours/disease have progressed to the point where first-line systemic anti-cancer therapies (SACT) for unresectable or advanced cancer are used, or palliation.

This stage of disease is downstream of the position of belzutifan in the treatment pathway as specified in its GB marketing authorisation i.e., where patients can end up if treatment with belzutifan or current management fails, and so such SACT are not relevant comparators to belzutifan. The SACT which are not relevant comparators to belzutifan in this indication include monotherapy or combination therapy with immunotherapies or kinase inhibitors for RCC, and monotherapy with lutetium (177Lu) oxodotreotide or combination therapy with everolimus and sunitinib for pNETs.

The place of belzutifan

There are currently no satisfactory existing treatment options for the patients in the position described above. Patients who are suitable for surgery should have surgery instead of belzutifan. Surgery is an effective option for such patients and belzutifan should not be considered as a treatment option for them.

The place of belzutifan in the clinical pathway is therefore to give patients for whom localised organ-sparing procedures are unsuitable or undesirable (as described previously) an alternative to surgeries where organ function will be significantly impaired or lost, or where there will be significant risk to neurological function for CNS lesions. Patients must have sufficient organ function (as described in the MK-6482-004 study participant eligibility criteria described in section B.2.3 later in this document) to be eligible to receive belzutifan. The objective of the treatment to prevent or reverse symptomatic disease progression thereby offering significant additional benefit over existing treatment options.

MSD were disappointed and surprised that the NICE Topic Selection Oversight Panel (TSOP) made the decision not to route this indication into the Highly Specialised Technologies (HST) programme. We disagree with the decision. However, in order to facilitate access for patients with VHL we are moving ahead the STA process.

B.1.4 Equality considerations

No equality considerations are anticipated. It should be noted that the onset of RCC, pNETs, and/or CNS hemangioblastomas can affect patients with VHL disease when they are very young. VHL disease is an inheritable genetic disease and so some families are disproportionately affected. There are inequities in the type of service available to VHL patients depending on which centre leads their care.

B.2 Clinical effectiveness

Summary of key clinical effectiveness information

Clinical trial:

- MK-6482-004 is a Phase 2, open-label, single-arm trial, that investigated the efficacy and safety of orally administered belzutifan, in patients with RCC associated with VHL disease. The primary end point was objective response (complete or partial response) in RCC tumours. Objective response to belzutifan in patients with non–renal cell carcinoma neoplasms was also assessed.

- Belzutifan provided a clinically meaningful overall response rate (ORR) in patients with VHL disease-associated RCC of 63.9% (95% CI: 50.6%, 75.8%).

- Among the subgroup of 50 participants with both RCC and CNS hemangioblastoma, belzutifan provided a clinically meaningful confirmed ORR of 44.0% (95% CI: 30.0%, 58.7%) in the CNS hemangioblastomas.

- Among the subgroup of 22 participants with both RCC and pNETs, belzutifan provided a clinically meaningful confirmed ORR of 90.9% (95% CI: 70.8%, 98.9%) in pNETs.

- Belzutifan provided a disease control rate (DCR, i.e. CR + PR + SD) in patients with VHL disease-associated RCC of 98.4% (95% CI: 91.2%, 100%).

- Among the subgroup of 50 participants with CNS hemangioblastoma, belzutifan provided a DCR of 90.0% (95% CI: 78.2%, 96.7%) in these tumours.

- Among the subgroup of 22 participants with pNETs, belzutifan provided a DCR of 100% (95% CI: 84.6%, 100%) in these tumours.

- Responses to belzutifan treatment were long and durable as shown by a median duration of response (DOR) not reached after a median follow-up of 37.7 months (range: 4.2 to 46.1 months). The median time-to-response (TTR)

was 11.1 months (range 2.7 to 30.5 months) among 39 patients with a confirmed best overall response of complete response (CR) or partial response (PR).

- Later data cuts for RCC tumour response indicate patients response improves over time. The number and proportion of patients with a best overall response of complete response (CR) or partial response (PR), increases in later data cuts. ORR increasing from 36.1% (95% CI: 24.2%, 49.4%) at the 01-JUN-2020 data cut-off date to 63.9% (95% CI: 50.6%, 75.8%) at 01-APR-2022 data cut-off date.

Clinical safety

- Belzutifan was generally well tolerated. Adverse events (AEs) leading to treatment discontinuations were reported in 4 (6.6%) participants.

- Belzutifan had a manageable safety profile. Most AEs were Grade 1 to 2.

Relative treatment effect – real world study:

- Relative effectiveness information versus UK standard of care was derived from a retrospective, non-interventional study (VHL Natural History Study). The Optum study (34) is used to align effectiveness data with real world UK SoC for the purposes of the cost-effectiveness analysis.

Health-related quality of life

- Health-related quality of life data were from a VHL patient survey, a non-interventional international, cross-sectional patient survey that collected European Quality of Life – 5 Dimension Survey (EQ-5D-5L) data.

B.2.1 Identification and selection of relevant studies

To identify and select relevant studies, a systematic literature review (SLR) search was carried out in accordance with NICE guidance, according to a previously prepared protocol to identify relevant studies that investigated pembrolizumab and

any relevant comparator treatments for this the indication of interest for this appraisal. Please refer to Appendix D for full details of the process and methods undertaken.

B.2.2 List of relevant clinical effectiveness evidence

A SLR was performed to identify all relevant published and unpublished randomised controlled trials (RCTs) and non-randomised clinical trials (non-RCTs) relating to belzutifan as per the final scope.

A single trial was identified from the SLR that provided clinical effectiveness information on belzutifan in the patient population of relevance to this submission. At the time of the SLR search, evidence from the MK-6482-004 study was available from a peer-reviewed journal publication (35), an European Society for Medical Oncology (ESMO) presentation (36), as well as from internal MSD data (37).

Table 4 Clinical effectiveness evidence

Study	An Open-Label Phase 2 Study to Evaluate PT2977 for the Treatment of Von Hippel Lindau Disease-Associated Renal Cell Carcinoma (MK-6482-004)
Study design	Phase 2, open-label, multicentre, single-arm, interventional study
Population	Patients with VHL disease who have at least one measurable RCC tumour
Intervention(s)	Belzutifan
Comparator(s)	None
Indicate if study supports application for marketing authorisation	Yes (GB marketing authorisation for belzutifan in this indication has already been received from the MHRA (5))
Indicate if study used in the economic model	Yes
Rationale if study not used in model	N/A. This study as well as data from other sources (the VHL Natural History Study, Optum study, and VHL patient survey) are used to supplement the cost-effectiveness analysis.
Reported outcomes specified in the decision problem	• Progression-free survival • Response rates • Adverse effects of treatment
All other reported outcomes	• Duration of response • Time to response • Time to surgery
Data cut-off dates of results existent for the study	• 01-JUN-2020 (an unpublished ad-hoc analysis) • 01-DEC-2020 (35) • 15-JUL-2021 (an unpublished ad-hoc analysis)

Study	An Open-Label Phase 2 Study to Evaluate PT2977 for the Treatment of Von Hippel Lindau Disease-Associated Renal Cell Carcinoma (MK-6482-004)
	• 01-APR-2022 (36, 37), the key set of results presented in this submission

B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

Summary of the methodology of the MK-6482-004 study

The methodology of the MK-6482-004 study is summarised in Table 5 and described in detail in the following subsections.

The MK-6482-004 study enrolled patients with VHL-associated RCC. While some patients in the study also had VHL-associated CNS hemangioblastomas and/or VHL-associated pNETs, all patients had VHL-associated RCC. This therefore means that the population of the MK-6482-004 study does not align with (i.e. is narrower than, in this respect) the marketing authorisation for belzutifan as described previously in section B.1), the population under consideration in this assessment.

The MHRA marketing authorisation decision and wording were based on the evidence from the MK-6482-004 study, which despite it being limited to patients with VHL disease who must have at least one RCC tumour, the MHRA found that the effect of treatment with belzutifan on the other VHL disease-associated tumours (in particular CNS hemangioblastomas and pNETs) that the patients with VHL-associated RCC achieved were sufficiently impressive, and the unmet need in patients with VHL disease sufficiently high, to warrant the marketing authorisation extending to patients with CNS hemangioblastomas and/or pNETs without RCC.

There is also misalignment between the GB marketing authorisation and the MK-6482-004 study in terms of:

- The marketing authorisation is for patients "who require therapy" for the VHL-associated tumours, whereas the MK-6482-004 study's participant eligibility criteria specified that patients who had an immediate need for surgical intervention for tumour treatment were excluded.

- The marketing authorisation is for patients "for whom localised procedures are unsuitable or undesirable" whereas this was not part of the participant eligibility criteria for the MK-6482-004 study.

The Medicines and Healthcare products Regulatory Agency (MHRA) granted this marketing authorisation in this indication based on the MK-6482-004 study and its results, based on which patients would benefit from treatment with belzutifan based on what was demonstrated in this very study. Such misalignments are not unusual for rare and highly specialised indications such as this one.

Table 5 Summary of trial methodology

Study name	MK-6482-004
Trial design	Open-label, multicentre, single-arm, non-randomised, interventional, Phase 2 study.
Eligibility criteria for participants	Patients with renal cell carcinoma associated with VHL disease.
Settings and locations where the data were collected	Patients were enrolled at 11 centres in the United States, Denmark, France, and the United Kingdom.
Trial drugs	Belzutifan
Primary outcomes	• Overall response rate (ORR)
Other outcomes used in the economic model/specified in the scope	• Duration of response (DOR) • Time to response (TTR) • Progression-free survival (PFS) • Time to surgery (TTS) • Adverse events
Pre-planned subgroups	Primary tumours other than RCC at screening.
Data cut-off dates of results existent for the study	• 01-JUN-2020 (an unpublished ad-hoc analysis) • 01-DEC-2020 (35) • 15-JUL-2021 (an unpublished ad-hoc analysis) • 01-APR-2022 (36), presented in this submission

Trial design

The MK-6482-004 open-label, multicentre, single-arm, non-randomised, interventional, Phase 2 study evaluated the efficacy and safety of belzutifan as treatment for participants with VHL disease who had at least 1 measurable RCC tumour. Participants may have also had VHL disease-associated tumours in other organ systems at screening which could have been measurable and/or non-measurable lesions. Patients were enrolled at 11 centres in the United States,

Denmark, France, and the United Kingdom between May 31, 2018, and March 29, 2019. Patients received belzutifan administered orally at a dose of 120 mg once daily (in three 40-mg tablets) unless unacceptable adverse events or disease progression occurred.

The primary end point was objective response to treatment with belzutifan (complete response or partial response), as defined according to RECIST, version 1.1 (described in Appendix N), in patients with VHL-disease–associated renal cell carcinoma. Participants were evaluated with imaging every 12 weeks. Measurements of VHL-associated RCC target lesions at 2 timepoints prior to the screening imaging were collected to determine the tumour growth rate before treatment with belzutifan. All images obtained were submitted to an independent review committee (IRC) to assess objective response and progression-free survival (PFS) using Response Evaluation Criteria in Solid Tumors (RECIST) 1.1. Radiographic response assessments were made separately for each VHL-associated organ system using RECIST 1.1 (e.g. RCC, pancreas lesions, and CNS hemangioblastomas). Retinal angiomas were assessed by an independent central committee of ophthalmologists, reviewing fundoscopic images.

Assignment, randomisation, and blinding

The MK-6482-004 study was an open-label single-group trial and so had no assignment, randomisation, or blinding.

Eligibility criteria

Patient inclusion criteria

Male and female participants of at least 18 years of age were eligible for enrolment in this study. Key inclusion criteria were as follows:

- Had a diagnosis of VHL disease based on a germline VHL alteration.
- Had at least 1 measurable solid RCC tumour and no RCC tumour greater than 3.0 cm that requires immediate surgical intervention. The diagnosis of RCC could be radiologic (histologic diagnosis not required). Participants could have VHL disease-associated tumours in other organ systems.

- Had an ECOG performance status of 0 to 1.
- Had adequate organ function as defined below:
    - Absolute neutrophil count ≥1,000/μL, haemoglobin level ≥10 g/dL and platelet count ≥100,000/μL without transfusion or growth factor support within 2 weeks prior to obtaining the haematology values at screening.
    - Serum creatinine level ≤2.0 x upper limit of normal (ULN).
    - AST and ALT <2.5 x ULN, total bilirubin <1.5 x ULN (<3 x ULN in patients with Gilbert's disease), and alkaline phosphatase ≤2.5 x ULN.

Patient exclusion criteria

Participants were excluded from the study if they met any of the following criteria:

- Had any systemic anticancer therapy (included anti-VEGF therapy or a systemic investigational anticancer agent).
- Had a surgical procedure for VHL disease or any major surgical procedure completed within 4 weeks prior to study enrolment.
- Had received prior treatment with PT2977 or another HIF-2α inhibitor.
- Had radiotherapy within 4 weeks prior to study enrolment.
- Had an immediate need for surgical intervention for tumour treatment.
- Had evidence of metastatic disease on screening imaging.
- Had malabsorption due to prior GI surgery or GI disease.
- Had any major cardiovascular event within 6 months prior to study drug administration including but not limited to myocardial infarction, unstable angina, cerebrovascular accident, transient ischemic event, pulmonary embolism, clinically significant ventricular arrhythmias (e.g., sustained ventricular tachycardia, ventricular fibrillation, torsades de pointes) or New York Heart Association Class III or IV heart failure.

Settings and locations where the data were collected

Patients were enrolled at 11 centres in the United States, Denmark, France, and the United Kingdom. One patient received treatment in the UK.

Trial drugs and concomitant medications

Trial treatment

Use of belzutifan as specified in MK-6482-004 is presented in Table 6.

Table 6 MK-6482-004 study intervention

Treatment:	MK-6482 (belzutifan)
Regimen:	120 mg QD (3 x 40 mg tablets)
Route of administration:	Oral
Duration of treatment:	Until unacceptable treatment-related toxicity or unequivocal disease progression
Use in study:	Experimental

QD: once daily

Concomitant medications

The protocol specified that patients may not receive any approved or any additional investigational anti-neoplastic agent during the course of this study.

Pre-treatment administration of prophylactic anti-emetics was not allowed, but may be given if needed during study treatment. There will be no constraint on the use of growth factors, including erythropoietin, during treatment; however, prophylactic use is discouraged and adherence to the American Society of Clinical Oncology (ASCO) or European Society for Medical Oncology (ESMO) guidelines is recommended.

Patients should receive all necessary supportive care, including blood products, transfusions, antibiotics, pain medications, bisphosphonates, and replacement hormonal therapies (insulin, thyroid hormones, oestrogen/progesterone).

Belzutifan has been shown to induce the enzymes CYP3A4 at concentrations about 2-times the expected plasma concentration of PT2977 at the clinical dose of 120 mg/day. Belzutifan may decrease the exposure of medications metabolized by CYP3A4. If co-administration with drugs of a narrow therapeutic index that are metabolized by CYP3A4 cannot be avoided, additional monitoring of drug effect is recommended i.e. if a patient absolutely must receive a concomitant medication that

is metabolised by the CYP3A4 enzyme, such a patient a patient would be monitored more frequently to check if the drug is working.

All concomitant medication(s) used during the study and within 28 days before the start of study drug administration were reported.

Outcomes assessed

A summary of the objectives and associated endpoints/outcomes assessed in the MK-6482-004 study is shown in Table 7.

Table 7 Objectives and associated endpoints of the MK-6482-004 study

Primary Objective	**Primary Endpoint**	**Relevant VHL disease-associated tumour(s) for the objective and endpoint**
To evaluate the efficacy of MK-6482 (belzutifan) for the treatment of VHL disease-associated RCC as measured by overall response rate (ORR) per RECIST 1.1 (described in Appendix N).	ORR: the proportion of participants who have achieved a complete response (CR) or partial response (PR)	RCC
Secondary Objectives	**Secondary Endpoints**	
To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by duration of response (DOR) per RECIST 1.1.	DOR: the time from first documented evidence of CR or PR until either disease progression or death due to any cause, whichever occurs first	RCC
To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by time to response rate (TTR) per RECIST 1.1.	TTR: the time from the start of study intervention to the first documentation of a response, calculated for participants with a best confirmed response of CR or PR	RCC
To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by progression-free survival (PFS) per RECIST 1.1.	PFS: the time from the start of study intervention to the first documented disease progression or death due to any cause, whichever occurs first.	RCC

To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by time to surgery (TTS).	TTS: the time from the start of study intervention to the date of surgery	RCC
To evaluate efficacy of MK-6482 for the treatment of VHL disease-associated non-RCC tumours (retinal and central nervous system [CNS] hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)*	ORR, DOR, TTR, PFS, and TTS	Non-RCC tumours (retinal and central nervous system [CNS] hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)*
To evaluate the safety and tolerability of MK- 6482 for the treatment of VHL disease-associated RCC	Adverse events (AEs) and study intervention discontinuation due to AEs.	All patients
To assess the pharmacokinetics (PK) of MK- 6482	Plasma concentrations of MK-6482 and its metabolite(s)	All patients

*Subgroup data on tumours other than VHL-associated RCC, CNS hemangioblastomas, or pNETs are presented in detail in this document as they are not included in the GB marketing authorisation for belzutifan in this indication or the scope of this appraisal.

Pre-planned subgroups

Summaries of the response endpoints were planned to be provided for subgroups of the Efficacy Analysis Set defined based on site of primary tumours other than RCC at screening. The analyses of ORR, DOR, TTR, PFS, and TTS were planned to be performed in each of these subgroups

Analysis populations

The data sets analysed as part of the study are described in Table 8.

Table 8 Analysis populations

Data set	**Population**	**n at 01-APR-2022 data cut-off date**
Efficacy Analysis Set	The All Participants as Treated (APaT) population will be used for the analyses of efficacy. The APaT population consists of all allocated patients who received at least one dose of belzutifan.	61
Safety Analysis Set	The APaT population will be used for the analysis of safety data in this study. The APaT population consists of all allocated patients who received at least one dose of belzutifan.	61

Pharmacokinetic Analysis Set	Pharmacokinetic Analysis Set will include all patients who received at least 1 dose of belzutifan and have at least one post-dose pharmacokinetic sample collection.	61
Pharmacodynamic Analysis Set	Pharmacodynamic Analysis Set will include all patients who received at least one dose of study drug and have evaluable pharmacodynamics data above the limit of quantification.	61

Baseline characteristics of trial participants

The baseline characteristics of the participants of the MK-6482-004 study are shown in Table 9 to Table 176. All patients included in the study had at least one concurrent non-RCC tumour at baseline. Median time from original diagnosis of VHL-associated RCC to initiation of treatment with belzutifan was 77.60 months (6.5 years). Participants had a mean (SD) of 5.5. (3.34) prior surgeries. Based on investigators assessment, the common concurrent non-RCC tumour types were CNS hemangioblastoma (n=51, 83.6%), pancreatic lesions (n=32, 52.5%), and retinal hemangioblastomas (n=17, 27.9%). The median time from last surgery to initiation of belzutifan was 23.49 months (1.96 years).

The baseline characteristics of this study have been shown to UK clinical experts who treat patients with VHL disease, who broadly agreed that these were representative of/applicable to the patients in the UK who would be treated with belzutifan in accordance with the marketing authorisation.

All participants

Table 9 MK-6482-004 study demographic and baseline characteristics (safety analysis set) - all patients

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	32		29		61	
Age (Years)						
Participants with data	32		29		61	
Mean	38.8		43.3		41.0	
SD	12.7		14.1		13.5	
Median	36.0		44.0		41.0	
Range	22.0 to 65.0		19.0 to 66.0		19.0 to 66.0	
Ethnicity						
Hispanic or Latino	3	(9.4)	3	(10.3)	6	(9.8)

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Not Hispanic or Latino	28	(87.5)	26	(89.7)	54	(88.5)
Unknown	1	(3.1)	0	(0.0)	1	(1.6)
Race						
Asian	1	(3.1)	0	(0.0)	1	(1.6)
Black or African American	1	(3.1)	1	(3.4)	2	(3.3)
Native Hawaiian or Other Pacific Islander	1	(3.1)	0	(0.0)	1	(1.6)
White	28	(87.5)	27	(93.1)	55	(90.2)
Unknown	1	(3.1)	1	(3.4)	2	(3.3)
Weight (kg)						
Participants with data	32		29		61	
Mean	86.7		72.1		79.7	
SD	21.4		23.4		23.4	
Median	81.5		65.0		74.4	
Range	63.0 to 147.6		47.7 to 147.0		47.7 to 147.6	
Height (cm)						
Participants with data	32		27		59	
Mean	176.6		161.1		169.5	
SD	8.7		6.7		11.0	
Median	175.5		160.1		169.0	
Range	159.5 to 195.0		148.0 to 174.0		148.0 to 195.0	
BMI (kg/m2)						
Participants with data	32		27		59	
Mean	27.7		27.8		27.8	
SD	6.0		8.8		7.4	
Median	27.0		24.5		26.3	
Range	18.4 to 42.7		17.2 to 52.0		17.2 to 52.0	
ECOG Performance Status						
0	24	(75.0)	26	(89.7)	50	(82.0)
1	8	(25.0)	2	(6.9)	10	(16.4)
2	0	(0.0)	1	(3.4)	1	(1.6)

Database Cutoff Date: 01APR2022
Number of participants: Safety Population
Note: Baseline is defined as the last available measurement prior to the first dose administered.

Table 10 MK-6482-004 study demographic and baseline characteristics (safety analysis set) - all patients – additional data

	Belzutifan (N=61)
Age at time of VHL diagnosis (years)	
N	61
Mean	31.3 (14.29)
Median	32.0

	Belzutifan (N=61)
Min, Max	4, 66
VHL Subtype, n (%)	
Type 1	51 (83.6)
Type 2A	2 (3.3)
Type 2B	6 (9.8)
Type 2C	0
Missing	2 (3.3)
VHL-associated Non-RCC tumours, n (%)	
Pancreatic Lesions	32 (52.5)
- Pancreatic lesions of which were pNETs	22 (36.1)
Adrenal Lesions (Pheochromocytomas)	3 (4.9)
CNS Hemangioblastoma [3]	51 (83.6)
Endolymphatic Sac Tumours	1 (1.6)
Epididymal Cystadenomas	10 (16.4)
Retinal Lesions	17 (27.9)
Other	2 (3.3)
Time from Original Diagnosis of VHL associated RCC to First Dose (months) [1]	
n	45
Mean (SD)	103.43 (96.231)
Median	77.60
Q1, Q3	24.54, 136.97
Min, Max	0.5, 389.4
Time from Last Surgery to First Dose (months)	
n	59
Mean (SD)	37.01 (38.493)
Median	23.49
Q1, Q3	9.66, 41.13
Min, Max	0.6, 137.6
Number of Prior Surgeries per Subject	
n	59
Mean (SD)	5.5 (3.34)
Median	5.0
Min, Max	1, 15
Age at time of VHL associated RCC diagnosis (years) [2]	
n	45
Mean (SD)	33.8 (13.06)
Median	32.0
Min, Max	15, 62
Histology, n (%)	
Renal Cell Carcinoma of Clear Cell Subtype	43 (70.5)
Other	2 (3.3)
Not Done	16 (26.2)

	Belzutifan (N=61)
Histological Grade	
GX - Grade cannot be assessed	2 (3.3)
G1 - Nucleoli absent or inconspicuous and basophilic at 400x magnification	10 (16.4)
G2 - Nucleoli conspicuous and eosinophilic at 400x magnification, visible but not prominent at 100x magnification	23 (37.7)
G3 - Nucleoli conspicuous and eosinophilic at 100x magnification	8 (13.1)
G4 - Marked nuclear pleomorphism and/or multinucleate giant cells and/or rhabdoid and/or sarcomatoid differentiation	0
Missing	2 (3.3)
TNM Stage T	
TX	1 (1.6)
T0	0
T1	5 (8.2)
T1a*	48 (78.7)
T1b	2 (3.3)
T2	0
T2a	0
T2b	1 (1.6)
T3	0
T3a	0
T3b	0
T3c	0
T4	0
TNM Stage N	
NX	NX 13 (21.3)
N0	N0 46 (75.4)
N1	N1 0
TNM Stage M	
cM0	59 (96.7)
cM1	0
pM1	0

[1] (First Dose Date-Date of first positive biopsy+1)/30.4375
[2] (Date of VHL associated RCC diagnosis-Birthdate+1)/365.25
[3] The number patients with CNS hemangioblastomas shown in this table is according to investigator assessment, study results are reported later on in this document in terms of the number of patients with CNS hemangioblastoma according to independent review committee determination where this was found to be n=50.
*T1a means that the tumour is less than 4cm across, and is completely inside the kidney.
Date of Data Cut-off: 01APR2022

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

Statistical analysis and definition of study groups in the MK-6482-004 study

Key information on the statistical analysis and definition of study groups in the MK-6482-004 study are summarised in Table 11 with details presented in Appendix M.

Table 11 Summary of MK-6482-004 study statistical methods

Study design and overview	The MK-6482-004 study was a single-arm open-label Phase 2 study that evaluated the efficacy and safety of belzutifan in patients with VHL disease who have at least 1 measurable RCC tumour (as defined by RECIST 1.1).
Treatment assignment and stratification	This was an open-label single-group trial and so had no assignment, randomisation, or stratification.
Study hypotheses	No formal hypothesis testing. For the purposes of sample size determination only, null hypotheses and alternative hypotheses were formulated (described later in the "Sample size and power" section of this table)

Study objectives	**Specific to VHL RCC tumours:**
	• Primary objective:
	  ○ To evaluate the efficacy of belzutifan for the treatment of von Hippel-Lindau (VHL) disease-associated renal cell carcinoma (RCC) as measured by overall response rate (ORR) per Response Evaluation Criteria in Solid Tumours Version 1.1 (RECIST 1.1)
	• Secondary objectives:
	  ○ To evaluate efficacy of belzutifan for the treatment of VHL disease-associated RCC measured as follows:
	  ▪ Duration of response (DOR)
	  ▪ Time to response (TTR)
	  ▪ Progression-free survival (PFS)
	  ▪ Time to Surgery (TTS)
	
	Specific to VHL non-RCC tumours:
	• Secondary objectives:
	  ○ To evaluate efficacy of belzutifan for the treatment of VHL disease associated non-RCC tumours (retinal and CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)
	
	Applies to all patients in the study:
	• Secondary objectives:
	  ○ To evaluate safety and tolerability of belzutifan
	  ○ To assess the pharmacokinetics (PK) of belzutifan
	• Exploratory objective:
	  ○ To evaluate changes in pharmacodynamic markers (e.g., serum erythropoietin)

Analysis populations	• Efficacy Analysis Set: The All Participants as Treated (APaT) population will be used for the analyses of efficacy. The APaT population consists of all allocated patients who received at least one dose of belzutifan. • Safety Analysis Set: The APaT population will be used for the analysis of safety data in this study. The APaT population consists of all allocated patients who received at least one dose of belzutifan. • Pharmacokinetic Analysis Set: Pharmacokinetic Analysis Set will include all patients who received at least 1 dose of belzutifan and have at least one post-dose pharmacokinetic sample collection. • Pharmacodynamic Analysis Set: Pharmacodynamic Analysis Set will include all patients who received at least one dose of study drug and have evaluable pharmacodynamics data above the limit of quantification.
Primary endpoint	**Specific to VHL RCC tumours:** • Overall response rate (ORR) in VHL disease-associated RCC tumours, defined as proportion of patients with a best confirmed response of Complete Response (CR) or Partial Response (PR) as determined by RECIST 1.1
Key secondary endpoint	**Specific to VHL RCC tumours:** • Secondary endpoints: o Duration of response (DOR) in VHL disease-associated RCC tumours, defined as the interval from the first documentation of response, as determined by RECIST 1.1, to the earlier of the first documentation of disease progression or death from any cause, and calculated for patients with a best confirmed response of CR or PR. o Time to response (TTR) in VHL disease-associated RCC tumours, defined as the interval from the start of study treatment to the first documentation of a response, as determined by RECIST 1.1, and calculated for patients with a best confirmed response of CR or PR. o Progression-free survival (PFS) in VHL disease-associated RCC tumours, defined as the interval from the start of study treatment until the earlier of the first documentation of disease progression determined by RECIST 1.1 or death from any cause. o Time to surgery (TTS) for VHL disease-associated RCC tumours, defined as the interval from the start of study treatment to the date of surgery. **Specific to VHL non-RCC tumours:** • Secondary endpoints: o ORR, DOR, TTR, PFS, and TTS for non-RCC tumours associated with VHL disease in individual organ systems (retinal lesions, CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas). **Applies to all patients in the study:** • Exploratory endpoints:

	○ Changes in pharmacodynamic markers • Safety endpoints: ○ Physical examinations ○ Vital sign measurements (including pulse oximetry) ○ 12-lead electrocardiograms (ECG) with QTc interval determination ○ Clinical laboratory measurements ○ Concomitant medications ○ Incidence, intensity, and relationship of AEs and serious adverse events (SAEs) ○ Effects on fertility in males (semen analysis, and measurement of testosterone, follicle-stimulating hormone, luteinizing hormone, and inhibin B levels)
Sample size and power	This study will enrol approximately 50 patients. Even though no formal hypothesis testing will be performed for this study, the required sample size for this study is based on the following assumptions. The null hypothesis is that the ORR is 15% (P0 = 0.15). The alternative hypothesis is that the ORR is 30% (P1 = 0.3). A sample size of 50 patients will provide greater than 80% power to reject the null under the alternative hypothesis using a one-sided test at a 0.05 level of significance.
Interim and final analyses	Periodic review of the trial data will be performed. Any analysis for the study will only take place after all patients have had the opportunity to complete at least two imaging assessments on study or have discontinued study therapy by the time of analysis data cut-off. The final analyses for the study will utilize a data cut-off date which will be at least 36 weeks after enrolment of the last patient.
Data management, patient withdrawals	Patients who discontinue from study treatment would complete the safety follow-up and long-term follow-up assessments according to the Schedule of Events. During the safety follow-up visit the patient would be evaluated for continuation or resolution of any AEs/SAEs. Patients who discontinue study treatment for any reason would undergo long term follow-up every 6 months for up to 3 years following enrolment of last patient into the study.

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

The quality assessment of the MK-6482-004 study is provided in Appendix D section D1.3. The main limitation of the MK-6482-004 study is that a single-arm trial does not directly compare the effects of belzutifan to that of UK clinical practice in this disease. Consequently, the relative effectiveness of belzutifan to current UK clinical practice is assessed via a control arm from a US VHL registry, methods described in section B.2.9 and section B.3.

B.2.6 Clinical effectiveness results of the relevant studies

Clinical effectiveness data from the ongoing MK-6482-004 study at the 01-APR-2022 database cut-off date are presented in this section.

Patient disposition and follow-up duration

A total of 61 participants were allocated and received at least 1 dose of belzutifan. As of the 01-APR-2022 database cut-off date, 38 participants (62.3%) were receiving belzutifan, 23 participants (37.7%) had discontinued belzutifan and 6 participants 9.8%) had discontinued from the study (Table 12). The median duration of follow-up among the 61 participants with RCC in the safety analysis set was 37.7 months (range: 4.2 to 46.1 months) (Table 13).

Table 12 MK-6482-004 summary of patient disposition (safety analysis set)

	Belzutifan (N = 61) n (%)
• Treatment Ongoing at Data Cut-Off Date	38 (62.3)
• Discontinued Treatment	23 (37.7)
Reason for Treatment Discontinuation	
• Disease progression per RECIST 1.1 for VHL disease-associated RCC tumours	6 (9.8)
• Disease progression due to symptomatic deterioration of the patient's health status	0
• Adverse event that in the opinion of the investigator or medical monitor would lead to undue risk if study treatment were continued	2 (3.3)
• Study drug interruption for more than 3 consecutive weeks due to a grade 3-4 or intolerable toxicity that is attributed to study drug	0
• Gross noncompliance with protocol	0
• Pregnancy in a female patient during the study	1 (1.6)
• Death*	2 (3.3)
• Lost to follow-up	0
• Patient decision to discontinue study drug	11 (18.0)
• Sponsor discontinuation of study	0
• Other	1 (1.6)
• On Study at Data Cut-Off Date [1]	55 (90.2)
• Off Study	6 (9.8)
Reason for Study Discontinuation	
• Death	2 (3.3)
• Informed Consent Withdrawn	2 (3.3)
• Lost To Follow-up	0
• Sponsor discontinuation of study	0
• Other	2 (3.3)

	Belzutifan (N = 61) n (%)
Completed Safety Follow-up Visit	13 (21.3)
On Long Term Follow-up Period at Data Cut-off Date	10 (16.4)

[1] Patients are still on study treatment or in long term follow-up as of the cutoff date.
Date of Data Cut-off: 01APR2022
*The two deaths (suicide attempt and toxicity to various agents) were assessed a not drug-related by the investigator.

Table 13 MK-6482-004 summary of follow-up duration (safety analysis set)

Follow-up duration (months)	Belzutifan (N = 61)			
Date of data cut-off	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Median (Range)	[redacted]	21.8 (4.2-30.1)	[redacted]	37.7 (4.2-46.1)
Mean (SD)	[redacted]	22.4 (3.35)	[redacted]	38.1 (5.01)

Follow-up duration is defined as the time from first dose to the date of death or the database cut-off date if the subject is still alive.

Extent of exposure

Duration of exposure is summarised in Table 14. The median duration of exposure to belzutifan was [redacted] at the 01-APR-2022 database cut-off date.

Table 14 MK-6482-004 study drug exposure (safety analysis set)

	Belzutifan (N=61)			
Date of data cut-off*	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Number of patients exposed	[redacted]	61	[redacted]	[redacted]
Duration of exposure (weeks)				
N	[redacted]	61	[redacted]	[redacted]
Mean (SD)	[redacted]	92.77 (23.561)	[redacted]	[redacted]
Median	[redacted]	94.14	[redacted]	[redacted]
Min, Max	[redacted]	8.4, 130.9	[redacted]	[redacted]
Cumulative dose received (mg/subject)				
N	[redacted]	61	[redacted]	[redacted]
Mean (SD)	[redacted]	72937.7 (21453.74)	[redacted]	[redacted]
Median	[redacted]	77760.0	[redacted]	[redacted]
Min, Max	[redacted]	4680, 106680	[redacted]	[redacted]

Summary of MK-6482-004 study efficacy results

The efficacy results of the MK-6482-004 study at the 01-APR-2022 data cutoff date are summarised in Table 15 and are presented in greater detail in the subsections of this section.

Table 15 Summary of MK-6482-004 study efficacy results (01-APR-2022 data cutoff)

Outcome	Summary of results
RCC (all patients)	
Overall response rate (ORR)	63.9% (95% CI: 50.6%, 75.8%)
Disease control rate (DCR)	98.4% (95% CI: 91.2%, 100.0%)
Duration of response (DOR)	Median DOR not reached (range: 5.4+ to 35.8+ months)
Time to response (TTR)	Median TTR was 11.1 months (range: 2.7 to 30.5 months) among 39 participants with response
Progression-free survival (PFS)	[redacted]
Time to surgery (TTS)	Not evaluable
Subgroup of patients with CNS hemangioblastoma	
Overall response rate (ORR)	44.0% (95% CI: 30.0%, 58.7%)
Disease control rate (DCR)	90.0% (95% CI: 78.2%, 96.7%)
Duration of response (DOR)	Median DOR not reached (range: 3.7+ to 38.7+ months)
Time to response (TTR)	[redacted]
Progression-free survival (PFS)	[redacted]
Time to surgery (TTS)	Not evaluable
Subgroup of patients with pNET	
Overall response rate (ORR)	90.9% (95% CI: 70.8%, 98.9%)
Disease control rate (DCR)	100% (95% CI: 84.6%, 100.0%)
Duration of response (DOR)	Median DOR not reached (range: 11.0+ to 37.3+ months)
Time to response (TTR)	[redacted]
Progression-free survival (PFS)	[redacted]
Time to surgery (TTS)	Not evaluable

Date of Data Cut-off: 01APR2022

VHL RCC

Summary of key RCT clinical effectiveness results

In patients with VHL disease-associated RCC:

- MK-6482-004 study results from the latest available 01-APR-2022 cut-off date (not yet published) are presented in this section. The median duration of follow-up among the 61 participants with RCC in the APaT population was 37.7 months (range: 4.2 to 46.1 months). The median duration of exposure to belzutifan was [redacted].
- Belzutifan provides a clinically meaningful confirmed ORR among the 61 participants with RCC in the Efficacy Analysis Set (APaT population) of 63.9% (95% CI: 50.6%, 75.8%).
- Belzutifan provided a disease control rate (DCR, i.e. CR + PR + SD) in patients with VHL disease-associated RCC of 98.4% (95% CI: 91.2%, 100%).
- Responses to belzutifan treatment were long and durable as shown by a median DOR that was not reached as of the 01-APR-2022 database cut-off date. The range of DOR was 5.4+ to 35.8+ months.
- The median TTR was 11.1 months (range: 2.7 to 30.5 month) among 39 participants with a confirmed best observed response (BOR) of CR or PR.
- The median (95% CI) PFS was [redacted] months. The PFS rate at Month 36 was [redacted] and at Month 42 was [redacted].
- At the 01-APR-2022 database cut-off date, 7 patients (11.5%) had undergone surgery. Consequently, the median time to surgery is not evaluable.

Overall response rate

The confirmed ORR among the 61 participants with RCC in the Efficacy Analysis Set was 63.9% (95% CI: 50.6, 75.8), with a rate and associated lower 95% CI >50% (i.e. even at the lowest estimate of efficacy at least half of patients experience CR or PR),

this is demonstrative of the efficacy of belzutifan in treating these tumours, as such tumours do not shrink/respond spontaneously in the absence of effective treatment. Detailed response results are shown in Table 16 and Figure 3. Note that that best overall response according to RECIST 1.1 criteria (described in more detail in Appendix N) is determined by more than only the change in tumour size between baseline and last measurement (38). Therefore it is possible for, e.g. the best overall response [BOR] of a tumour to be SD even though it may have a greater percentage reduction in diameter than a different tumour with a BOR of PR, as shown in certain patients in the figure.

Tumour response results from several data cut-off dates are shown in Table 16, it can be seen that the number and proportion of patients with complete response (CR), and with partial response (PR), increases with later data cut-off dates, indicating that more patients experienced a better response as time went on. These are further illustrated in Appendix N.

Of the four patients who experienced a complete response in their target RCC tumour by the 01-APR-2022 data cut-off date, their target RCC tumour [redacted] from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with belzutifan persist. For the 35 patients who had experienced a partial response by the 01-APR-2022 data cut-off date, the change in their target RCC tumour size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date are shown in Figure 5. [redacted]

No patients had a best overall response of PD (shown in Table 16) despite there being 6 patients who discontinued treatment due to disease progression per RECIST 1.1 for VHL disease-associated RCC tumours (shown in Table 12), and 9 patients are shown to have disease progression in the analysis of PFS (discussed later in this document and shown in Table 19). This because *best overall response* is the best response status recorded in the target tumour in the patient at any point during their

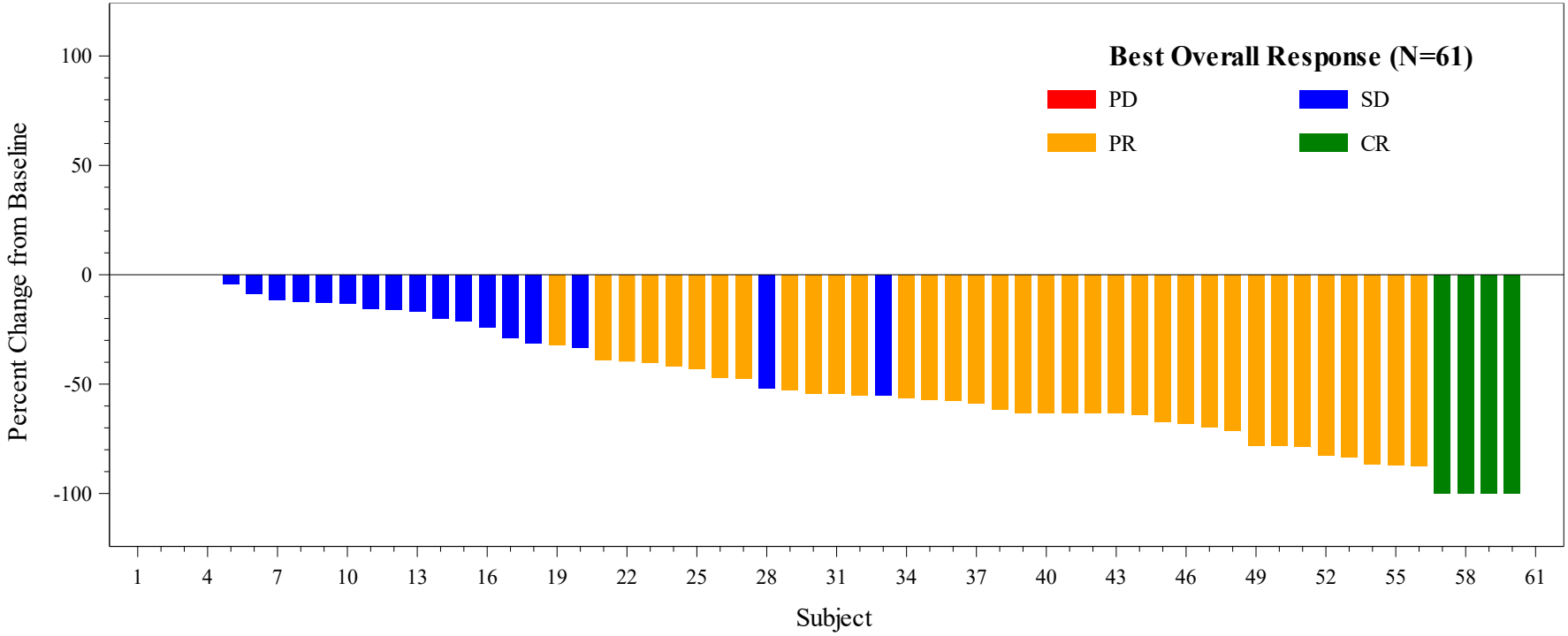
follow-up period i.e. if a patient had an overall tumour response state of PR or SD at any point prior to disease progression their best overall response would be PR or SD (whichever was the better one had). The only way a patient could have a best overall response of PD is if they were recorded to have PD at their first post-baseline follow-up scan and then discontinued the trial or never subsequently experienced a CR, PR, or SD.

Table 16 MK-6482-004 summary of best overall tumour response for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Best Overall Response, n (%)				
• Complete Response (CR)		0		4 (6.6)
• Partial Response (PR)		30 (49.2)		35 (57.4)
• Stable Disease (SD)		30 (49.2)		21 (34.4)
• Progressive Disease (PD)*		0		0
• Not Evaluable (NE)		1 (1.6)		1 (1.6)
Ongoing with unconfirmed response, n (%)		4 (6.6)		3 (4.9)
Ongoing without a response, n (%)		20 (32.8)		7 (11.5)
Objective response rate CR + PR (ORR), n (%)		30 (49.2)		39 (63.9)
• 95% Confidence interval		(36.1, 62.3)		(50.6, 75.8)
• 90% Confidence interval		(38.0, 60.4)		(52.6, 74.2)
Disease Control Rate CR + PR + SD (DCR), n (%)		60 (98.4)		60 (98.4)
• 95% Confidence interval		(91.2, 100.0)		(91.2, 100.0)
• 90% Confidence interval		(92.5, 99.9)		(92.5, 99.9)

Note: 95% and 90% confidence intervals are constructed using 2-sided Clopper-Pearson method.
Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.

Figure 3 Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)



Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).
Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8), i.e. 98.1% of patients had their tumour reduce in size at some point during follow-up in their RCC target lesions.
Date of Data Cut-off: 01APR2022

Figure 4 Waterfall plot - percentage change in total sum of target lesions diameters for RCC, CNS hemangioblastoma, and pancreatic neuroendocrine tumours, from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)

Date of Data Cut-off: 01APR2022.
Note: Only data from 60 of the 61 participants in the efficacy analysis set are shown as one participant was without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses for RCC and CNS tumours, and therefore the percentage change could not be calculated.

Figure 5 Spider plot - percentage change in total sum of RCC target lesion diameters from date of partial response (RECIST 1.1) – independent review committee (efficacy analysis set)

Date of Data Cut-off: 01APR2022

Duration of response

In the 39 patients for whom CR or PR was recorded (also shown in Table 16), the median DOR was not reached as of the 01-APR-2022 database cut-off date (50% of the patients who had CR or PR need to have subsequently had disease progression or death in order for median DOR to be calculated, but only 5 such patients [12.8%] had progressed or died by the 01-APR-2022 data cutoff date). Bearing in mind that at the 01-APR-2022 data cut-off date the median length of follow-up is 37.7 months and the median time-to-response is 11.1 months (detailed in a later subsection), the fact that only 12.8% of patients who had CR or PR have subsequently had disease progression or death at this data cut-off date is indicative of a durable response. The range of DOR was 5.4+ to 25.8+ months (Table 17 and Figure 6).

Table 17 MK-6482-004 summary of duration of response for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with Confirmed Response, n (%)	39 (63.9)
Responders who Progressed or Died (%)	5 (12.8)
Duration of Response (Months) 95% CI	
n	39
Mean [1]	23.5
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (19.3, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	5.4+, 35.8+
Number (%) of Patients with Extended Response Duration [2]	
>=6 Months	36 (100.0)
>=12 Months	35 (100.0)
>=18 Months	29 (93.5)
>=24 Months	22 (86.6)
>=30 Months	10 (86.6)
>=36 Months	0 (NR)

NE: Not Estimable.
Duration of Response is analysed using the Kaplan-Meier estimator. Median, first and third quartiles of Duration of
Response is reported along with 95% Brookmeyer-Crowley confidence intervals.
[1] Arithmetic mean.
[2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
+ indicates there was no progressive disease by the time of last disease assessment.
Date of Data Cut-off: 01APR2022

Figure 6 MK-6482-004 Kaplan-Meier plot of duration of response for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
NE = Not estimable.
Date of Data Cut-off: 01APR2022

Time to response

The median TTR was 11.1 months (range: 2.7 to 30.5 months) among 39 participants with a confirmed best overall response (BOR) of CR or PR (Table 18).

Table 18 MK-6482-004 summary of time to response for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with Confirmed Response, n (%)	39 (63.9)
Time to Response (Months)	
n	39
Mean (SD)	12.4 (8.08)
Median	11.1
Min, Max	2.7, 30.5

Date of Data Cut-off: 01APR2022

Progression-free survival

The median (95% CI) PFS was [redacted] months. The PFS rate at Month 36 was [redacted] (Table 17 and Figure 7).

Table 19 MK-6482-004 summary of progression-free survival for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Subjects with Events, n (%)	[redacted]
Progression Disease	[redacted]
Death	[redacted]
Censored Subjects, n (%)	[redacted]
New Anticancer Therapy Initiated	[redacted]
No Baseline or Post-Baseline Tumour Assessment	[redacted]
Death or Progression after More than One Missed Assessments	[redacted]
No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1]	
Median (95% CI)	[redacted]
Q1 (95% CI)	[redacted]
Q3 (95% CI)	[redacted]
Progression-Free Survival Rate (%) (95% CI) [number at risk] at	
Month 6	[redacted]
Month 12	[redacted]
Month 18	[redacted]
Month 24	[redacted]
Month 30	[redacted]
Month 36	[redacted]
Month 42	[redacted]
Month 48	[redacted]

NE: Not Estimable.
[1] Progression-Free Survival are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley confidence intervals.
Date of Data Cut-off: 01APR2022

Figure 7 Kaplan-Meier plot of progression-free survival for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

Progression-Free Survival is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalized Brookmeyer-Crowley method. Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
NE = Not estimable.
Date of Data Cut-off: 01APR2022

Time to surgery

At the 01-APR-2022 database cut-off date, 7 patients (11.5%) had undergone surgery, the median time to surgery is not evaluable.

Table 20 Summary of time to surgery for RCC tumours (efficacy analysis set)

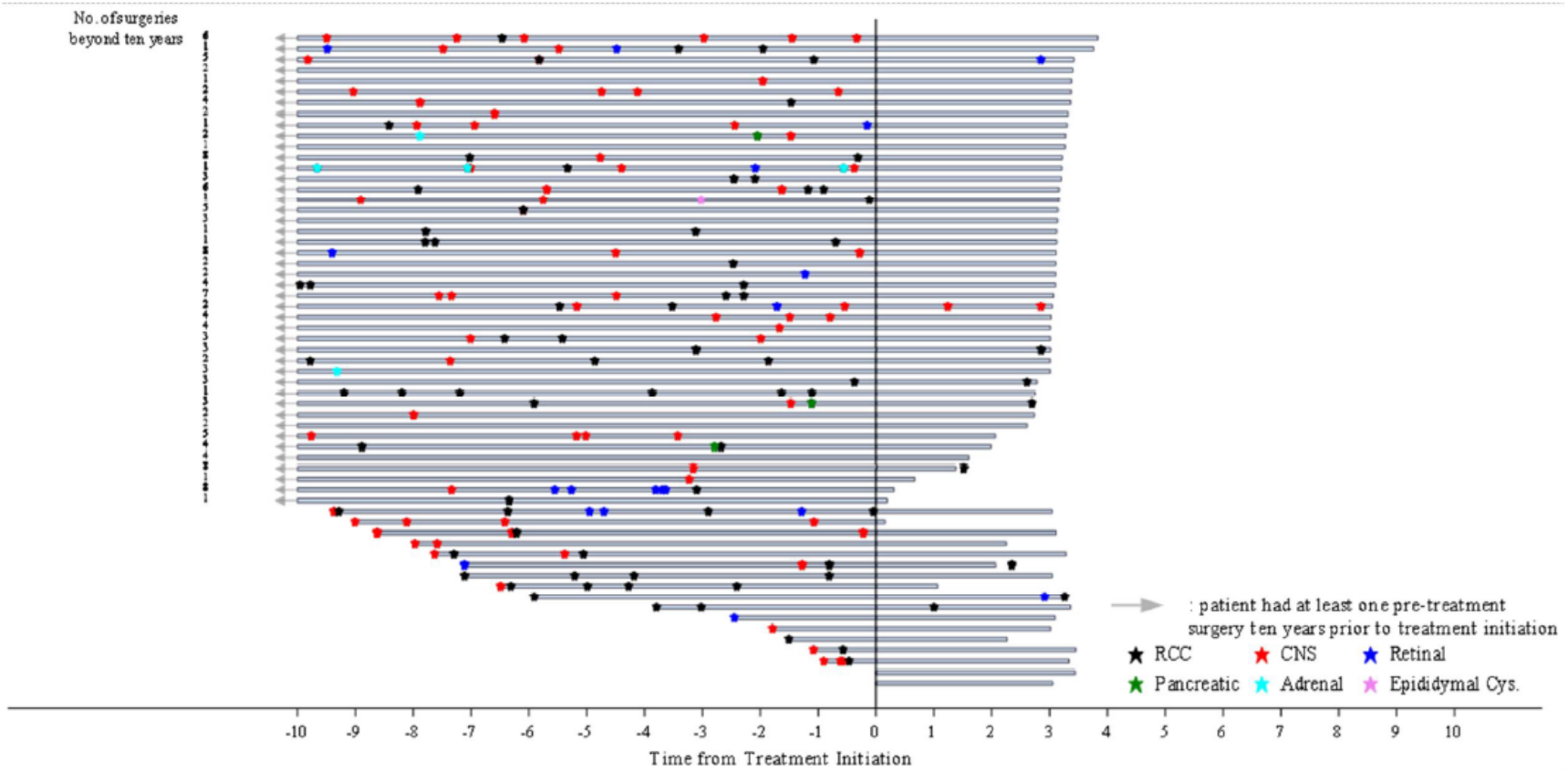
	Belzutifan (N=61)
Number of Subjects Undergo Surgeries, n (%)	7 (11.5)
Time to Surgery (Months) 95% CI	
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (39.2, NE)
Q3 (95% CI)	NE (NE, NE)

The Q1, median, Q3, and 95% CI are obtained from Kaplan-Meier estimates.
Surgery includes any procedure, excluding radiation, which leads to reduction of RCC tumor size.
NE: Not Estimable.
Date of Data Cut-off: 01APR2022

Rate of surgeries

A comparison of the VHL disease-associated tumour-related surgeries patients underwent before and after initiation of treatment with belzutifan is shown in Figure 8 (note that only pre-treatment surgeries less than 10 years prior to treatment initiation are presented). From this it can be seen that that the frequency of VHL disease-associated surgeries in the time period after initiation of treatment with belzutifan is lower than observed in the time period before, which is indicative of a potentially practice-changing favourable effect of belzutifan treatment on subsequent rate of VHL disease-associated surgeries.

Figure 8 Distribution of all surgeries pre- and post-treatment initiation over time for individual patients - safety analysis set



Horizontal bars represent each patient.
Only pre-treatment surgeries less than 10 years prior to treatment initiation are presented.
Length of the bars on the right side of the y-axis represents duration of treatment at time of data cut-off.
Surgery is defined as a tumour reduction procedure excluding radiation.
Date of Data Cut-off: 01-APR-2022.

B.2.7 Subgroup analysis

Summary of key subgroup analyses results:

Patients with VHL-associated RCC and CNS hemangioblastomas:

- Belzutifan provides a clinically meaningful confirmed ORR among the 50 participants with CNS hemangioblastoma at baseline per IRC assessment of 44.0% (95% CI: 30.0, 58.7). Four patients (8.0%) achieved a BOR of CR and 18 participants (36.0%) achieved a BOR of PR.
- Belzutifan provided a DCR of 90.0% (95% CI: 78.2%, 96.7%) in CNS hemangioblastomas.
- Responses to belzutifan treatment were long and durable as shown by a median DOR that was not reached as of the 01-APR-2022 database cut-off date. The range of DOR was 3.7+ to 38.7+ months, 12 patients achieved a DOR ≥30 months.
- The median TTR was [redacted] participants with a confirmed BOR of CR or PR.
- The median PFS for patients with CNS hemangioblastoma [redacted] at the 01-APR-2022 database cut-off date, [redacted] had a PFS event (9 had disease progression and 2 died).
- At the 01-APR-2022 database cut-off date, only one patient with CNS hemangioblastoma had undergone surgery.

In patients with VHL disease-associated RCC and pancreatic neuroendocrine tumours:

- Belzutifan provides a clinically meaningful confirmed ORR among 22 participants with pancreatic neuroendocrine tumours at baseline per IRC assessment of 90.9% (95% CI: 70.8, 98.9). As of the 01-APR-2022 database

cut-off date, 7 participants (31.8%) achieved a BOR of CR and 13 participants (59.1%) achieved a BOR of PR.

- Among the subgroup of 22 participants with pNETs, belzutifan provided a DCR of 100% (95% CI: 84.6%, 100%) in these tumours.
- Responses to belzutifan treatment were long and durable as shown by a median DOR that was not reached as of the 01-APR-2022 database cut-off date. The range of DOR was 11.0+ to 37.3+ months, 15 participants achieved a DOR ≥24 months.
- The median TTR was [redacted] participants with a confirmed BOR of CR or PR.
- The median PFS for patients with pancreatic neuroendocrine tumours [redacted] at the 01-APR-2022 database cut-off date, [redacted] had a PFS event.
- At the 01-APR-2022 database cut-off date, no patient with pancreatic neuroendocrine tumour had undergone surgery.

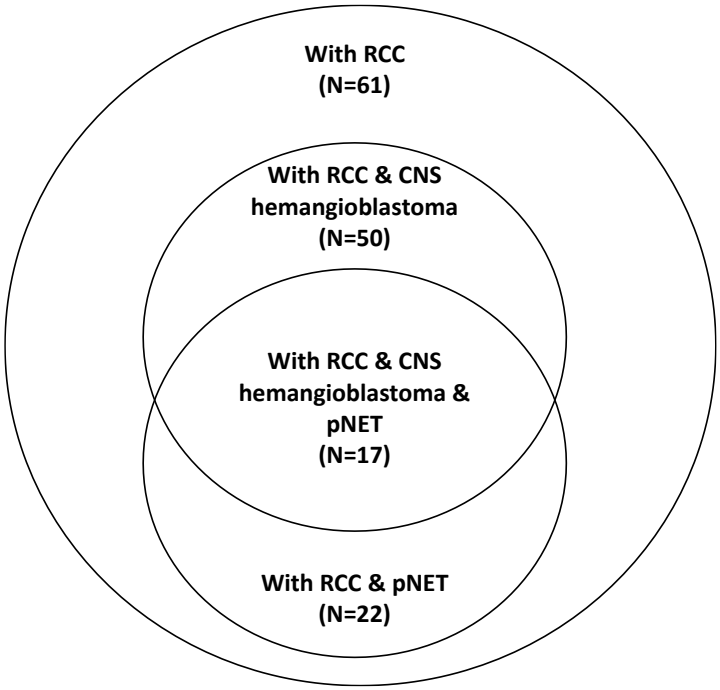
Results are presented in this subsection for the subgroups of patients with VHL RCC who also had central nervous system (CNS) hemangioblastomas or pancreatic neuroendocrine tumours (pNETs). In terms of numbers of patients in the MK-6482-004 study (also summarised in Figure 9):

- Of the total 61 patients with RCC (results for this population reported in section B.2.6):
    - 50 of the 61 patients with RCC also had CNS hemangioblastomas (results for this population reported later in this section B.2.7. Please note that 50 of the 61 patients with RCC also had CNS hemangioblastomas according to IRC, this number differs slightly to the number of patients with RCC also had CNS hemangioblastomas reported in the baseline characteristics table in Table 10 as being 51 as the 51 is the number according to only investigator assessment and not IRC).

- 17 of the 50 patients with RCC and CNS hemangioblastomas also had pNETs (results for this population not reported separately)
- 33 of the 50 patients with RCC and CNS hemangioblastomas did not also have pNETs (results for this population not reported separately)

- 22 of the 61 patients with RCC also had pNETs (results for this population reported later in this section B.2.7)
  - 17 of the 22 patients with RCC and pNETs also had CNS hemangioblastomas (results for this population not reported separately)
  - 5 of the 22 patients with RCC and pNETs did not also have CNS hemangioblastomas (results for this population not reported separately)

Figure 9 Summary of subgroups in the MK-6482-004 study (not to scale)



Central nervous system hemangioblastomas

Overall response rate

The confirmed ORR among the 50 participants with CNS hemangioblastoma at baseline per IRC assessment was 44.0% (95% CI: 30.0, 58.7). Four patients (8.0%) achieved a BOR of CR and 18 participants (36.0%) achieved a BOR of PR (Table 21 and Figure 10). The specific locations in the CNS of the tumours with each category of vest overall response are shown in Appendix P, due to the very low sample sizes, no conclusive patterns/trends can be observed with regard to precise locations of tumours in the CNS.

Of the four patients in whom a complete response was reported in their target CNS hemangioblastoma by the 01-APR-2022 data cut-off date, their target tumour [redacted] from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with belzutifan persist. For the 18 patients in whom a partial response was reported by the 01-APR-20223 data cut-off date, the change in their target CNS hemangioblastoma size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date are shown in Figure 11. It can be seen that [redacted].

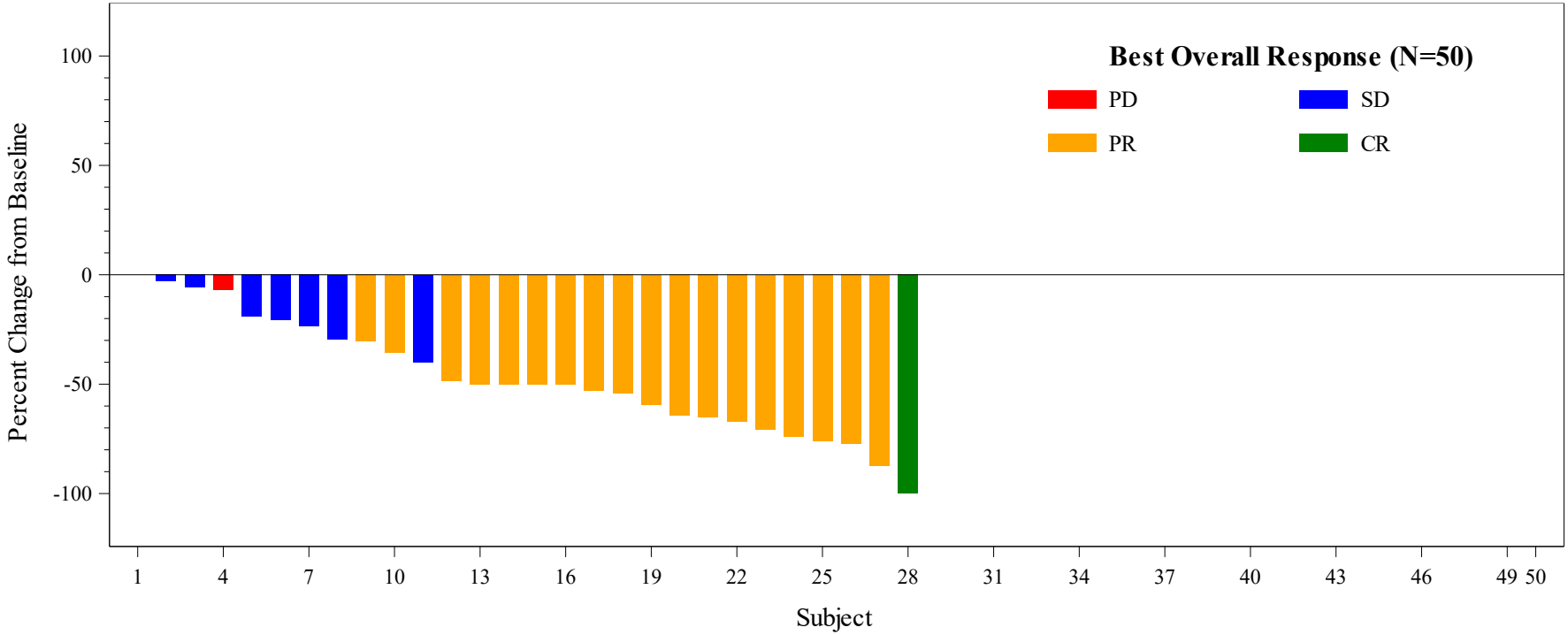
No patients had a best overall response of PD (shown in Table 21) despite there being 9 patients shown to have disease progression in the analysis of PFS (discussed later in this document and shown in Table 24). This because *best overall response* is the best response status recorded in the patient at any point during their follow-up period i.e. if a patient had an overall tumour response state of PR or SD at any point prior to disease progression in their target tumour their best overall response would be PR or SD (whichever was the better one had). The only way a patient could have a best overall response of PD is if they were recorded to have PD at their first post-baseline follow-up scan and then discontinued the trial or never subsequently experienced a CR, PR, or SD.

Table 21 MK-6482-004 summary of best overall tumour response for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1 (N1/N%)	[redacted]	50 (82.0)	[redacted]	50 (82.0)
Best Overall Response, n (n/N1%)				
• Complete Response (CR)	[redacted]	3 (6.0)	[redacted]	4 (8.0)
• Partial Response (PR)	[redacted]	12 (24.0)	[redacted]	18 (36.0)
• Stable Disease (SD)	[redacted]	31 (62.0)	[redacted]	23 (46.0)
• Progressive Disease (PD)	[redacted]	2 (4.0)	[redacted]	3 (6.0)
• Not Evaluable (NE)	[redacted]	2 (4.0)	[redacted]	2 (4.0)
Ongoing with unconfirmed response, n (n/N1%)	[redacted]	2 (4.0)	[redacted]	1 (2.0)
Ongoing without a response, n (n/N1%)	[redacted]	28 (56.0)	[redacted]	13 (26.0)
Objective response rate CR + PR (ORR), n (n/N1%)	[redacted]	15* (30.0)	[redacted]	22 (44.0)
• 95% Confidence interval	[redacted]	(17.9, 44.6)	[redacted]	(30.0, 58.7)
• 90% Confidence interval	[redacted]	(19.5, 42.4)	[redacted]	(32.0, 56.6)
Disease Control Rate CR + PR + SD (DCR), n (n/N1%)	[redacted]	46 (92.0)	[redacted]	45 (90.0)
• 95% Confidence interval	[redacted]	(80.8, 97.8)	[redacted]	(78.2, 96.7)
• 90% Confidence interval	[redacted]	(82.6, 97.2)	[redacted]	(80.1, 96.0)

Note: 95% and 90% confidence intervals are constructed using 2-sided Clopper-Pearson method.
Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.
Patients evaluable at baseline per IRC are included.
* There were changes in assessments of imaging data that resulted in an overall decrease in the number of participants with responses for CNS hemangioblastoma (from 16 to 15 participants with confirmed response) and an overall decrease in the number of PFS events for pancreatic neoplasms (1 less PFS event) and CNS hemangioblastomas (1 less PFS event) since submission of initial application, at the 01-DEC-2020 data cut-off date.

Figure 10 Waterfall plot - percentage change in total sum of CNS hemangioblastoma target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)



Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).
Number (%) of patients with maximum % reduction in sum of diameters of target lesions < 0 = 27 (54.0), i.e. 54.0% of patients had their tumour reduce in size at some point during follow-up in their CNS hemangioblastoma target lesions.
Date of Data Cut-off: 01APR2022

Figure 11 Spider plot - Percentage change in total sum of CNS hemangioblastoma target lesion diameters from date of partial response (RECIST 1.1) – independent review committee (efficacy analysis set)

Database cut-off date: 01APR2022

Duration of response

The median DOR was not reached as of the 01-APR-2022 database cut-off date. The range of DOR was 3.7+ to 38.7+ months, 12 patients achieved a DOR ≥30 months (Table 22 and Figure 12).

Table 22 MK-6482-004 summary of duration of response for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1/N	50 (82.0)
Patients with Confirmed Response, n (n/N1%)	22 (44.0)
Responders who Progressed or Died (%)	4 (18.2)
Duration of Response (Month) 95% CI	
Mean [1]	23.9
Median (95% CI)	NE (30.9, NE)
Q1 (95% CI)	31.3 (5.5, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	3.7+, 38.7+
Number (%) of Patients with Extended Response Duration [2]	
>=6 Months	19 (95.2)
>=12 Months	16 (90.2)
>=18 Months	14 (90.2)
>=24 Months	13 (90.2)
>=30 Months	12 (90.2)
>=36 Months	2 (72.2)
>=42 Months	0 (NR)

NE: Not Estimable.
Duration of Response is analyzed using the Kaplan-Meier estimator. Median, first and third quartiles of Duration of Response are reported along with 95% Brookmeyer-Crowley confidence intervals.
[1] Arithmetic mean.
[2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
+ indicates there was no progressive disease by the time of last disease assessment.
Date of Data Cut-off: 01APR2022

Figure 12 MK-6482-004 Kaplan-Meier plot of duration of response for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
NE = Not estimable.
Date of Data Cut-off: 01APR2022

Time to response

The median TTR was [redacted] participants with a confirmed BOR of CR or PR (Table 23).

Table 23 MK-6482-004 summary of time to response for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1 (N1/N%)	[redacted]
Patients with Confirmed Response, n (n/N1%)	[redacted]
Time to Response (months) n Mean (SD) Median Min, Max	[redacted]

Date of Data Cut-off: 01APR2022

Progression-free survival

The median PFS for patients with CNS hemangioblastoma [redacted] at the 01-APR-2022 database cut-off date, [redacted] patients had a PFS event (Table 24 and Figure 13).

Table 24 MK-6482-004 summary of progression-free survival for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1/N	[redacted]
Subjects with Events, n (n/N1 %)	[redacted]
Progression Disease	[redacted]
Death	[redacted]
Censored Subjects, n (n/N1 %)	[redacted]
New Anticancer Therapy Initiated	[redacted]
No Baseline or Post-Baseline Tumour Assessment	[redacted]
Death or Progression after More than One Missed Assessments	[redacted]
No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1]	
Median (95% CI)	[redacted]
Q1 (95% CI)	[redacted]
Q3 (95% CI)	[redacted]

NE: Not Estimable.
[1] Progression-Free Survival are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley confidence intervals.
Date of Data Cut-off: 01APR2022

Figure 13 Kaplan-Meier plot of progression free survival for CNS hemangioblastomas (RECIST 1.1) – independent review committee (efficacy analysis set)

Progression-Free Survival is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method. Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
NE = Not estimable.
Date of Data Cut-off: 01APR2022

Time to surgery

At the 01-APR-2022 database cut-off date, only one patient with CNS hemangioblastoma had undergone surgery. Consequently, the median time to surgery is not evaluable for this subgroup.

Rate of surgeries

A comparison of the VHL disease-associated tumour-related surgeries these patients underwent before and after initiation of treatment with belzutifan is shown in Figure 14 (note that only the blue bars indicate patients with CNS hemangioblastomas at baseline per IRC).

Figure 14 Distribution of all surgeries pre- and post-treatment initiation over time for individual patients for individual patients with baseline CNS hemangioblastomas per independent review committee - safety analysis set

Horizontal bars represent each patient.
Blue bars indicate patients with CNS Hemangioblastomas at baseline per IRC.
Only pre-treatment surgeries less than 10 years prior to treatment initiation are presented.
Length of the bars on the right side of the y-axis represents duration of treatment at time of data cut-off.
Surgery is defined as a tumour reduction procedure excluding radiation.
Date of Data Cut-off: 01-APR-2022.

Pancreatic neuroendocrine tumours

Overall response rate

The confirmed ORR among 22 participants with pancreatic neuroendocrine tumours at baseline per IRC assessment was 90.9% (95% CI: 70.8, 98.9). As of the 01-APR-2022 database cut-off date, 7 participants (31.8%) achieved a BOR of CR and 13 participants (59.1%) achieved a BOR of PR (Table 25 and Figure 15).

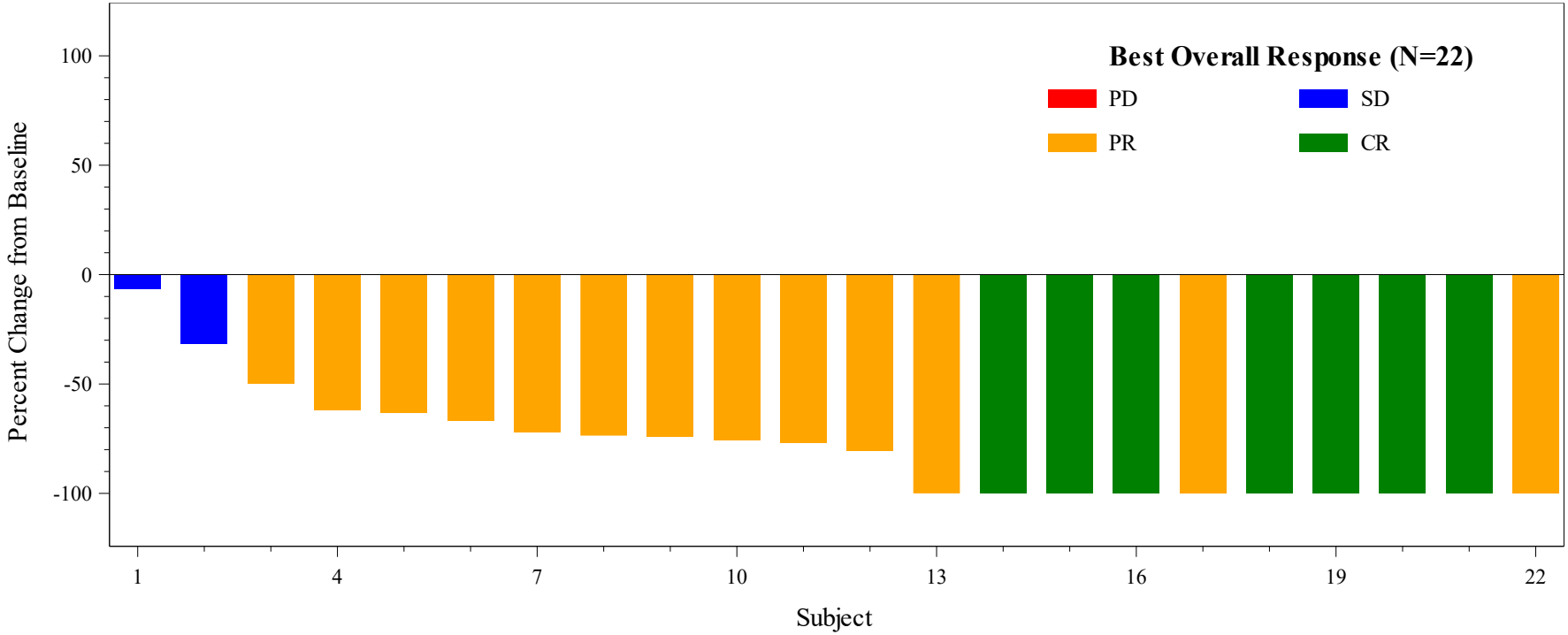
Of the seven patients who experienced a complete response in their target pNET by the 01-APR-2022 data cut-off date, their target tumour [redacted] from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with belzutifan persist. For the 13 patients who experienced a partial response by the 01-APR-20223 data cut-off date, the change in their target pNET size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date are shown in Figure 16. It can be seen that [redacted].

Table 25 MK-6482-004 summary of best overall tumour response for pancreatic neuroendocrine tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1 (N1/N%)	[redacted]	22 (36.1)	[redacted]	22 (36.1)
Best Overall Response, n (n/N1%)				
• Complete Response (CR)	[redacted]	3 (13.6)	[redacted]	7 (31.8)
• Partial Response (PR)	[redacted]	17 (77.3)	[redacted]	13 (59.1)
• Stable Disease (SD)	[redacted]	2 (9.1)	[redacted]	2 (9.1)
• Progressive Disease (PD)	[redacted]	0	[redacted]	0
• Not Evaluable (NE)	[redacted]	0	[redacted]	0
Ongoing with unconfirmed response, n (n/N1%)	[redacted]	1 (4.5)	[redacted]	0
Ongoing without a response, n (n/N1%)	[redacted]	0	[redacted]	0
Objective response rate CR + PR (ORR), n (n/N1%)	[redacted]	20 (90.9)	[redacted]	20 (90.9)
• 95% Confidence interval	[redacted]	(70.8, 98.9)	[redacted]	(70.8, 98.9)
• 90% Confidence interval	[redacted]	(74.1, 98.4)	[redacted]	(74.1, 98.4)
Disease Control Rate CR + PR + SD (DCR), n (n/N1%)	[redacted]	22 (100.0)	[redacted]	22 (100.0)
• 95% Confidence interval	[redacted]	(84.6, 100.0)	[redacted]	(84.6, 100.0)
• 90% Confidence interval	[redacted]	(87.3, 100.0)	[redacted]	(87.3, 100.0)

Note: 95% and 90% confidence intervals are constructed using 2-sided Clopper-Pearson method.
Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.

Figure 15 Waterfall plot - percentage change in total sum of target lesions diameters for pancreatic neuroendocrine tumours from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)



Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).
Number (%) of patients with maximum % reduction in sum of diameters of target lesions < 0 = 22 (100.0), i.e. 100% of patients had their tumour reduce in size at some point during follow-up in their pNET target lesions.
Date of Data Cut-off: 01APR2022

Figure 16 Spider plot - Percentage change in total sum of target lesion diameters for pNETs from date of partial response (RECIST 1.1) – independent review committee (efficacy analysis set)

Database cut-off date: 01APR2022

Duration of response

The median DOR was not reached as of the 01-APR-2022 database cut-off date, no patients in this subgroup had progression or died by the 01-APR-2022 data cutoff date. The range of DOR was 11.0+ to 37.3+ months, 15 participants achieved a DOR ≥24 months (Table 26 and Figure 17).

Table 26 MK-6482-004 summary of duration of response for pancreatic neuroendocrine tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	22 (36.1)
Patients with Confirmed Response, n (n/N1%)	20 (90.9)
Responders who Progressed or Died (%)	0
Duration of Response (Months) 95% CI	
Mean [1]	27.4
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (NE, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	11.0+, 37.3+
Number (%) of Patients with Extended Response Duration [2]	
>=6 Months	20 (100.0)
>=12 Months	19 (100.0)
>=18 Months	19 (100.0)
>=24 Months	15 (100.0)
>=30 Months	8 (100.0)
>=36 Months	1 (100.0)
>=42 Months	0 (NR)

NE: Not Estimable.
Duration of Response is analysed using the Kaplan-Meier estimator. Median, first and third quartiles of Duration of Response are reported along with 95% Brookmeyer-Crowley confidence intervals.
[1] Arithmetic mean.
[2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
+ indicates there was no progressive disease by the time of last disease assessment.
Date of Data Cut-off: 01APR2022

Figure 17 MK-6482-004 Kaplan-Meier plot of duration of response for pancreatic neuroendocrine tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
NE = Not estimable.
Date of Data Cut-off: 01APR2022

Time to response

The median TTR was [redacted] participants with a confirmed BOR of CR or PR (Table 27).

Table 27 MK-6482-004 summary of time to response for pancreatic neuroendocrine tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	[redacted]
Patients with Confirmed Response, n (n/N1%)	[redacted]
Time to Response (Months) n Mean (SD) Median Min, Max	[redacted]

Date of Data Cut-off: 01APR2022

Progression-free survival

The median PFS for patients with pancreatic neuroendocrine tumours [redacted] at the 01-APR-2022 database cut-off date, [redacted] had a PFS event (Table 28).

Table 28 MK-6482-004 summary of progression-free survival for pancreatic neuroendocrine tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	[redacted]
Subjects with Events, n (n/N1 %) Progression Disease Death	[redacted]
Censored Subjects, n (n/N1 %) New Anticancer Therapy Initiated No Baseline or Post-Baseline Tumour Assessment Death or Progression after More than One Missed Assessments No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1] Median (95% CI) Q1 (95% CI) Q3 (95% CI)	[redacted]

NE: Not Estimable.

[1] Progression-Free Survival are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley confidence intervals.
Date of Data Cut-off: 01APR2022

Time to surgery

At the 01-APR-2022 database cut-off date, no patient with pancreatic neuroendocrine tumour had undergone surgery. Consequently, the time to surgery is not evaluable for this subgroup. Data sources used to estimate time to surgery in this population for the cost-effectiveness analyses are described in section B.3.

Other tumours

The MK-6482-004 study collected data on several other tumour types in addition to the tumour types covered in the MHRA marketing authorisation and the scope of this current appraisal. Key results for these tumours are summarised in the following subsections. The results indicate treatment with belzutifan were associated with valuable positive effects in these tumours.

Pancreatic lesions

Results from the MK-6482-004 study were collected for the pancreatic lesions subgroup, which included both pNET and non-pNET lesions, pNET lesions were defined as solid parenchymal lesions that do not communicate with the pancreatic duct, while non-pNET lesions were defined as all pancreatic lesions that were not pNET lesions.

Treatment with belzutifan showed [redacted] ORR in participants with pancreatic lesions; the ORR by IRC was [redacted]. The DCR for pancreatic lesions was [redacted].

The median DOR for participants with pancreatic lesions was [redacted], and based on Kaplan-Meier estimation, [redacted] of responders had an ongoing response at 30 months. The median TTR was [redacted]. Median PFS and TTS were [redacted]. [redacted] underwent surgery for pancreatic lesions as of the 01-APR-2022 data cut-off date.

Retinal hemangioblastoma

Twelve of 17 participants in the MK-6482-004 study with baseline retinal hemangioblastomas were evaluable for response with follow-up evaluations. Of 12 participants with retinal hemangioblastoma, evaluable retinal hemangioblastomas were determined in 16 eyes at baseline per IRC assessment.

Treatment response in retinal hemangioblastoma, per IRC, was assessed using multiple parameters such as number/size/location, degree of feeder/drainer engorgement (mild/prominent), presence of intraretinal heme, presence of preretinal heme, presence of vitreous heme, presence of lipid exudation, presence of subretinal fluid, and presence of fibrosis (37).

An improvement of retinal hemangioblastoma was observed after treatment with belzutifan. The response of 'Improved' was 100% (95% CI: 79.4, 100.0) in all 16 eyes and 100% (95% CI 73.5, 100.0) in all 12 participants. Median DOR was not reached. All 12 participants had an improvement for ≥12 months, and of these, 9 participants had improvement for ≥30 months. Median TTR was [redacted].

Visual acuity of participants with retinal hemangioblastoma underwent ophthalmologic evaluation (by investigator assessment). Visual acuity in most participants [redacted].

Adrenal lesions and endolymphatic sac tumours

As of the 01-APR-2022, per investigator assessment, [redacted] with adrenal lesions (n=3) and endolymphatic sac tumours (n=1) had a BOR of [redacted]; median TTR [redacted]. Median PFS was [redacted].

Epididymal cystadenomas

Sixteen participants had epididymal cystadenomas at baseline and were followed up by ultrasound examination. Per investigator assessment, at Week 49, [redacted] had improvement in lesions compared with baseline, [redacted] had stable lesions, and [redacted] had progressed.

B.2.8 Meta-analysis

The MK-6482-004 study is the only study that reports clinical effectiveness data on the treatment effect of belzutifan in the relevant indication, therefore no meta-analysis possible. Information on the effectiveness of the comparator (standard of care in UK clinical practice) was derived from data collected in a retrospective non-interventional study conducted in the United States (the VHL Natural History Study) described in section B.2.9.

B.2.9 Indirect and mixed treatment comparisons

- A matching adjusted indirect comparison (MAIC) using data from the MK-6482-004 study and a retrospective non-intervention VHL Natural History Study was conducted to compare the outcomes of treatment with belzutifan with the outcomes observed in the standard of care for this patient population, in order to inform the cost-effectiveness analyses. The inclusion criteria for the real-world study reflected the inclusion criteria for the MK-6482-004 study, which is different to the final MHRA indication wording as described earlier. Therefore, there are some generalisability considerations required as the patients in these two studies are not quite the same as those covered by the MHRA label.

- The MAIC found that treatment with belzutifan, compared to standard of care, was associated with a lower exponential rate parameter for the cause-specific hazards of pre-surgery to first surgery (0.03692 events/person-year compared to 0.25324 events/person-year) and a lower incidence of non-RCC VHL-related surgeries with therapeutic intent (0.02119 events/person-year compared to 0.178984 events/person-year).

- Additionally, the Optum study was used to align effectiveness data sourced from the VHL Natural History Study with real world UK SoC for the purposes of the cost-effectiveness analysis (34). Details of this are provided in section B.3.3.

Propensity score weighting-based matching-adjusted indirect comparison versus standard of care

Objective and rationale

For the purposes of the cost-effectiveness analyses described in section B.3, it is necessary to compare the outcomes of treatment with belzutifan with the outcomes observed in the standard of care for this patient population. This required data to inform the comparator arm of patients (i.e. patients managed via standard of care) that are comparable to the population of the MK-6482-004 study. This was done by generating a comparator population from the population assessed in the VHL Natural History Study from which a matching-adjusted indirect comparison can be performed.

The VHL Natural History Study was commissioned prior to GB marketing authorisation was finalised. While it provides useful data it does not describe well the standard of care population specific to current indication under assessment.

Data source for the comparator arm

The VHL Natural History study was a retrospective non-interventional study of existing medical records, with supplemental electronic medical record (EMR) data abstraction and central imaging review of abdominal imaging scans obtained during routine clinical care. Patients with VHL disease who had ≥1 measurable renal solid tumour measured during the study period and met the other patient eligibility criteria of the VHL Natural History Study were identified and followed until the end of the assessment window (July 31, 2004 to June 30, 2020).

The study used data collected by the United States National Cancer Institute's (NCI) Urologic Oncology Branch (UOB) of patients with VHL disease managed and treated at the United States National Institutes of Health (NIH) Clinical Center in Bethesda, Maryland, a global leader in comprehensive care for VHL disease patients. Data collected on VHL patient characteristics, treatment, and follow-up information is available in medical records for all patients treated at the NCI's UOB since approximately 1987 (though the assessment window for the current analyses was July 31, 2004 to June 30, 2020). For the current analyses, the data source included

data registered by the NCI using a hybrid of existing structured data fields in the UOB Hereditary Database, linked to other structured NCI and external database (e.g., NCI laboratory and prescription database, and US National Death Index [NDI]), supplemented with available serial tumour measurements and additional medical record abstraction of unstructured data fields. The study leveraged existing serial tumour measurements that the NCI registered in the Hereditary Database to assess growth rates of renal solid tumours. Additional details on the methodology of the VHL Natural History Study are provided in Appendix O.

MAIC methods

Sample selection criteria

In order to benchmark against the results of the MK-6482-004 phase 2 trial of belzutifan and generate key parameters for elements of the standard of care arm of the cost-effectiveness model, a sub-population of patients was identified from the VHL Natural History Study who met inclusion and exclusion criteria that closely matched that of the trial. Patients were followed from the patient-level index date, defined as the earliest date that a measurable renal solid tumour was detected during the study period (July 31, 2004 to June 30, 2020). The available follow-up time for each patient spanned from their patient-level index date until the first of: mortality date; or last clinical encounter date during the study period.

The following study inclusion and exclusion criteria were applied to identify the patient population used to generate cost-effectiveness model inputs for the VHL-RCC indication:

Inclusion criteria:

- Patients treated at the NCI with VHL syndrome who are residents of the US or Canada
- Patients with ≥1 renal solid tumour identified and measured during the study period (July 31, 2004 to June 30, 2020)
- Patients with a diagnosis of VHL disease based on germline VHL alteration

Exclusion criteria:

- Patients with any renal procedure in the 30 days on or prior to patient-level index date
- Patients whose last follow-up date was on or prior to patient-level index date
- If the largest renal solid tumor at patient-level index date is ≥30 millimetres (mm), patients with a renal surgical procedure with therapeutic intent performed within 60 days on or after patient-level index date
- Patients who received treatment with belzutifan or another hypoxia inducible factor 2 alpha (HIF-2α) inhibitor any time prior to patient-level index date
- Patients who received systemic oncologic or investigational therapy any time prior to patient-level index date
- Patients with evidence of VHL disease-associated metastatic disease prior to patient-level index date

In order to generate cost-effectiveness model parameters for the VHL-CNS hemangioblastoma and VHL-pNET indications, patients who met the above inclusion/exclusion criteria were further restricted to those with a recorded history of CNS hemangioblastoma and pNET, respectively. As a limitation, it was not feasible using the available Natural History Study data to identify whether patients in these subsets had CNS hemangioblastoma and pNET at the patient-level index date (i.e., it was only feasible to identify patients with a recorded history of CNS hemangioblastoma or pNET at some point prior to the patient-level index date). In contrast, patients in the corresponding CNS hemangioblastoma and pNET sub-populations of the MK-6482-004 trial were confirmed to have CNS hemangioblastoma and pNET tumours at the baseline visit (the MK-6482-004 study baseline visit is equivalent to the VHL Natural History Study patient-level index date for the purposes of designating the time point from which follow-up should begin for these analyses). Due to this imbalance, certain parameter inputs that were estimated using VHL Natural History Study data for the VHL-RCC model cohort were obtained

from other data sources for the VHL-CNS hemangioblastomas and VHL-pNET model cohorts (these are detailed later in section B.3).

The resulting sample selection process is presented in Table 29 below.

Table 29 Sample selection process: Trial Population Subgroup

Step #	Criterion	N Patients
1	INCLUSION: Patients with VHL syndrome who are residents of the US or Canada	776
2	INCLUSION: Patients with ≥1 renal solid tumor identified and measured during the study period (July 31, 2004 to June 30, 2020)	313
3	INCLUSION: Patients with a diagnosis of Von Hippel-Lindau (VHL) syndrome based on germline VHL alteration	297
4	EXCLUSION: Patients with any renal procedure in the 30 days on or prior to Patient-level index date	296
5	EXCLUSION: Patients whose follow-up date was on or prior to Patient-level index date	296
6	EXCLUSION: If the largest tumor at Patient-level index date is ≥30 millimeters (mm), patients with a renal surgical procedure with therapeutic intent performed within 60 days on or after Patient-level index date	278
7	EXCLUSION: Patients who received treatment with MK-6482 or another hypoxia inducible factor 2 alpha (HIF-2α) inhibitor any time prior to Patient-level index date	278
8	EXCLUSION: Patients who received systemic oncologic or investigational therapy any time prior to Patient-level index date	272
9	EXCLUSION: Patients with evidence of VHL disease-associated metastatic disease prior to Patient-level index date	260
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL-RCC cohort**	260
10.a	INCLUSION: Patients at step #9 with ≥1 concomitant CNS hemangioblastoma on or before Patient-level index date	228
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL-CNS Hb cohort** *(subset of the 260 patients in the VHL-RCC Natural History Study sample)*	228
10.b	INCLUSION: Patients at step #9 with ≥1 concomitant pNET on or before Patient-level index date	94
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL-pNET cohort** *(subset of the 260 patients in the VHL-RCC Natural History Study sample)*	94

CNS Hb, central nervous system hemangioblastoma; HIF-2α: hypoxia inducible factor 2 alpha; mm, millimeters; pNET, pancreatic neuroendocrine tumor; RCC, renal cell carcinoma; US, United States; VHL, Von Hippel-Lindau

Baseline risk adjustment through propensity score reweighting

After applying inclusion/exclusion criteria similar to those used in the MK-6482-004, the key population characteristics of the selected samples for the relevant cohorts

(all [RCC] patients, the CNS hemangioblastoma subgroup, and the pNETs subgroup) are those summarised in in Table 30 to Table 32.

Table 30 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before – VHL RCC cohort

Baseline characteristics	VHL Natural History Study (N=260)		MK-6482-004 (N=61)	
Age at patient-level index date (years)				
Mean	42.1		41.0	
Standard deviation	12.3		13.5	
Sex, N (%)				
Female	120	46.2 %	29	47.5 %
Male	140	53.9 %	32	52.5 %
Number of renal surgeries with therapeutic intent prior to patient-level index date[2]				
Mean	1.4		2.4	
Standard deviation	1.5		1.6	
Tumour size of the largest renal solid tumour at patient-level index date (cm)				
Mean	2.1		2.5	
Standard deviation	1.0		0.9	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.
[2] This “Number of renal surgeries with therapeutic intent prior to patient-level index date” variable and its definition/criteria is not the same as (it is more restrictive than) that of the "Number of Prior Surgeries per Subject" variable shown in Table 10.

Table 31 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before matching – VHL CNS hemangioblastoma cohort (patients with VHL-associated RCC and CNS hemangioblastoma)

Baseline characteristics	VHL Natural History Study subgroup with CNS hemangioblastoma history (N=228)		MK-6482-004 CNS hemangioblastoma subgroup (N=50)	
Age at patient-level index date (years)				
Mean	42.3		40.4	
Standard deviation	11.7		12.8	
Sex, N (%)				
Female	102	44.7%	20	40.0%
Male	126	55.3%	30	60.0%
Number of CNS surgeries with therapeutic intent prior to patient-level index date				

Baseline characteristics	VHL Natural History Study subgroup with CNS hemangioblastoma history		MK-6482-004 CNS hemangioblastoma subgroup	
	(N=228)		(N=50)	
Mean	1.0		2.8	
Standard deviation	1.3		2.6	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.

Table 32 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before matching – VHL pNET cohort (patients with VHL-associated RCC and pNET)

Baseline characteristics	VHL Natural History Study subgroup with pNET history		MK-6482-004 pNET subgroup	
	(N=94)		(N=22)	
Age at patient-level index date (years)				
Mean	45.5		42.7	
Standard deviation	11.7		15.0	
Sex, N (%)				
Female	54	57.5%	12	54.5%
Male	40	42.6%	10	45.5%
Number of pancreatic surgeries with therapeutic intent prior to patient-level index date				
Mean	0.3		0.2	
Standard deviation	0.6		0.5	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.

The patients in these selected samples were then reweighted to further match the distribution of key baseline covariates among patients in the MK-6482-004 trial. Population-level baseline characteristics for patients in the MK-6482-004 trial were used for this analysis. The Natural History Study cohorts were reweighted to match the corresponding patient samples from MK-6482-004 using the matching-adjusted indirect comparison (MAIC) method, an inverse propensity weighting method which allows the logistic regression for propensity score to be estimated without individual patient data in one of the populations. Given that an anchor-based comparison of belzutifan vs. SOC was not feasible, the use of the MAIC method was considered *a priori* to be important for mitigating potential confounding due to baseline differences between the MK-6482-004 and VHL Natural History Study populations.

The baseline variables used for weighting adjustment are listed below for each patient cohort. These variables were selected after eliciting input from clinical experts on patient baseline characteristics that are prognostic of transition probabilities starting from the pre-surgery state, or that may modify the effect of belzutifan on these transition probabilities:

- Baseline covariates adjusted in the VHL-RCC cohort:
    - Age at patient-level index date (patient-level index date as defined earlier in the sample selection criteria subsection)
    - Gender
    - Number of renal surgeries with therapeutic intent prior to patient-level index date
    - Tumour size of the largest renal solid tumour at the patient-level index date
- Baseline covariates adjusted in the VHL-CNS Hb cohort:
    - Age at patient-level index date
    - Gender
    - Number of CNS surgeries with therapeutic intent prior to patient-level index date
- Baseline covariates adjusted in the VHL-pNET cohort:
    - Age at patient-level index date
    - Gender
    - Number of pancreatic surgeries with therapeutic intent prior to patient-level index date

The following additional covariates were noted as potentially relevant by the consulted experts, but could not be included in the matching adjustment due to data

limitations: VHL type and VHL gene alteration type (excluded due to a high proportion of missing values in the Natural History Study); number of concomitant measured tumours (unavailable at the time of analysis); and size of the largest CNS tumour at the patient-level index date in the VHL-CNS Hb cohort or size of the largest pancreatic tumour in the VHL-pNET cohort (the presence/absence of CNS Hb and pNET at the patient-level index date could not be identified using the available data).

For each cohort, the distribution of baseline characteristics before and after reweighting are displayed in the following tables. After reweighting, the effective sample size from the Natural History Study decreased by 65% (from 260 to 92.2) in the VHL-RCC cohort and by 83% (from 228 to 37.9) in the VHL-CNS hemangioblastoma cohort, as patients in these cohorts had fewer prior surgeries on average than their corresponding MK-6482-004 trial samples prior to matching. The effective sample size decreased by 36% (from 94 to 60.4) after reweighing the VHL-pNET Natural History Study cohort, as the number of prior pancreatic surgeries in this cohort was relatively well-balanced with the corresponding MK-6482-004 trial subgroup before matching.

For each cohort, the distribution of baseline characteristics after reweighting are displayed below in Table 29 to Table 31.

Table 33 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before and after reweighting – VHL RCC cohort

Baseline characteristics	**VHL Natural History Study**				**MK-6482-004**	
	Before reweighting (N=260)		**After reweighting (effective N=92.2)[1]**		**(N=61)**	
Age at patient-level index date (years)						
Mean	42.1		41.0		41.0	
Standard deviation	12.3		13.5		13.5	
Sex, N (%)						
Female	120	46.2 %	43.8	47.5 %	29	47.5 %
Male	140	53.9 %	48.4	52.5 %	32	52.5 %
Number of renal surgeries with therapeutic intent prior to patient-level index date[2]						

Baseline characteristics	VHL Natural History Study				MK-6482-004	
	Before reweighting (N=260)		After reweighting (effective N=92.2)[1]		(N=61)	
Mean	1.4		2.4		2.4	
Standard deviation	1.5		1.6		1.6	
Tumour size of the largest renal solid tumour at patient-level index date (cm)						
Mean	2.1		2.5		2.5	
Standard deviation	1.0		0.9		0.9	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.
[2] This “Number of renal surgeries with therapeutic intent prior to patient-level index date” variable and its definition/criteria is not the same as (it is more restrictive than) that of the "Number of Prior Surgeries per Subject" variable shown in Table 10.

Table 34 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before and after reweighting – VHL CNS hemangioblastoma cohort

Baseline characteristics	VHL Natural History Study subgroup with CNS hemangioblastoma history				MK-6482-004 CNS hemangioblastoma subgroup	
	Before reweighting (N=228)		After reweighting (effective N=37.9)[1]		(N=50)	
Age at patient-level index date (years)						
Mean	42.3		40.4		40.4	
Standard deviation	11.7		12.8		12.8	
Sex, N (%)						
Female	102	44.7%	15.1	40.0%	20	40.0%
Male	126	55.3%	22.7	60.0%	30	60.0%
Number of CNS surgeries with therapeutic intent prior to patient-level index date						
Mean	1.0		2.8		2.8	
Standard deviation	1.3		2.6		2.6	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.

Table 35 Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before and after reweighting – VHL pNET cohort

Baseline characteristics	VHL Natural History Study subgroup with pNET history	MK-6482-004 pNET subgroup

	Before reweighting (N=94)		After reweighting (effective N=60.4)[1]		(N=22)	
Age at patient-level index date (years)						
Mean	45.5		42.7		42.7	
Standard deviation	11.7		15.1		15.0	
Sex, N (%)						
Female	54	57.5%	32.9	54.5%	12	54.5%
Male	40	42.6%	27.5	45.5%	10	45.5%
Number of pancreatic surgeries with therapeutic intent prior to patient-level index date						
Mean	0.3		0.2		0.2	
Standard deviation	0.6		0.5		0.5	

[1] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.

MAIC results

Table 36 summarises key parameter inputs that were estimated from the reweighted VHL Natural History Study RCC cohort, alongside input values that were estimated analogously in the MK-6482-004 population. As shown, the weekly exponential rate of pre-surgery → 1st surgery transition was 0.00487 (0.25324 events/person-year) in the matched Natural History Study sample versus 0.00071 (0.03692 events/person-year) in the MK-6482-004, implying an 85% reduction in the cause-specific hazards of renal surgery with belzutifan. The incidence of non-RCC VHL-related surgeries in the two populations was 0.03692 events/person-year in the matched VHL Natural History Study sample versus 0.02119 events/person-year in the MK-6482-004 trial, a percentage reduction of 88%.

Table 36 Selected model parameters estimated in the reweighted VHL Natural History Study RCC cohort and the MK-6482-004 trial population

Outcomes	VHL Natural History Study	MK-6482-004
	After matching (effective N=92.2)	(N=61)
Exponential rate parameter for the cause-specific hazards of pre-surgery → 1st surgery		
Rate (events/person-year)	0.25324	0.03692
Standard error	(0.01768)	(0.0156)

Incidence of non-RCC VHL-related surgeries with therapeutic intent (events/person-year)		
Number of VHL-related surgeries	2116.4	208
Total person-years at risk	227.35	194.41
Incidence rate (events/person-year)	0.178984	0.02119

Uncertainties in the indirect and mixed treatment comparisons

The MAIC used the data from the MK-6482-004 study (N=61) and a relatively small post-selection and post-matching sample of the VHL Natural History Study (effective N=92.2) and so was based on a relatively small underlying data set. The number of relevant events (non-RCC VHL-related surgeries with therapeutic intent) observed in the MK-6482-004 study for belzutifan was also very small (only four events), which also limits the robustness of the MAIC results.

The inclusion criteria for the real-world study reflected the inclusion criteria for the MK-6482-004 study, which is different to the final MHRA indication wording as described earlier. Therefore, there are some generalisability considerations required as the patients in these two studies are not quite the same as those covered by the MHRA label.

Health-related quality-of-life data used in the cost-effectiveness analysis – VHL patient survey

- As the MK-6482-004 study did not collect HRQoL data, such data to inform the cost-effectiveness analyses were sourced from a non-interventional cross-sectional VHL patient survey that collected EQ-5D-5L that was converted to UK-specific EQ-5D-3L values using the standard crosswalk method.

- This is the first study of its kind in patients with VHL disease - through this study we have filled a gap in the current understanding of the HRQoL impact of VHL disease on patients.

Health-related quality of life data used in the cost-effectiveness analyses were obtained from a VHL patient survey, a non-interventional international, cross-sectional patient survey spanning the US, Canada, the UK, France, and Germany.

Over 200 patients were asked to complete one structured survey, designed to be completed in one sitting. Participation in the survey was voluntary, data collection took place between December 2021 and June 2022.

Objective

The objective of the survey was to understand the patient experience and burden of disease with current management practices in the treatment of VHL disease and the treatment preferences of patients with RCC, CNS hemangioblastoma, or pNET manifestations, including:

- To assess the patients' health-related quality of life.
- To assess the patients' work productivity loss and activity impairment.

Research methods

The data collected in this survey included patient perceptions of the burden of VHL and their experiences of undergoing surgery. The survey also included patient-reported outcome tools to help describe the burden (direct and indirect) of the condition. Data collection was conducted by a contract research organisation through a patient advocacy group, the VHL Alliance who operates in the US, and screened online. Prior to the start of data collection, all participating patients were provided with the relevant survey materials and instructions and provided informed consent prior to commencing the survey. The survey took approximately 30 minutes to complete and was only completed once by each participant. The survey materials were developed in English and translated then validated as applicable for each country. The survey was administered online and delivered to participants via email for completion on tablet or laptop.

<u>Inclusion Criteria</u>

Patients needed to meet all of the following inclusion criteria to be eligible for inclusion in the study:

- Adult, aged ≥18 years

- Have a diagnosis of VHL with manifestations in at least one of the following areas:
    - Kidney or
    - Brain or spinal cord or
    - Pancreas

Exclusion Criteria

Patients who met the following exclusion criteria were not eligible for inclusion in the study:

- Aged <18 years
- Do not have a diagnosis of VHL

Patient reported outcome measures

Patient reported outcome measures collected as part of the survey were the European Quality of Life – 5 Dimension Survey (EQ-5D-5L) and the Work Productivity and Activity Impairment: Specific Health Problem (WPAI:SHP). However, only the EQ-5D data are of relevance to this appraisal and were incorporated into the cost-effectiveness analyses (as described in section B.3.)

The EQ-5D was derived in two ways, namely the direct EQ-5D-5L scoring, and the crosswalk method which maps the EQ-5D-3L value sets to the 5-level questions, resulting in crosswalk value sets. The following methods were used for each of the domains: UK (crosswalk and direct) (39), Canada (direct), US (direct), France (direct and crosswalk), Italy (crosswalk) and Spain (direct).

Results

Patient demographics

Patient demographics are described in Table 37. Mean patient age was 42.5 years, 68.2% of patients were female and the majority of patients were white (88.6%). Time since diagnosis was 210.4 months (overall mean). Of the overall sample, 66.4% of patients had RCC manifestations, 55.5% of patients had pNET manifestations, and

86.4% of patients had CNS-Hb manifestations. Their disease status and whether they were prescribed medication for their VHL-related cancer is shown in Table 38.

Table 37 Patient demographics

		Type of Tumour			
		Base	RCC	pNET	CNS-Hb
Age (years)	Base (N, %)	220	146	122	190
		100%	100%	100%	100%
	Mean	42.5	43.0	42.8	42.9
	Min	18	19	18	18
	Max	77	77	70	70
	SD	13.78	14.13	14.05	13.91
Sex	Base	220	146	122	190
		100%	100%	100%	100%
	Male	70	50	41	64
		31.8%	34.2%	33.6%	33.7%
	Female	150	96	81	126
		68.2%	65.8%	66.4%	66.3%
Ethnicity	Base (N, %)	220	146	122	190
		100%	100%	100%	100%
	White	195	128	108	171
		88.6%	87.7%	88.5%	90.0%
	Asian-Indian subcontinent	2	1	1	1
		0.9%	0.7%	0.8%	0.5%
	South-East Asian	2	2	2	-
		0.9%	1.4%	1.6%	-
	Chinese	1	1	1	1
		0.5%	0.7%	0.8%	0.5%
	Japanese	1	1	1	1
		0.5%	0.7%	0.8%	0.5%
	Korean	2	1	1	2
		0.9%	0.7%	0.8%	1.1%
	Asian (other)	1	-	1	1
		0.5%	-	0.8%	0.5%
	Hispanic/Latino	9	7	3	7
		4.1%	4.8%	2.5%	3.7%
	Mixed Race	2	1	1	2
		0.9%	0.7%	0.8%	1.1%
	Other (specify)	5	4	3	4
		2.3%	2.7%	2.5%	2.1%
Time since diagnosis (months)	Base (N, %)	220	146	122	190
		100%	100%	100%	100%
	Mean	210.4	222.4	213.5	210.5
	Min	0	8	10	0
	Max	650	650	650	650
	SD	143.0	150.4	152.1	142.0
	Don't know	21	13	11	20
		100.0%	100.0%	100.0%	100.0%
Type of tumour	Base (N, %)	220	146	122	190

		Type of Tumour			
		Base	RCC	pNET	CNS-Hb
		100%	100%	100%	100%
	RCC	146	146	87	120
		66.4%	100.0%	71.3%	63.2%
	pNET	122	87	122	108
		55.5%	59.6%	100.0%	56.8%
	CNS-Hb	190	120	108	190
		86.4%	82.2%	88.5%	100.0%
Tumour type overlap	Base (N, %)	220	146	122	190
		100%	100%	100%	100%
	RCC + pNET + CNS-Hb	77	77	77	77
		35.0%	52.7%	63.1%	40.5%
	RCC + pNET	10	10	10	-
		4.5%	6.8%	8.2%	-
	RCC + CNS-Hb	43	43	-	43
		19.5%	29.5%	-	22.6%
	pNET + CNS-Hb	31	-	31	31
		14.1%	-	25.4%	16.3%
	RCC	16	16	-	-
		7.3%	11.0%	-	-
	pNET	4	-	4	-
		1.8%	-	3.3%	-
	CNS-Hb	39	-	-	39
		17.7%	-	-	20.5%

Table 38 Disease status by current treatment status

		Are you currently prescribed medication for your VHL-related cancer?		
		Base	**Yes**	**No**
Disease status	**Base (N%)**	220	61	159
		100.00%	100.00%	100.00%
	Progressive disease	58	21	37
		26.4%	34.4%	23.3%
	Stable disease	113	26	87
		51.4%	42.6%	54.7%
	Complete response	1	-	1
		0.5%	-	0.6%
	Partial response	8	4	4
		3.6%	6.6%	2.5%
	Don't know	40	10	30
		18.2%	16.4%	18.9%

Overall, 61 patients were currently prescribed treatment for their VHL-related cancer, and 159 patients were not. Of patients currently prescribed treatment, 42.6% (n=26) had stable disease, 34.4% (n=21) had progressive disease, and 6.6% (n=4) had partial response. For patients not currently prescribed treatment, 54.7% (n=87) had stable disease, 23.3% (n=37) had progressive disease, 0.6% (n=1) patients had a complete response, and 2.5% (n=4) patients had a partial response.

EQ-5D results

All patients

Overall, patients with VHL-related tumours had a mean EQ-5D score of 0.699, using the UK crosswalk. Patients with metastatic disease (n=16) had a mean EQ-5D score of 0.550 and patients without metastatic disease (n=195) had a mean EQ-5D score of 0.714 (Table 39).

Table 39 EQ-5D by metastatic status

		Do you have metastatic cancer?			
		Base	Yes	No	Don't know
EQ5D Crosswalk – UK (Hernandez-Alava et al)	Base (N, %)	220	16	195	9
		100%	100%	100%	100%
	Mean	0.699	0.550	0.714	0.638
	Min	-0.240	-0.240	-0.238	-0.154
	Max	0.988	0.987	0.988	0.892
	SD	0.27	0.35	0.25	0.37

Patients with metastatic disease

Sixteen of the 220 patients had metastatic disease. Among these, patients who classified themselves as having progressive disease (n=6) had a mean EQ-5D score of 0.412 and patients who classified themselves as having stable disease (n=4) had a mean EQ-5D score of 0.525. Full data presented in Table 40.

Table 40 EQ-5D by disease status in patients with metastatic disease

		Disease status			
		Base	Progressive disease	Stable	Don't know
EQ5D Crosswalk – UK (Hernandez-Alava et al)	Base (N, %)	16	6	4	6
		100%	100%	100%	100%
	Mean	0.550	0.412	0.525	0.705
	Min	-0.240	-0.240	0.096	0.381
	Max	0.987	0.740	0.987	0.801
	SD	0.35	0.43	0.41	0.16

Patients without metastatic disease

Mean EQ-5D by disease status, surgery history, number of tumours among patients without metastatic disease are presented in Table 41. Among patients without metastatic disease (n=195), patients who classified themselves as having

progressive disease (n=49) had a mean EQ-5D of 0.665, patients who classified themselves as having stable disease or complete or partial response (n=116) had a mean EQ-5D of 0.754. Mean EQ-5D scores of 0.665 and 0.754 are indicative of poor HRQoL as they are 33.5% and 24.6% lower, respectively, than the score of 1 which represents perfect health (these values would be age and gender corrected to the relevant populations they are sourced from and applied to, in the case of the cost-effectiveness analyses of this appraisal, these are performed in the cost-effectiveness model itself, described in section B.3.

Patients who had their most recent surgery within the last 6 months had a mean EQ-5D of 0.666 and patients who had their most recent surgery over 6 months ago had a mean EQ-5D of 0.705. Patients with more than one tumour type had a mean EQ-5D of 0.708 and patients with only one tumour type had a mean EQ-5D of 0.757.

Table 41 EQ-5D by disease status, among patients without metastatic disease

		Disease status			
		Base	Progressive disease	Stable disease, CR, PR	Don't know
EQ5D Crosswalk - UK (Hernandez-Alava et al)	Base (N%)	195 100.00%	49 100.00%	116 100.00%	30 100.00%
	Mean	0.714	0.665	0.754	0.642
	Min	-0.238	-0.138	-0.199	-0.238
	Max	0.988	0.987	0.988	0.985
	SD	0.25	0.25	0.21	0.35

CR: complete response; PR: partial response

B.2.10 Adverse reactions

Summary of adverse events information

- The MK-6482-004 study showed that belzutifan had a manageable safety profile in participants with VHL disease-associated RCC.
- All participants experienced at least 1 AE, and all experienced an AE that was assessed as related to study intervention by the investigator.
- Belzutifan was generally well tolerated. The proportion of participants who discontinued study intervention due to an AE was low (4 [6.6%] participants).
- Two deaths have occurred during the study, one due to suicide and one due to an AE (acute fentanyl toxicity) that was reported as not related to study drug by the investigator.

The adverse events reported in the MK-6482-004 study are summarised in Table 42, these are presented in more detail in Appendix F. As would be expected with the passage of time, the number of patients with adverse events increased with later data cut-off dates, and the proportion of patients reporting a higher grade of severity for adverse events (and treatment-related adverse events) as their highest severity adverse event experienced thus far also increased with later data cut-off dates.

Table 42 MK-6482-004 study summary of adverse events

Category	Belzutifan (N=61) n (%)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Number of adverse events	[redacted]	945	[redacted]	1260
Subjects with any adverse events	[redacted]	61 (100.0)	[redacted]	61 (100.0)
Subjects with any treatment-related adverse events	[redacted]	61 (100.0)	[redacted]	61 (100.0)
Subjects with any adverse events of CTCAE Grade 3 and Above	[redacted]	20 (32.8)	[redacted]	27 (44.3)
Subjects with any serious adverse events	[redacted]	11 (18.0)	[redacted]	18 (29.5)
Subjects with any treatment-related serious adverse events	[redacted]	3 (4.9)	[redacted]	4 (6.6)
Severity grade (Refer to NCI-CTCAE V 4.03) [1]				
• Mild (Grade 1)	[redacted]	10 (16.4)	[redacted]	8 (13.1)
• Moderate (Grade 2)	[redacted]	31 (50.8)	[redacted]	26 (42.6)
• Severe (Grade 3)	[redacted]	18 (29.5)	[redacted]	22 (36.1)
• Life Threatening (Grade 4)	[redacted]	1 (1.6)	[redacted]	3 (4.9)
• Death (Grade 5)	[redacted]	1 (1.6)	[redacted]	2 (3.3)
Related Severity grade (Refer to NCI-CTCAE V 4.03) [1]				
• Mild (Grade 1)	[redacted]	25 (41.0)	[redacted]	21 (34.4)
• Moderate (Grade 2)	[redacted]	27 (44.3)	[redacted]	29 (47.5)
• Severe (Grade 3)	[redacted]	9 (14.8)	[redacted]	11 (18.0)
Subjects with adverse events leading to death	[redacted]	1 (1.6)	[redacted]	2 (3.3)
Subjects with adverse events leading to treatment discontinued	[redacted]	2 (3.3)	[redacted]	4 (6.6)
Subjects with adverse events leading to dose reduced	[redacted]	9 (14.8)	[redacted]	10 (16.4)
Subjects with treatment-related adverse events leading to dose reduced	[redacted]	7 (11.5)	[redacted]	8 (13.1)
Subjects with adverse events leading to dose interrupted	[redacted]	26 (42.6)	[redacted]	26 (42.6)
Subjects with treatment-related adverse events leading to dose interrupted	[redacted]	14 (23.0)	[redacted]	13 (21.3)

B.2.11 Ongoing studies

There is currently an ongoing phase 2 single-arm study to evaluate the efficacy and safety of belzutifan monotherapy in participants with advanced pheochromocytoma/paraganglioma, pNET or VHL disease-associated tumours (the MK-6482-015 study) (40), the primary objective of the study is to evaluate the ORR associated with treatment with belzutifan, the estimated primary completion date for this study is 12-AUG-2026.

Additionally, a condition of the MHRA marketing authorisation for belzutifan in this indication is for MSD to set up and report on results from a prospective patient registry with the objective to further characterise efficacy and understand long term safety of belzutifan, particularly in VHL-associated RCC and CNS hemangioblastomas. The protocol for this prospective patient registry is currently being assessed by the MHRA.

B.2.12 Interpretation of clinical effectiveness and safety evidence

There are currently no active systemic interventions that are used (or recommended by NICE or funded on the NHS) for the treatment of adult patients with VHL disease who require therapy for VHL associated RCC, CNS hemangioblastomas, or pNETs, and for whom localised procedures are unsuitable or undesirable. Whilst the marketing authorisation states surgery is undesirable or unsuitable in these patients, it may be the only option. The relevant comparator in this indication in current clinical practice is therefore established clinical management without belzutifan.

Given the substantial unmet need for effective treatment for patients with VHL-tumours and the burden this disease places on patients and their families and carers, the outcomes from the MK-6482-004 study represent a step change in the treatment of this disease. The benefit the regulators perceived in this treatment is clear from the expansion of the primary tumours eligible for treatment, from that expected in RCC only to including also CNS hemangioblastomas and pNETs. Response rates and disease control rates across the ITT population represent a clinically profound treatment effect.

As demonstrated in the MK-6482-004 study, after a median follow-up of 37.7 months (range: 4.2 months to 46.1 months), the percentage of patients with RCC who had an objective response was 63.9% (95% confidence interval: 50.6% to 75.8%). Responses were also observed in patients with pancreatic neuroendocrine lesions (20 of 22 patients [90.9%]) and central nervous system hemangioblastomas (22 of 50 patients [44.0%]). The results show that belzutifan is able to offer significant additional benefit compared to current standard of care (i.e. no systemic therapy or surgical options) by effectively slowing the growth of tumours, thereby significantly delaying the time at which surgical resections become necessary or avoiding it completely in patients who have a durable complete response, and consequently decreasing the frequency of necessary surgical resections. Doing so is expected to greatly improve patients' quality of life in the long term.

Belzutifan also shows a tolerable and manageable safety profile in these patients. Treatment-related adverse events were mostly grades 1 and 2 and consisted of anaemia, fatigue, headache, and dizziness. The rate of discontinuations is low with 62.3% of patients in the MK-6482-004 study still continuing with a daily-administered treatment after a >3 year (median 37.7 months) follow-up period. Adherence to treatment is particularly important in the context of a chronic disease like VHL disease, in which patients may need lifelong, regular medical intervention from a young age.

The immediate effect of treatment with belzutifan on the course of disease in terms of surgical interventions patients undergo is especially notable, the number of VHL-disease associated surgical procedures patients underwent decreased from >300 in the group included in the study in the 10 year period prior to treatment initiation to <10 in the >3 year (median 37.7 months) follow-up period after initiation of treatment with belzutifan (by the 01-APR-2022 data cut-off date, as shown in Figure 8). The results observed in the MK-6482-004 study therefore indicate that the use of belzutifan may spare patients with VHL disease multiple surgeries, decrease their risk of loss of organ function (such as renal and/or pancreatic failure), and reduce their risk of death from metastatic RCC, metastatic pNET, or CNS hemangioblastomas.

The MK-6482-004 study included patients with VHL disease-associated RCC, and subgroup data are reported for those patients with RCC who also had CNS hemangioblastomas and patients with RCC who also had pNETS. Strictly speaking this differs from the marketing authorisation for belzutifan and its scope in this appraisal which does not specify that patients with CNS hemangioblastomas or pNETs need to also have RCC. However, the patients with CNS hemangioblastoma and pNETS in the MK-6482-004 trial with RCC are likely to be representative of patients with CNS hemangioblastoma and pNETS in general as patients with VHL disease have a tendency to develop tumours in multiple sites and organ systems with RCC being the most common form of tumours (13), and so it is likely that many patients with VHL disease-associated CNS hemangioblastoma and pNETS will also develop or have RCC. Indeed, the MK-6482-004 study found that, in the population of VHL RCC patients recruited, 82% had CNS hemangioblastomas and 33% had pNETs.

The number of patients with CNS hemangioblastoma and pNETs in the MK-6482-004 trial and included in the analyses were low at 50 and 22 patients, respectively. However, the results in terms of ORR in these groups of patients are still impressive, with responses to belzutifan treatment being long and durable to the extent that median DOR that was not reached (i.e. less than half the patients with response had lost response) as of the 01-APR-2022 database cut-off date for either of these subgroups. We do not have composite data for the 17 patients who had all three tumours, therefore the potential for such patients to benefit due to response in tumours in multiple organs is not known.

It is also worthwhile to note that the MK-6482-004 trial showed positive effects of treatment with belzutifan on VHL disease-associated retinal hemangioblastomas, adrenal lesions, endolymphatic sac tumours, epididymal cystadenomas and pancreatic lesions (not just limited to pNETs). While these tumours do not fall within the marketing authorisation for belzutifan, the fact that patients with VHL disease associated RCC, CNS hemangioblastomas, and pNETs are likely to also have some or all of these tumours due to the nature of VHL disease, and belzutifan appears to

have a positive effect on these also, is an additional benefit of treatment with belzutifan that should not be overlooked when considering its true value.

There is also misalignment between the GB marketing authorisation and the MK-6482-004 study in terms of the fact that the marketing authorisation is for patients "who require therapy" for the VHL-associated tumours, whereas the MK-6482-004 study's participant eligibility criteria specified that patients who had an immediate need for surgical intervention for tumour treatment were excluded. Also, the marketing authorisation is for patients "for whom localised procedures are unsuitable or undesirable" whereas this was not part of the participant eligibility criteria for the MK-6482-004 study. The MHRA granted this marketing authorisation in this indication based on the MK-6482-004 study and its results, i.e. based on which patients would benefit from treatment with belzutifan was demonstrated by the evidence from this very study. Such misalignments between marketing authorisation wording and supporting clinical trial patient population characteristics are not unusual for rare and highly specialised indications such as this one.

The MK-6482-004 study was single-arm trial that did not compare the treatment effect of belzutifan to UK SOC directly. While the comparison to UK-relevant SOC was made via a MAIC, the results from such techniques are inevitable associated with a lower degree of certainty than direct head-to-head results from a randomised clinical trial.

B. 3 Cost effectiveness

Summary of key cost effectiveness information

Objective:

- The purpose of this cost-effectiveness analysis is to assess the cost-effectiveness of belzutifan for patients with VHL-associated RCC, CNS Hb or pNET tumours who require treatment and for whom surgery is unsuitable or undesirable against the current SOC in the UK.

Model overview:

- A de novo cost-effectiveness model has been developed using a Markov cohort structure to estimate health outcomes and costs for belzutifan compared to the UK SOC from a National Health Service (NHS) and Personal Social Services (PSS) perspective.

- The MHRA conditional marketing authorisation is in patients who require therapy for VHL associated RCC, CNS Hb and pNET, and for whom localised procedures are unsuitable or undesirable. This population misalignment is a source of uncertainty in the economic analysis; however, the MK-6482-004 trial is the most appropriate source of data for this appraisal.

Base-case results and sensitivity analyses:

- Belzutifan is cost-effective across each of the cohorts when compared to the current SOC with an ICER of £42,997 per QALY in the VHL-RCC cohort, £33,490 per QALY in the VHL-CNS Hb cohort and £45,676 in the VHL-pNET cohort with the severity weighting incorporated into the ICER calculation.

- Probabilistic sensitivity analyses produce a 0.3%, 7.8% and 0.2% likelihood of cost-effectiveness for the VHL-RCC, VHL-CNS Hb and VHL-pNET cohorts respectively at a willingness-to-pay threshold of £30,000 per QALY gained when parameters are varied simultaneously.

Scenario analyses:

- The ICERs across the three cohorts remained largely stable to the parameters and assumptions tested in extensive scenario analyses with the exception of time horizon and treatment effect waning.

Cost effectiveness conclusions:

- Considering the complexities of modelling VHL, the rarity of the disease and paucity of published data, the cost-effectiveness analysis presented in this appraisal robustly demonstrates that belzutifan delivers clinically meaningful QALY and LY gains compared with SOC in patients with VHL-associated RCC, pNET, and CNS Hb. The full value of belzutifan is not captured in the economic model and therefore produces results that are marginally above the decision-making threshold. The economic evaluation highlights that belzutifan is a step-change in the management of VHL, a rare disease with high unmet need.

B.3.1 Published cost-effectiveness studies

There are no published economic models evaluating the cost-effectiveness or budget impact of belzutifan

An SLR was conducted to identify published cost-effectiveness studies for belzutifan or other VHL therapies. Searches were performed in July 2020 and were subsequently updated in July 2022 to cover the period of July 2020 to July 2022. The cost-effectiveness SLR was designed and executed in line with NICE guidance and was run as part of a broader SLR designed to identify (i) RCTs and non-RCTs, (ii) utility data, and (iii) cost and resource use data. Further information on the SLR methodology, search strategy, and results is provided in the following section: *Appendix G: Published cost-effectiveness studies*.

Given the rarity of disease and absence of approved therapies with the exception of belzutifan, the SLR did not identify any economic models evaluating the cost-effectiveness or budget impact of any treatments for VHL, including belzutifan.

B.3.2 Economic analysis

In the absence of published economic analyses, a *de novo* cost-effectiveness analysis was developed for this appraisal. The economic model uses a Markov cohort structure to estimate health outcomes and costs for belzutifan compared to the UK standard of care (SOC) in patients with VHL who require therapy for VHL RCC, CNS Hb, or pNET, and for whom localised procedures are unsuitable or undesirable.

Patient population

The model population is aligned to the marketing authorisation for belzutifan

The economic analysis considers adult patients (aged 18 years or older) with VHL disease who require treatment for: (i) VHL-associated RCC, (ii) VHL-associated CNS Hb, or (iii) VHL associated pNET. In line with the characteristics of patients in the key trial informing the economic analysis (see *B.2 Clinical effectiveness*), patients enter the model at an age of 41.0 years, with a mean (SD) weight of 79.7 kg (23.4 kg) and a mean (SD) body surface area of 1.9 m^2 (0.3 m^2), and with 47.5% female patients.

The population assessed in the economic analysis is aligned to the population as specified in the marketing authorisation for belzutifan, with clinical evidence informed by the MK-6482-004 trial (37) and real-world data for SOC. Based on data from the MK-6482-004 trial in support of the license application for belzutifan, the MHRA decided to specify eligibility to adult patients with VHL “who require therapy” for VHL associated RCC, CNS Hb, or pNET, “and for whom localised procedures are unsuitable or undesirable”. This contrasts with the inclusion criteria of the MK-6482-004 trial, which did not specifically require participants to be considered unsuitable or undesirable for localised procedures.

Treatment decision point

The patient population stipulated by the MHRA label identifies a specific VHL patient group “who require therapy” and “for whom localised procedures are unsuitable or undesirable”. This patient population have exhausted alternative options to control VHL tumour manifestations and are at the “end of the road”, they must have sufficient organ function as per the inclusion criteria for the trial.

In routine clinical practice, the decision point for a patient meeting the criteria of belzutifan eligibility would have three options: 1) surgery that is unsuitable or undesirable because it results in loss of organ function, 2) active surveillance to monitor a tumour that is above 3cm (RCC) or 2cm (pNETs) and therefore there is an increased risk of metastatic disease and/or other symptoms of tumour burden (particularly in CNS Hb tumours), or 3) belzutifan. At this point on the disease/treatment pathway occurs, patients and their disease progression will have been carefully monitored over many years. Any treatment decision would be made very carefully between a patient and their treating clinician and for some, this decision may not be immediate. A simplifying assumption was taken to initiate the model at the point this treatment decision is 'enacted', i.e. when the patient is faced with a choice between receiving surgery, routine surveillance, or belzutifan.

Model structure

A de novo Markov cohort model was developed to estimate the cost-effectiveness of belzutifan versus UK SOC in patients with VHL who require therapy for VHL associated RCC, CNS Hb, or pNET, and for whom localised procedures are unsuitable or undesirable

A *de novo* cost-effectiveness model was developed using a Markov cohort structure to estimate health outcomes and costs for belzutifan compared to the UK SOC in patients with VHL who require therapy for VHL RCC, CNS Hb, or pNET, and for whom localised procedures are unsuitable or undesirable. Markov models feature an explicit structural linkage between intermediate health states and death and are therefore commonly used in economic analyses for chronic diseases with critical outcomes of relevance other than mortality.

The Markov model was initially built when the expectation for the marketing authorisation was for VHL-associated RCC only. Following the granting of the market authorisation, featuring both an expansion and restriction compared to the population assessed in the belzutifan pivotal trial, the model was adapted to reflect the updated eligible patient population. The model captures the most relevant clinical events in VHL disease i.e., surgery resulting in loss of organ function, development of metastatic disease, and death.

The model structure consists of five mutually exclusive health states aligned with important outcomes in the disease progression of VHL-associated RCC, CNS Hb, and pNET.

i. Pre-surgery (model entry)

ii. Surgery tunnel state, patients remain in this state for one week and receive surgery that results in loss of organ function (for RCC and pNET primary tumours) or results in brain injury (for CNS Hb primary tumours) as a 'last resort' intervention.

iii. Event-free after surgery state, which represents patients who have undergone one surgery, have not developed metastatic disease, and are still alive:

iv. Metastatic disease

v. Death

Each primary tumour site i.e. the VHL-RCC, VHL-CNS Hb, or VHL-pNET tumour with the greatest burden on the patient is modelled as a separate cohort using the same model structure. Reflective of the natural history of VHL disease, patients in each primary tumour site cohort may also have one, two or all three tumour types simultaneously. Although the incidence of non-primary tumours, and therefore related surgeries, is captured in the model, the additional burden on costs and quality of life of having multiple tumour manifestations simultaneously is not specifically captured.

Patients transition from the pre-surgery health state to either:

- Surgery then event-free after first surgery
- Metastatic disease
- Death

Patients transition from the event-free after surgery health state to either:

- Metastatic disease
- Death

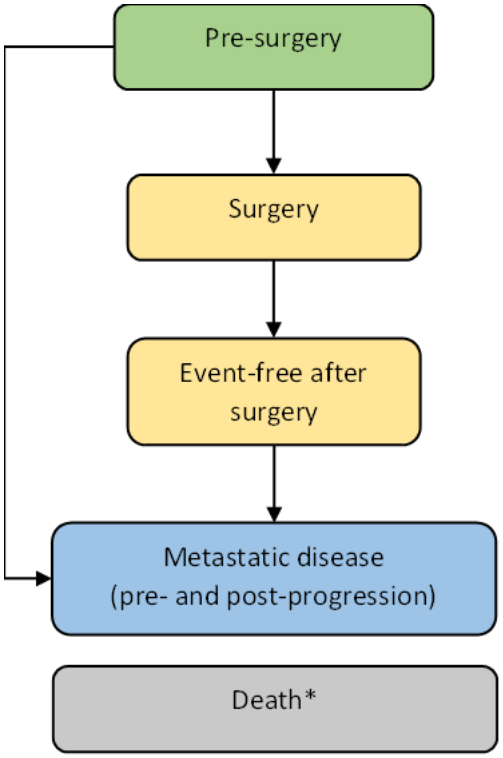
The surgery health state refers specifically to surgery relating to the primary tumour type. There is no surgery-specific health state for surgeries at non-primary tumour sites (e.g. a patient in the RCC cohort who has a subsequent pNET surgery is not captured in a specific health state transition). The cost and health implications of surgeries for non-primary tumours as well as their associated complications were reflected as per-event costs and QALY decrements applied on incidence of each non-primary tumour surgery. This approach to modelling primary and non-primary tumours differently is used to reflect typical disease progression in VHL: patients may require surgeries at multiple sites but clinicians focus on the highest risk tumour site, as this is where tumours are likely to be the most progressed. Given the incidence of non-primary surgeries is not reflected by explicit state transitions within the model, the cumulative health impact of having undergone multiple non-primary surgeries is therefore not captured. This is a conservative approach as evidence from MK-6482-004 has shown that belzutifan reduces tumour size, and would therefore have metastases and mortality benefits, not only in primary tumours.

The event-free after surgery health state corresponds to time spent post-surgery during which a patient is at risk of metastases, short-term and long-term consequences of surgery.

The metastatic disease health state is included to reflect the clinical pathway of patients with VHL-associated tumours. In each tumour type cohort, transitions to the metastatic disease state are assumed to be due to either RCC or pNET as the origin tumour (CNS Hb tumours do not metastasize), consistent with findings from the VHL Natural History Study.

The model's health states and available transitions is presented in Figure 18. Note that the pre-surgery health state describes patients who have not had surgery since belzutifan trial initiation, and for the purposes of the economic analysis, the treatment decision point. The majority of participants in the MK-6482-004 trial had multiple surgeries prior to trial initiation (mean 5.5).

Figure 18 Model schematic for the economic evaluation of belzutifan in VHL associated RCC, pNET, or CNS Hb



*Transitions to death are possible from all health states. Arrows to the death state are omitted from the diagram for simplicity.
Note: Analogous Markov cohort structures are used for each of the three tumour-specific populations (VHL-associated RCC, CNS Hb and pNET). In each of these populations, subsets of patients also have one or both of the other two tumour types. In the Markov model, the surgery states refer specifically to surgeries corresponding to the primary tumour type for each population. Costs and QALY decrements due other tumour types are modelled separately for each population and layered onto the costs and QALYs that are modelled accordingly to patients' Markov state residency over time.
CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; QALY: quality adjusted life year; VHL: Von Hippel Lindau

One general limitation of Markov models is their memoryless property, by which history of patients' movements through the model structure is not tracked over time. Transition probabilities, utilities, and costs at each model cycle therefore depend only on a patient's current health state, the cycle number, and time-constant factors instead of reflecting patient's past sequence of events or time spent in an intermediate health state. To mitigate this limitation while accounting for practical data constraints, functionality was incorporated into the Markov structure to track the occurrence of certain important clinical events. Specifically, surgery and event-free

after surgery health states were used to capture perioperative mortality and surgical complications to which costs and disutilities could be applied accordingly. Separate health states were also defined to track patients' treatment discontinuation status over time (i.e., on-treatment vs. off-treatment).

As described in *B.2 Clinical effectiveness*, the primary outcome assessed in the MK-6482-004 trial for belzutifan was ORR in VHL-associated RCC tumours per RECIST 1.1, with best overall response categorised as CR, PR, SD, or PD. Secondary outcomes included DOR (i.e. time from the first documented evidence of CR or PR until either disease progression or death from any cause), TTR, PFS, and TTS in VHL-RCC. Secondary outcomes also included ORR, DOR, TTR, PFS, and TTS evaluated for non-RCC VHL-associated tumours (CNS Hb, pNET, retinal Hb, adrenal endolymphatic sac tumours, and epididymal cystadenomas). Responder status is understood to have implications on patients' health-related quality of life (HRQoL), as patients with larger tumour burdens may be more likely to experience direct symptoms from their tumours (particularly in the case of CNS Hb) and may have greater anxiety about their VHL disease. Each non-metastatic health state (i.e., pre-surgery, surgery, and event-free after surgery) therefore allows three different categories of objective response for primary tumours (e.g., complete response [CR], partial response or stable disease [PR/SD], and progressive disease [PD]), to reflect a distinct utility value. In the model, this is implemented as an average utility value weighted by response status.

Of note, PD per RECIST 1.1 criteria does not necessarily indicate a need for surgery, nor does the occurrence of surgery necessarily imply PD. A tumour size threshold [for RCC and pNET] or the manifestation of symptoms [for CNS Hb] along with other clinical factors collectively determine the need for surgery in patients with VHL-associated tumours. Consequently, the model allowed for the possibility that some patients have a PD while residing in the pre-surgery state, and that some patients have CR, PR, or SD in the surgery and event-free after surgery states.

Table 43 summarises the features of the economic analysis.

Table 43 Features of the economic analysis

	Current evaluation	
Factor	**Chosen values**	**Justification**
Time horizon	Lifetime	To fully capture differences in costs and outcomes between belzutifan and SOC, in-line with the NICE reference case.
Cycle length	1 week	Allows for precise calculation of drug acquisition and administration costs.
Half-cycle correction	Yes	A half-cycle correction is applied to costs and effectiveness, to reflect the fact that patients may transition between health states at any point during a cycle. The half-cycle correction is not applied to cost components that are incurred at the beginning of a cycle, including belzutifan acquisition and administration costs (recurring costs starting from week 0) and AE-related costs (applied as a one-time cost at week 0).
Discounting for costs and QALYs	3.5% for costs and health outcomes. Given the chronic nature of the disease we test 3.5% for costs and 1.5% for health outcomes.	Consistent with NICE reference case
Intervention	Belzutifan	Belzutifan is the topic of this appraisal
Comparator	SOC without belzutifan defined according to each cohort: • For RCC and pNET cohorts, surgery resulting in loss of organ function in 90% of patients with the remaining 10% receiving symptom management • For CNS Hb cohort, surgery with risk of brain injury in 50% of patients, in the remaining 50% where tumour location renders it inoperable they receive symptom management but face the same risk of brain	The marketing authorisation is for patients “who require therapy” for their disease. Therefore, active surveillance regimens are not relevant comparators. A patient for whom active surveillance is appropriate does not require therapy. The marketing authorisation is for patients “for whom localised procedures are unsuitable or undesirable” therefore, minimally invasive treatments, surgery, and radiotherapy that are organ-sparing are not relevant comparators. If a patient can have such a procedure, they should have such a procedure. The surgery received by the respective cohorts results in either loss of organ function or brain injury in line with clinical expert opinion of the

	Current evaluation	
Factor	**Chosen values**	**Justification**
	injury due to tumour size/location.	definition of “unsuitable or undesirable localised procedures”.
Clinical effectiveness (belzutifan)	Health state TPs and surgery incidence from the MK-6482-004 trial (37)	Clinical effectiveness is based on data collected in a single-arm trial. Due to the rarity of VHL, sourcing sufficient patient numbers to power a placebo-controlled RCT is challenging and may be considered unethical in a rare disease. Thus a single-arm trial was used to inform the required efficacy data for patients receiving belzutifan.
Clinical effectiveness (SOC)	Initial risk of surgery is estimated based on the definition of the target population and reflects the treatment decision point at which belzutifan becomes a treatment option. The remaining TPs are sourced from the VHL Natural History Study (adjusted using the Optum Clinformatics Data Mart claims study) and analysis of the pre-treatment period of the MK-6482-004 trial (37, 41)	MK-6482-004 is a single-arm trial. As such, comparator data were sourced from a large natural history study (VHL Natural History Study) and a large real-world study in patients closely aligned to the MK-6482-004 population (Optum Clinformatics Data Mart claims study). This approach has been accepted by NICE in appraisals for other rare conditions e.g. NICE HST14 Metreleptin for treating lipodystrophy [ID861]. (42). Data from the Optum Clinformatics Data Mart claims study were included to provide an adjustment to the rates of surgery (for 10% of RCC & pNET cohorts) and rates of metastatic disease from the VHL Natural History Study, to account for the potentially elevated SOC in the VHL Natural History Study (see Clinical efficacy: section for more details).
Safety (belzutifan)	AE rates sourced from the MK-6482-004 trial. The model considers Grade ≥3 AEs that occurred in the MK-6482-004 trial with a frequency of ≥5% or Grade ≥3 TRAEs in all patients. Risks of surgery complication are sourced from the Optum Clinformatics Data Mart claims study (41).	Grade ≥3 AEs at a frequency of ≥5% are considered in the cost-effectiveness model as they are expected to impact cost and utility and a 5% threshold has been accepted in previous NICE appraisals, and is considered standard practice. Grade ≥3 TRAEs at a frequency of ≥0% are included in the cost-effectiveness model as they are considered to be attributable to treatment, and expected to incur a cost and utility impact. This definition for considering AEs is standard practice for NICE appraisals and has

	Current evaluation	
Factor	**Chosen values**	**Justification**
		been previously accepted by NICE (e.g. NICE TA531 (43)). Risks of surgery complication are adjusted in line with clinical expert opinion to align with the licensed population.
Safety (SOC)	Surgery complication risks sourced from the Optum Clinformatics Data Mart claims study (41). These surgery complication risks are adjusted in line with clinical expert opinion to align with the licensed population.	As VHL is a rare disease, MK-6482-004 is a single arm trial. As such, comparator data were sourced from a large natural history study and a large real-world study in patients closely aligned to the MK-6482-004 population. This approach has been accepted by NICE in appraisals for other rare conditions e.g. NICE HST14 Metreleptin for treating lipodystrophy [HST14] (42). Surgical complication risk rates are sourced from the Optum Clinformatics Data Mart claims study. To align with the licensed population for whom localised procedures are unsuitable or undesirable, these risk rates were adjusted in line with clinical expert opinion.
Source of utilities	Health state utilities in the pre-surgery, surgery and event-free after surgery states were calculated as a weighted average of EQ-5D utility values for the CR (sourced from the KEYNOTE-564 trial), PR/SD (sourced from the VHL RW QoL Disease Burden Study (44); PR and SD assumed to be equal), and PD (sourced from the VHL RW QoL Disease Burden Study (44)) response levels. Disutilities relating to short-term and long-term consequences of surgery were applied to the event-free-after surgery state. For the metastatic disease health state, EQ-5D-	Since no HRQoL data were collected in the MK-6482-004 clinical trial nor VHL Natural History Study, the VHL RW QoL Disease Burden Study and KEYNOTE-564 were used as alternative sources to provide EQ-5D utility data for each health state in the economic model. The use of EQ-5D data is in line with the NICE reference case. Utility values for the CR response level were previously accepted by NICE in TA830 with a population of patients with non-metastatic RCC following full or partial nephrectomy which is similar to the VHL-RCC cohort of this appraisal (45).

	Current evaluation	
Factor	**Chosen values**	**Justification**
	based utility values were sourced from the KEYNOTE-426 trial.	
Source of costs	NHS Reference Costs, PSSRU, BNF and Monthly Index of Medical Specialties (cost year 2020/2021).	In line with NICE reference case.

*No previous evaluations have been performed in VHL.
AE: adverse event; BNF: British National Formulary; HRQoL: health related quality of life; NHS: national health service; NICE: national institute for health and care excellence; PSSRU: Personal Social Services Research Unit; QALY: Quality adjusted life year; QoL: Quality of Life; RW: Real-World; SOC: standard of care; TP: transition probability; TRAE: treatment-related adverse event; VHL: Von Hippel Lindau

Intervention technology and comparators

Belzutifan is compared with SOC in patients who require therapy for VHL associated RCC, CNS Hb, or pNET, and for whom localised procedures are unsuitable or undesirable

Intervention: belzutifan

The recommended dose of belzutifan is 120 mg administered orally once daily until disease progression or unacceptable toxicity (5) (see Appendix C1.1 SmPC for more information). Dose adjustments and modifications have been reflected as per the MK-6482-004 trial data as described in the *Time to treatment discontinuation* section.

Comparator: SOC

There are currently no approved systemic treatments specifically for VHL other than belzutifan. The comparator in the economic analysis is therefore SOC, defined as established clinical practice without belzutifan in line with the final scope. Also per the final scope, localised procedures are unsuitable or undesirable for the patient population eligible for belzutifan. Clinical experts consulted for this appraisal described the trade-off between the risks associated with localised procedures and versus the risks of metastatic and/or symptomatic disease that arise in the absence of urgently needed surgical procedures. These experts therefore define the patient population for whom localised procedures are unsuitable or undesirable as those

patients who have exhausted all options and for whom localised procedures would be a ‘last resort’ and likely result in loss of organ function with extremely poor outcomes (3).

For patients who require therapy and for whom localised procedures are unsuitable or undesirable, SOC interventions they receive includes localised procedures; however, these will not be able to preserve organ function and so result in problematic sequalae that can have a negative impact on HRQoL (3).

- For RCC tumours, the localised procedures that are no longer capable of preserving organ function are any further partial/full bilateral nephrectomies after which patients will require renal replacement therapy. Following this procedure, no further surgeries related to RCC tumours would take place or be considered as organ function would be lost.

- For pNET, the localised procedures that are no longer capable of preserving organ function are Whipple procedures/ pancreatectomies and splenectomies. Laparoscopic surgery (e.g. distal pancreatectomy) is considered if the tumour is in the pancreas tail or peripheral pancreas. The Whipple procedure is considered if the tumour is in the pancreas head, or if the tumour is at least 2 centimetre in diameter and not amenable to distal pancreatectomy. Following this procedure, no further surgeries related to pNET would take place as organ function would be lost.

- For CNS Hb, these localised procedures are only considered where the tumour is peripherally located in the cerebellum; however, with or without the unsuitable/undesirable localised procedure the outcome is likened to symptoms of patients with motor neurone disease or following major stroke either because of the surgery risk or tumour burden (3).

In routine clinical practice in the UK, for such patients “who require therapy” these unsuitable/undesirable localised procedures would take place immediately as these patients have reached the point in which active surveillance or watchful waiting is no longer an option due to unbearable symptomatic disease or unacceptable risk of metastases. For VHL-RCC and VHL-pNET cohorts, immediate surgery is assumed

for 90% of patients. For VHL-CNS Hb, immediate surgery is assumed for 50% of patients; however, the outcomes associated with surgery is assumed for 100% of the cohort due to tumour burden creating neurological disability for the remaining 50% not operated on, which is assumed to have similar impact on HRQoL as the serious complications from CNS surgery. See *Cost of surgery and complications* for further details on how the costs are applied to reflect this.

Following immediate surgery, efficacy inputs for the SOC arm were informed by data from the VHL Natural History Study (see Section B.2.9 for more information on the VHL Natural History Study) (46). The VHL Natural History Study included a large patient population (n=247) and similar eligibility criteria to the MK-6482-004 trial, enabling reasonable comparability to the belzutifan eligible population. To ensure sufficient matching between the VHL Natural History Study and the MK-6482-004 trial populations, the VHL Natural History Study cohorts were reweighted to match the corresponding patient samples from MK-6482-004 trial using the MAIC method, an inverse propensity weighting method which allows the logistic regression for propensity score to be estimated without individual patient data in one of the populations, as described in greater detail in *B.2.9 Indirect and mixed treatment comparisons*.

- Similar to the belzutifan arm, the VHL Natural History Study population informing the SOC arm was not limited to only those patients for whom localised procedures were unsuitable or undesirable. Therefore, the rationale for using the MK-6482-004 trial to inform the belzutifan arm described in the Patient population section also applies to the use of the VHL Natural History Study in the SOC arm.

- The VHL Natural History Study collected data from US-based centres of excellence and patients in the study may therefore have received a different SOC compared to standard UK clinical practice. It is expected that patients treated at these sites had better access to surgery, and as a result, higher rates of surgery and therefore lower rates of metastasis were observed than would be expected in UK clinical practice. To account for this, the hazards of 1) surgery from the pre-surgery health state and 2) metastatic disease from the

pre-surgery and event-free after surgery states were adjusted in the model using real-world metastatic disease rates data from an additional study (Optum Clinformatics Data Mart claims study) of VHL patients treated in a setting more reflective of real-world practice. This adjustment is described further in *Aligning risk of surgery and metastatic disease to real-world SOC*. Of note, the Optum study could not be used in place of the VHL Natural History study as it did not have the breadth of availability of matching variables to the MK-6482-004 trial.

B.3.3 Clinical parameters and variables

Clinical efficacy: transition probabilities

Overview of data sources

MK-6482-004 trial data and real-world evidence have been used to estimate the clinical parameters used in the economic model. A summary is provided in Table 44 below.

MK-6482-004 trial

TPs from the pre-surgery → surgery state for belzutifan are informed by TTS data for the VHL-RCC cohort from the MK-6482-004 trial (37). For the VHL-pNET cohort, the pre-treatment period from the MK-6482-004 trial is used to inform the TP from the pre-surgery → surgery state for the 10% who do not receive immediate surgery in the SOC arm (further described in *Transition from the pre-surgery and event-free after surgery states* below). The baseline characteristics from the MK-6482-004 trial were also used to reweight the VHL Natural History Study population to match the eligibility criteria of the trial for the SOC arm as described in section B.2.9.

VHL Natural History Study

Given the single-arm design of the MK-6482-004 trial, clinical efficacy data for SOC were primarily sourced from external data collected in the VHL Natural History Study, which informed the following transitions in the model: (46):

- For the 10% of the VHL-RCC cohort who do not receive immediate surgery in the SOC arm, TPs from pre-surgery → surgery. The corresponding TPs for the

VHL-pNET and VHL-CNS Hb cohorts are informed by the pre-treatment period from the MK-6482-004 trial, described below

- TPs from pre-surgery → metastatic disease and event-free after surgery → metastatic disease in all cohorts
- TPs from pre-surgery → death and event-free after surgery → death in all cohorts

Although the inclusion criteria limited the study sample to US and Canadian patients, the VHL Natural History Study included a large study sample (308 patients) followed up for a mean of 8.75 years (46). This study was designed to generate model inputs for the SOC arm following immediate surgery that would allow for a balanced comparison against belzutifan-treated patients with VHL-associated RCC in the MK-6482-004 trial. Based on the data fields available in the VHL Natural History Study, it was feasible to apply inclusion/exclusion criteria analogous to those used in the trial to identify eligible patients with VHL-associated RCC at the patient-level index date. However, it was not feasible to identify whether patients in the VHL Natural History Study had CNS Hb or pNET tumours on the patient-level index date. The criteria for identifying the CNS Hb and pNET cohorts of the MK-6482-004 trial could not be well-replicated within the VHL Natural History Study. Given this limitation, TTS for the proportion of SOC patients who do not receive immediate surgery was estimated using a retrospective analysis of the pre-treatment period of the MK-6482-004 trial.

Retrospective analysis of the MK-6482-004 trial pre-treatment period

History of surgery prior to trial initiation was collected retrospectively based on patient medical records collected in the MK-6482-004 trial (i.e. before patients began receiving belzutifan) (37) and this is used as a proxy for outcomes associated with the SOC arm. This “pre-treatment” data was used to inform the TP for pre-surgery → surgery in the SOC arm for the proportion who do not receive immediate surgery in the VHL-pNET cohort and following treatment effect waning in the belzutifan arm for the VHL-CNS Hb and VHL-pNET cohorts. Parametric survival models were fitted to patient-level data on time (looking backwards) to the most recent primary tumour surgery prior to belzutifan initiation in patients with VHL-CNS Hb and VHL-pNET

tumours in the MK-6482-004 trial. This retrospective analysis was also used to inform the incidence rate (in the SOC arm) and distribution of non-primary-VHL-related tumour surgeries (assumed equal in both arms) for the VHL-CNS Hb and VHL-pNET cohorts.

Whilst the MK-6482-004 pre-treatment analysis provides a robust source of data to inform the parameters described above for the VHL-CNS Hb and VHL-pNET cohorts, as per the eligibility criteria of the trial, patients were (by definition) alive and metastases-free prior to belzutifan initiation. Hence, the pre-treatment period data from MK-6482-004 could not be used to estimate the TP from pre-surgery → metastatic disease nor event-free after surgery → metastatic disease or death in these cohorts for the SOC arm. Instead, the TPs from pre-surgery and event-free after surgery → metastatic disease or death were estimated based on data collected in patients in the VHL Natural History Study who had a history of CNS Hb and pNET prior to the patient-level index date (although these tumours may have not been present on the index date).

The methods used to estimate transition probabilities described above are summarised in Table 44.

Table 44 Summary of clinical parameters sourced from the VHL Natural History Study and the pre-treatment period data from the MK-6482-004 trial.

Parameter	Source of parameter by model cohort			Report section(s) describing estimation of parameter
	VHL-RCC	VHL-CNS Hb	VHL-pNET	
Transition probability parameters				
• **Parameter estimates for pre-surgery→surgery** (for the proportion who do not receive immediate surgery in the SOC arm and after treatment effect waning in the belzutifan arm)	Parametric multistate modelling of patient-level data from the reweighted Natural History Study sample (effective sample size = 92.2)	Parametric modelling of pre-treatment period data from the MK-6482-004 trial (subgroup with CNS Hb at baseline visit)*	Parametric modelling of pre-treatment period data from the MK-6482-004 trial (subgroup with pNET at baseline visit)	Section: *Transitions from the pre-surgery and event-free after surgery states*.
• **Parameter estimates for pre-**		Parametric multistate	Parametric multistate	

surgery→metastatic disease (in the SOC arm and after treatment effect waning in the belzutifan arm)		modelling of patient-level data from the reweighted Natural History Study sample (subgroup with history of CNS Hb; effective sample size = 37.9)	modelling of patient-level data from the reweighted Natural History Study sample (subgroup with history of pNET; effective sample size = 60.4)	
• **Parameter estimates for pre-surgery→death** (assumed equal in both arms)				
• **Parameter estimates for event-free after surgery→metastatic disease** (in the SOC arm and after treatment effect waning in the belzutifan arm)	Parametric multistate modelling of patient-level data from the reweighted Natural History Study sample, based on the subset of patients who had a renal surgery after their patient-level index date (effective sample size = 75.7)	Assumed equal to pre-surgery→ metastatic disease, due to the small sample size of the reweighted Natural History Study sample (subgroup with history of CNS Hb) when restricted to those who had a CNS Hb surgery after their patient-level index date	Assumed equal to pre-surgery→ metastatic disease, due to the small sample size of the reweighted Natural History Study sample (subgroup with history of pNET) when restricted to those who had a pNET surgery after their patient-level index date	Section: *Transitions from the pre-surgery and event-free after surgery states.*
• **Parameter estimates for event-free after surgery→death** (in the SOC arm and after treatment effect waning in the belzutifan arm)		Assumed equal to pre-surgery→ death, due to the same limitation noted above	Assumed equal to pre-surgery→ death, due to the same limitation noted above	
Incidence and distribution of surgeries for non-primary tumours				
• **Incidence rate of surgeries (events/person-week) for non-primary VHL-related tumours** (in the SOC arm and after treatment effect waning in the belzutifan arm)	Calculated using the reweighted Natural History Study sample (effective sample size = 92.2)	Calculated using pre-treatment period data from the MK-6482-004 trial (subgroup with CNS Hb at baseline visit)	Calculated using pre-treatment period data from the MK-6482-004 trial (subgroup with pNET at baseline visit)	Section *Surgery incidence*
• **Distribution of surgeries for non-primary VHL-related**	Calculated using the reweighted Natural History Study sample	Calculated using pre-treatment period data from the MK-6482-	Calculated using pre-treatment period data from the MK-6482-	Section *Surgery incidence*

tumours (assumed equal in both arms)	(effective sample size = 92.2)	004 trial (subgroup with CNS Hb at baseline visit)	004 trial (subgroup with pNET at baseline visit)	
Other model parameters				
• **Distribution of metastatic disease cases by origin tumour in each model cohort** (assumed equal in both arms)	Calculated using the reweighted Natural History Study sample (effective sample size = 92.2)	Calculated using the reweighted Natural History Study sample (subgroup with history of CNS Hb; effective sample size = 37.9)	Calculated using the reweighted Natural History Study sample (subgroup with history of pNET; effective sample size = 60.4)	Section *Transitions from metastatic disease to death*

*Note: Under the SOC arm in the VHL-CNS Hb cohort, all patients are assumed to have either immediate CNS Hb surgery or experience equivalent sequelae (i.e., the same costs and disutilities of surgical complications) to reflect the severity of inoperable CNS Hb. Patients who experience the equivalent sequalae of CNs Hb surgery are modelled as entering the surgery state, but do not incur the costs of the surgical procedure itself.

Optum Clinformatics Data Mart claims study

The Optum Clinformatics Data Mart claims database study conducted by Jonasch et al. (2022) collected real-world data from 160 patients with VHL from a wide geographic area in the US (34, 41). Patients with VHL-RCC were identified within the claims database using a customised algorithm. They were then matched to population controls without VHL disease or RCC, and comparisons were made with the matched cohort to conduct analysis on the prevalence, treatment patterns and healthcare resource use (HCRU) associated with VHL. Analysed outcomes included the rates of metastasis, surgery, and short- and long-term surgical complications for all three primary tumour sites.

In contrast to the Optum study, patients included in the VHL Natural History Study were managed by a multidisciplinary team with broad experience in detection, prevention and treatment of hereditary genitourinary malignancies in line with the intended treatment-eligible population in MK-6482-004. It is expected that patients at the NCI received an elevated SOC and thus had improved access to surgery and lower rates of metastatic disease due to improved control of tumours relative to real-world treatment patterns expected for many patients with VHL managed in the UK. To account for this difference in SOC, data from the Optum study were used to adjust surgery rates downwards and metastatic disease rates upwards to better align

with real world practice of VHL treatment. The adjustments are described in further detail in *Aligning risk of surgery and metastatic disease to real-world SOC*.

Transitions from the pre-surgery and event-free after surgery states

This section describes how TPs estimated for each transition from the pre-surgery and event-free after surgery states for each of the 3 VHL cohorts (i.e. RCC, CNS Hb, pNET).

Parametric models were fitted to time-to-event data to estimate the cause-specific hazards of each transition starting from the pre-surgery state (i.e., pre-surgery → surgery, pre-surgery → metastatic disease, and pre-surgery → death) and event-free after surgery state (i.e., event-free after surgery → metastatic disease, and event-free after surgery → death) over time within the belzutifan and SOC arms. In each weekly cycle of the model, the probability of each of these transitions (as well as the composite probability of any transition from the pre-surgery state) was calculated as a function of all three cause-specific hazards.

Cause-specific hazards for transitions from the pre-surgery health state and hazards for transitions from the event-free after surgery health states are summarised below in Table 45 and Table 46 respectively. Further details on how the derivation of each TP by tumour site and treatment arm are provided in the proceeding sections.

Table 45 Cause-specific hazards and data sources for transitions from the pre-surgery state, by model cohort and treatment arm

Cohort / Treatment arm	Cause-specific hazards (yearly exponential rates)			Sources & estimation approaches		
	Pre-surgery → surgery	Pre-surgery → metastatic disease	Pre-surgery → death	Pre-surgery → surgery	Pre-surgery → metastatic disease	Pre-surgery → death
VHL-RCC						
Belzutifan	0.03692	0.000312	0.00364	MK-6482-004 trial (data cut-off date: 01 April 2022)	MK-6482-004 trial (data cut-off date: 01 April 2022)	Assumed equal to SOC, except in deaths attributable to CNS Hb (see VHL-CNS Hb cohort)
SOC	0.25324*	0.00208	0.00624	VHL Natural History Study (2021)	VHL Natural History Study (2021)	VHL Natural History Study (2021)
VHL-CNS Hb						
Belzutifan	0.0052	0.000156	0.00728	MK-6482-004 trial (data cut-off date: 01 April 2022)	MK-6482-004 trial (data cut-off date: 01 April 2022)	HR of belzutifan vs. SOC assumed equal to pre-surgery to surgery
SOC	0.10504*	0.00312	0.01456	Pre-treatment period data from MK-6482-004 trial	VHL Natural History Study (2021)	VHL Natural History Study (2021)
VHL-pNET						
Belzutifan	0.000312	0.00026	0.00624	MK-6482-004 trial (data cut-off date: 01 April 2022)	MK-6482-004 trial (data cut-off date: 01 April 2022)	Assumed equal to SOC, except in deaths attributable to CNS Hb (see VHL-CNS Hb cohort)
SOC	0.00884*	0.00676	0.01092	Pre-treatment period data from MK-6482-004 trial	VHL Natural History Study (2021)	VHL Natural History Study (2021)

*For the pre-surgery → surgery transition in the VHL-RCC and pNET cohorts this cause-specific hazard is used for the remaining 10% who do not receive immediate surgery in the SOC arm. In the VHL-CNS Hb cohort all patients are assumed to have the outcomes from

surgery, therefore this cause-specific hazard is only used following treatment effect waning in the belzutifan arm.
CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Table 46 Cause-specific hazards and data sources for transitions from the event-free after surgery state, by model cohort and treatment arm

Cohort / Treatment arm	Cause-specific hazards (yearly exponential rates)		Sources & estimation approaches	
	Event-free after surgery → metastatic disease	Event-free after surgery → death	Event-free after surgery → metastatic disease	Event-free after surgery → death
VHL-RCC				
Belzutifan	0.000468	0.00728	MK-6482-004 trial (data cut-off date: 01 April 2022)	Assumed equal to SOC
SOC	0.00312	0.01196	VHL Natural History Study (2021)	VHL Natural History Study (2021)
VHL-CNS Hb				
Belzutifan	0.000156	0.00728	Assumed equal to pre-surgery → metastatic disease, as estimated from MK-6482-004 trial (data cut-off date: 01 April 2022)	Assumed equal to pre-surgery → death, as estimated from the VHL Natural History Study (2021)
SOC	0.00312	0.01456	Assumed equal to pre-surgery → metastatic disease, as estimated from the VHL Natural History Study (2021)	
VHL-pNET				
Belzutifan	0.00026	0.00624	Assumed equal to pre-surgery → metastatic disease, as estimated from MK-6482-004 trial (data cut-off date: 01 April 2022)]	Assumed equal to pre-surgery → death, as estimated from the VHL Natural History Study (2021)
SOC	0.00676	0.01092	Assumed equal to pre-surgery → metastatic disease, as estimated from the VHL Natural History Study (2021)	

CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Belzutifan arm

- *VHL-RCC Cohort:*

Transition probabilities from the non-metastatic health states to a subsequent health state for the VHL-RCC cohort for the belzutifan arm were estimated from time-to-event data from the MK-6482-004 trial and, for transitions where trial data was limited, the VHL Natural History Study.

Pre-surgery to surgery: In the belzutifan arm of the VHL-RCC cohort, the pre-surgery → surgery transition was estimated based on parametric fitting to observed time-to-event data from baseline visit to first post-baseline renal surgery in the overall MK-6482-004 trial population, all of whom had VHL-related RCC tumours at baseline (n=61). Given that only 7 renal surgeries were observed in the VHL-RCC cohort of the MK-6482-004 trial (as of the 01 April 2022 data cut-off date), an exponential distribution was applied to the observed KM data from the trial. The exponential distribution is used for this transition in the base case due to the following factors: **1**) the small number of pre-surgery → surgery events; **2**) the fact that the exponential requires fewer assumptions about underlying risk over time; and **3**) the suitability of the exponential distribution to model TTS based on data from the VHL Natural History Study based on visual inspection, statistical goodness-of-fit and clinical plausibility.

Pre-surgery to metastatic disease: As no cases of metastatic disease were observed in patients prior to surgery in the MK-6482-004 trial (as of the 01 April 2022 data cut-off date), the treatment effect on the risk of surgery was assumed to be equal to the treatment effect on the risk of metastases. Therefore, the transition from pre-surgery → metastatic disease in the belzutifan arm was estimated by applying the HR of pre-surgery → surgery (for belzutifan vs VHL Natural History Study) to the hazard of developing metastatic disease estimated for SOC. This assumption was considered clinically plausible given belzutifan is expected to reduce both the risks of surgeries and metastatic disease by decreasing the size of tumours and/or halting their growth (47).

Pre-surgery to death: Given the absence of evident VHL-tumour related deaths in MK-6482-004 and the low mortality rates observed in the VHL Natural History Study, the per-cycle TP from pre-surgery → death was set equal to the maximum of (i) the background mortality, using national mortality rates based on the age and sex distribution of the model cohort in each cycle, and (ii) the mortality rate of the VHL Natural History Study RCC cohort. The rate of pre-surgery to death is assumed to be equal between the belzutifan and SOC arms. Although belzutifan is expected to reduce the risk of death indirectly by preventing surgeries and incidence of metastases, data was not available (in the VHL population for whom surgery is unsuitable or undesirable) prior to surgery to model this transition separately by treatment arm. Therefore, the potential benefit of belzutifan reducing mortality risk due to a reduction in tumour size is not captured which likely underestimates the clinical benefit of belzutifan. For the pre-surgery → death transitions that are attributable to primary CNS Hb tumour progression and secondary CNS Hb tumour progression in the VHL-RCC and VHL-pNET cohorts, a treatment effect on overall survival is assumed (see *VHL-CNS Hb Cohort* below). It is important to note that two deaths due to other causes occurred in the trial population: one due to suicide and one due to toxicity from fentanyl and other agents. While these deaths may have been influenced by the patients' VHL disease, they were not considered be to be directly related to underlying VHL-tumour manifestations. Mortality rates for all cohorts were adjusted to fit the licensed population and is further described below in *Aligning VHL outcomes estimated with the population eligible for belzutifan per the MHRA license*.

Event-free after surgery transitions: As the pre-surgery → surgery transitions relate to 'last resort' primary tumour surgeries, transitions to subsequent surgery states were not permitted. In the absence of metastatic disease events in the MK-6482-004 trial, the transition from event-free after surgery → metastatic disease was estimated in a similar way to pre-surgery → metastatic disease. The TP from event-free after surgery state → metastatic disease in the belzutifan arm was estimated by calculating the ratio of the exponential hazard rates of pre-surgery → surgery for belzutifan versus the same risk in VHL Natural History Study and multiplying this

ratio by the hazard rate of event-free after surgery → metastatic disease in the SOC arm, which was also estimated directly from patient-level data in the VHL Natural History Study (further details of how the SOC arm TPs are calculated are described in the *SOC arm* section below).

As with the pre-surgery → death transition, belzutifan was assumed to affect direct transitions from the event-free after surgery state → death that are attributable to secondary CNS Hb tumours (as described in the *VHL-CNS Hb Cohort:* section). All other mortality causes were assumed equal between the belzutifan and SOC arms.

The exponential distribution was selected to estimate TPs from event-free after surgery given the exponential generates a time-constant hazard which is often most appropriate for transitions from an intermediate health state in a Markov model which are not time dependent.

- *VHL-CNS Hb Cohort:*

Transition probabilities from the non-metastatic health states for the VHL-CNS Hb cohort for the belzutifan arm were estimated from time-to-event data from the MK-6482-004 trial and, for transitions where trial data was limited, the VHL Natural History Study. For the VHL CNS Hb cohort, a mortality benefit of belzutifan has been accounted for.

Pre-surgery to surgery: In the VHL-CNS Hb population treated with belzutifan, the pre-surgery → surgery transition was estimated similarly to the VHL-RCC population, using the observed time from baseline visit to first post-baseline CNS surgery in the subset of the MK-6482-004 population with CNS Hb tumours (n=50). It should be noted that only 2 CNS Hb surgeries were observed in the VHL-CNS Hb cohort of the MK-6482-004 trial (both performed on the same patient) as of the 01 April 2022 data cut-off. An exponential distribution was fitted to the observed KM TTS data, as this distribution showed the best fit to the observed data based on visual inspection and statistical goodness-of-fit, provided clinically plausible long-term extrapolations and only required one assumption (constant hazards).

Pre-surgery to metastatic disease: CNS Hb tumours do not directly metastasise and any metastases in patients with CNS Hb tumours develop from non-primary tumour sites (i.e. RCC and pNET tumours). All metastases in the CNS Hb cohort therefore originated from non-primary tumours. The TP from pre-surgery → metastatic disease for the VHL-CNS Hb cohort (i.e. metastases for non-primary tumours) in the belzutifan arm was estimated by assuming the percentage reduction (belzutifan vs SOC) in the hazard rate of pre-surgery → metastatic disease to be equal to the percentage reduction in the hazard rate of pre-surgery → surgery (belzutifan vs SOC).

Pre-surgery to death: In the VHL-CNS Hb cohort, belzutifan is anticipated to provide an OS benefit compared to SOC. In the MK-6482-004 trial, the ORR was 44.0% among patients who had CNS Hb tumours at baseline visit (n=50), and the disease control rate was 90.0%. Patients with CNS-Hb in the MK-6482-004 trial experienced a 95% reduction in the weekly exponential hazard rate for the transition from pre-surgery to surgery when comparing the *pre-treatment* to *post-treatment* with belzutifan periods (transition of 0.00010 following belzutifan in the MK-6482-004 trial vs. 0.00202 during the pre-treatment period). As of the 01 April 2022 data cut-off, there have been no deaths due to CNS Hb progression in the trial (37). Although CNS Hb tumours are benign, CNS Hb tumours are a significant cause of severe neurological disability and mortality due to the pressure asserted on nearby CNS structures by the sheer mass of the tumours (7). In the MK-6482-004 trial, belzutifan demonstrated efficacy in reducing the size of CNS Hb and is therefore expected to reduce the risk of death due to CNS Hb progression. These data provide compelling evidence of the clinical benefits belzutifan offers to patients with CNS Hb (37).

To reflect the reduction in mortality associated with belzutifan for patients with CNS Hb supported by evidence from the MK-6482-004 trial, the risk of death from the pre-surgery health state attributable to the presence and growth of CNS Hb tumours was estimated by assuming the percentage reduction in mortality attributable to CNS Hb to be equal to the percentage reduction in the risk of pre-surgery → surgery with belzutifan vs. the VHL Natural History Study in the CNS Hb cohort. This assumption is supported by the observed reduction of CNS Hb surgeries and death due to CNS

Hb progression, which results from belzutifan's mode of action in decreasing the size or halting the growth of CNS Hb tumours.

The proportion of pre-surgery → death transitions that are attributable to CNS Hb progression in each model cohort are presented below in Table 47.

Accordingly, the following equation was used to calculate the cause-specific HR of pre-surgery → death for belzutifan in each model cohort:

(cause-specific hazard rate of pre-surgery → death under SOC) × [100% – (% of pre-surgery → death transitions attributable to CNS Hb progression in cohort)]

+

(cause-specific hazard rate of pre-surgery → death under SOC) × (% of pre-surgery → death transitions attributable to CNS Hb progression in cohort) × (% of pre-surgery → death transitions attributable to CNS Hb progression in cohort) × (hazard ratio of pre-surgery → surgery with belzutifan vs. SOC in the CNS Hb cohort)

In the above formula, the percentage of pre-surgery → death transitions attributable to CNS Hb progression was calculated for each cohort as follows: Among patients with CNS Hb, 50% of pre-surgery → death transitions were estimated to be attributable to CNS Hb progression, based on evidence from a prospective study of CNS Hb in VHL disease (see Table 47) (48). In the VHL-CNS Hb cohort, 100% of patients have CNS Hb (by definition). In the VHL-RCC and VHL-pNET cohorts, an estimated 82% and 89% of patients have CNS Hb, respectively, based on evidence from a cross-sectional survey study (Table 47). Thus, the percentage of pre-surgery → death transitions attributable to CNS Hb progression was calculated as 50% (= 50% × 100%) in the VHL-CNS Hb cohort, 41% (= 50% × 82%) in the VHL-RCC cohort, and 44.5% (= 50% × 89%) in the VHL-pNET cohort.

Table 47 Parameters used to estimate the proportion of pre-surgery → death transitions that are attributable to CNS Hb progression in each model cohort

Parameter	Value	Source
Among patients with VHL-CNS Hb, proportion of deaths that are attributable to CNS Hb progression	50%	Lonser et al. (2014)* (48)
Proportion of VHL-RCC cohort who have CNS Hb	82%	VHL RW QoL Disease Burden Study
Proportion of VHL-CNS Hb cohort who have CNS Hb	100%	By definition of the VHL-CNS Hb cohort
Proportion of VHL-pNET cohort who have CNS Hb	89%	VHL RW QoL Disease Burden Study

*In a prospective study of CNS Hb in VHL disease (N=225), Lonser et al. (2014) reported that 4 deaths were caused by CNS Hb progression (out of 8 deaths with known causes), implying that 50% of deaths among patients with VHL-associated CNS Hb are due to CNS Hb progression.
CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; VHL: Von Hippel Lindau

Event-free after surgery transitions: In the absence of metastatic disease events in the MK-6482-004 trial, the TP for event-free after surgery → metastatic disease was assumed to be equal to the TP for pre-surgery → metastatic disease. Consistent with the transition from pre-surgery → death, belzutifan was assumed to only directly affect the hazards of transitions from event-free after surgery → death attributable to CNS Hb progression; the hazards of transitions from event-free after surgery → death due to all other causes are assumed equal between belzutifan and SOC arms.

- *VHL-pNET Cohort:*

Transition probabilities for the VHL-pNET cohort for the belzutifan arm were estimated from time-to-event data from the MK-6482-004 trial and, for transitions where trial data was limited, the VHL Natural History Study. The estimation approach for metastases and mortality rates are largely similar to the VHL-RCC cohort.

Pre-surgery to surgery: In the VHL-pNET cohort of the MK-6482-004 trial (n=22), no pNET surgeries were observed as of the 01 April 2022 data cut-off. Rather than assume zero risk of pNET surgeries, the percentage reduction observed for belzutifan in the hazard of pre-surgery → surgery (vs. VHL Natural History Study) in the VHL-pNET cohort was assumed to be equal to the percentage reduction in the

hazard rate of pre-surgery → surgery with belzutifan (vs. VHL Natural History Study) in the VHL-RCC population, multiplied by (1-ORR_{pNET})/(1-ORR_{RCC}), where ORR_{pNET} and ORR_{RCC} are the ORRs of belzutifan with respect to the pNET and RCC tumours, respectively. This assumption was considered clinically plausible as it accounts for the higher ORR observed with belzutifan with respect to pNET tumours (91% compared with 64% for RCC tumours). As both the ORR and the need for surgery are determined by tumour size, the higher ORR with respect to pNET can be reflected in a greater percentage reduction in the hazard rate of pre-surgery → surgery.

Pre-surgery to metastatic disease: No metastases were observed in the MK-6482-004 trial in patients prior to surgery. Therefore, the transition from pre-surgery → metastatic disease in the belzutifan arm is estimated by applying the HR of pre-surgery → surgery (with belzutifan vs VHL Natural History Study) to the hazard rate of developing metastases estimated for SOC based on the VHL Natural History Study (see *VHL-RCC Cohort*: section above for further description). Pre-surgery to death: The TP of pre-surgery → death was calculated using the same approach as in the VHL-RCC cohort, assuming equal TPs between the belzutifan and SOC arms and accounting for mortality attributable to secondary CNS Hb tumours for which belzutifan is expected to provide clinical benefit (see *VHL-CNS Hb* Cohort: above).

Event-free after surgery transitions: In the absence of data, TPs from the event-free after surgery state were assumed equal to those from the pre-surgery state. This is a conservative assumption, as patients in the event-free after surgery state may have long-term complications as a result of surgery and therefore have a higher mortality risk as a result. This is not reflected in the current modelling approach to avoid undue complexity. The costs and utilities of surgical complications are also factored into this analysis (see *Surgical complications* section).

SOC arm

Due to the single arm design of the MK-6482-004 trial, RWE sources were used to estimate TPs for the SOC arm for all 3 VHL cohorts. For the VHL-CNS Hb and VHL-

pNET cohorts, the pre-treatment period of the MK-6482-004 trial was used to estimate TTS as RWE data was scarce for these cohorts.

As these patients have exhausted all options but still require therapy, immediate surgery is necessary, and results in loss of organ function and/or problematic sequalae. There is a lack of available evidence to estimate TTS associated with the current UK SOC in the population indicated per the MHRA label. Therefore, the pre-surgery → surgery transition is immediate following model entry in the SOC arm for 90% of the VHL-RCC and VHL-pNET cohorts, and for 100% of the VHL-CNS Hb cohort. TTS data from the VHL Natural History Study (VHL-RCC cohort) or the pre-treatment period of the MK-6482-004 trial (VHL-CNS Hb and VHL-pNET cohorts) are used for the remaining proportion who do not receive immiedate surgery in the SOC arm. These two data sources are the best available to proxy estimated TTS for the proportion who do not receive immediate surgery in the SOC arm. The estimation approach for this transition for these patients are described below.

- *VHL-RCC Cohort:*

Transition probabilities for the VHL-RCC cohort for the SOC arm were estimated from the VHL Natural History Study using a similar approach to the estimation of metastases and mortality rates to other primary tumour sites.

Pre-surgery to surgery *(for the 10% who do not receive immediate surgery)*: The cause-specific hazard of the pre-surgery to surgery transition was estimated by fitting a parametric distribution to the observed data on time from patient-level index-date to first surgery from the VHL Natural History Study cohort. An exponential distribution was selected for consistency with the belzutifan cohort and because this distribution showed the best fit to the observed data based on visual inspection, statistical goodness-of-fit and clinical plausibility. TTS data from the VHL Natural History Study were reweighted using propensity scores to match the baseline characteristics of patients in the VHL Natural History Study to those in the MK-6482-004 trial population (see B.2.9 Indirect and mixed treatment comparisons for more detail).

Pre-surgery to metastatic disease: The transition from pre-surgery → metastatic disease was estimated by fitting an exponential distribution to the observed rate of pre-surgery → metastasis in the VHL Natural History Study. An exponential distribution was selected given statistical goodness-of-fit, visual inspection and clinical plausibility.

Pre-surgery to death: The TP from pre-surgery → death was estimated using the same methods as for the belzutifan arm (please see *VHL-RCC Cohort:* above for the methods); however, it excludes the belzutifan-attributable reduction in the rate of death attributable to secondary CNS Hb progression.

Event-free after surgery transitions: Consistent with the approach in the belzutifan arm, transitions to subsequent surgeries were not permitted. In the VHL RCC cohort, the TP of event-free-after-first-surgery → metastatic disease for the SOC arm was estimated using patient-level time-to-event data from the VHL Natural History cohort, which was reweighted based on propensity score matching to the baseline characteristics of the MK-6482-004 trial population (see B.2.9 Indirect and mixed treatment comparisons for more information). The analytical sample included patients with ≥1 renal surgery at the patient-level index date (effective sample size = 75.7). Specifically, exponential models were fitted to observed data on time from the first post-index renal surgery date until each of the two transitions starting from event-free after surgery. The TP for event-free-after-surgery → death were estimated from patient-level mortality data from the VHL Natural History Study. The per cycle probability of death was set to the maximum of the estimated probability estimated using parametric models fitted to observed mortality data from the VHL Natural History Study and background mortality.

- *VHL-CNS Hb Cohort:*

Transition probabilities for the VHL-CNS Hb cohort for the SOC arm were estimated from the VHL Natural History Study using a similar approach to other tumour types. The pre-treatment period of the MK-6482-004 trial was used to estimate surgery rates following cessation of treatment effect in the belzutifan arm

(since 100% of the cohort receive the outcomes associated with immediate surgery).

Pre-surgery to surgery *(following treatment effect waning in the belzutifan arm)*: As detailed in the *Overview of data sources* section, the VHL Natural History Study could not be used to inform TTS in the VHL-CNS Hb cohort, as the study eligibility criteria only specified that patients were diagnosed with an RCC tumour; therefore a retrospective analysis of the pre-treatment period in MK-6482-004 was conducted (see *Retrospective analysis of the MK-6482-004 pre-treatment period* for further details). The pre-treatment period of the MK-6482-004 trial was therefore preferred as the best available source of data to inform TTS in CNS Hb patients. Specifically, the rate of pre-surgery to surgery for CNS Hb was fitted using time from the baseline visit (looking backwards) to the most recent pre-baseline CNS Hb surgery in the subset of patients in the MK-6482-004 trial with CNS Hb tumours at baseline. An exponential distribution was used to model this transition based on visual inspection, statistical goodness-of-fit and clinical plausibility.

Pre-surgery to metastatic disease: The transition from pre-surgery → metastatic disease uses the same approach as the *VHL-RCC Cohort*; however, it used a subset of the VHL Natural History Study cohort who had a pre-index history of CNS Hb. All metastases in the CNS Hb cohort arise from non-primary tumours (RCC and pNET tumours) as CNS Hb tumours do not themselves metastasize.

Pre-surgery to death: No patients in the MK-6482-004 trial died during the pre-treatment period (by definition, they were alive at study initiation), therefore the rate of pre-surgery to death in the SOC arm was estimated from the subset of patients in the VHL Natural History Study with a pre-index history of CNS Hb. The per-cycle TP from pre-surgery → death is set equal to the maximum of (i) the background mortality, using national mortality rates based on the age and gender distribution of the model cohort In each cycle, and (ii) the mortality rate of the VHL Natural History Study CNS Hb cohort. Few pre-surgery deaths were observed in the VHL Natural History Study, however, this population had disease which was less severe compared to the belzutifan-eligible population.

Event-free after surgery transitions: In the CNS Hb cohort of the VHL Natural History Study, the effective sample size of patients with ≥1 post-index primary tumour surgery was too small (n=3.8) to fit exponential models to transitions starting from the event-free after surgery state, as was done for the RCC cohort. Therefore, the TPs to the metastatic surgery state and the death state were assumed to be equal to the respective transitions from the pre-surgery state, as estimated from the VHL Natural History Study.

- *VHL-pNET Cohort:*

Transition probabilities for the VHL-pNET cohort for the SOC arm were estimated from the VHL Natural History Study using a similar approach to estimate metastases and mortality rates as the VHL-RCC cohort. The pre-treatment period of the MK-6482-004 trial was used to estimate surgery rates in the 10% who do not receive immediate surgery.

Pre-surgery to surgery *(for the 10% who do not receive immediate surgery)*:
Consistent with the approach for the CNS Hb cohort, the pre-treatment period of the MK-6482-004 was used to inform TTS in VHL-pNET patients. Specifically, the rate of pre-surgery to surgery for pNET was fitted using time from the baseline visit (looking backwards) to the most recent pre-baseline pNET surgery in the subset of patients in the MK-6482-004 trial with pNET tumours at baseline. An exponential distribution was used for this transition based on visual inspection, statistical goodness-of-fit and clinical plausibility.

Pre-surgery to metastatic disease: The transition from pre-surgery → metastatic disease uses the same approach as with the *VHL-RCC Cohort:*, however, it uses a subset of the VHL Natural History Study cohort who had a pre-index history of pNET.

Pre-surgery to death: This transition was estimated using the same methods as the *VHL-CNS Hb Cohort:* for the SOC arm above, however, it uses the subset of patients in the VHL Natural History Study with a pre-index history of pNET.

Event-free after surgery transitions: Consistent with the approach for the VHL-CNS Hb cohort, in the pNET cohort of the VHL Natural History Study, the effective sample size of patients with ≥1 post-index primary tumour surgery was too small (n=14.1) to fit exponential models to transitions starting from the event-free after surgery states, as was done for the RCC cohort. Therefore, TPs to the metastatic disease state and to the death state were assumed to be equal to the respective transitions from the pre-surgery state as estimated from the VHL Natural History Study.

Summary of transition probability estimation approaches from pre-surgery and event-free after surgery health states: Belzutifan and SOC

Table 48 Summary of transition probability estimation approaches from pre-surgery and event-free after surgery health states

Transition probability	VHL-RCC	VHL-CNS Hb	VHL-pNET
Pre-surgery → surgery	*Belzutifan:* TTS is estimated from MK-6482-004 trial applying exponential distribution	*Belzutifan:* TTS is estimated from MK-6482-004 trial applying exponential distribution	*Belzutifan:* % reduction in the hazard rate of this TP (for belzutifan vs. SOC) for VHL-pNET cohort is assumed equal to the % reduction in the hazard rate of this TP for VHL-RCC cohort multiplied by the complement ORR ratio
	SOC: 90% receive immediate surgery. For the remaining 10%, TTS is estimated from re-weighted VHL Natural History Study applying exponential distribution	*SOC:* 100% receive the outcomes associated with immediate surgery.	*SOC:* 90% receive immediate surgery. For the remaining 10%, TTS is estimated from the pre-treatment (looking backwards) applying exponential distribution
Pre-surgery → metastatic disease	*Belzutifan:* Hazard ratio of pre-surgery → surgery (for belzutifan vs. VHL Natural History Study) is applied to the hazard rate of pre-surgery → metastatic disease from the SOC arm	*Belzutifan:* % reduction in the hazard rate of pre-surgery → surgery (for belzutifan vs. VHL Natural History Study) is assumed equal to the % reduction in the hazard rate of pre-	*Belzutifan:* Hazard ratio of pre-surgery → surgery (for belzutifan vs. VHL Natural History Study) is applied to the hazard of pre-surgery → metastatic disease from the SOC arm

		surgery → metastatic disease which is applied to the SOC hazard rate*	
	SOC: TTM from the re-weighted VHL Natural History Study applying exponential distribution	*SOC:* TTM from the re-weighted VHL Natural History Study applying exponential distribution*	*SOC:* TTM from the re-weighted VHL Natural History Study applying exponential distribution
Pre-surgery → death	*Belzutifan:* Maximum of background mortality and VHL Natural History Study mortality and accounting for CNS Hb mortality benefit of belzutifan.	*Belzutifan:* % reduction in the hazard rate of pre-surgery → surgery (for belzutifan vs. VHL Natural History Study) is assumed equal to the % reduction in the hazard rate of pre-surgery → death which is applied to the SOC hazard rate	*Belzutifan:* Maximum of background mortality and VHL Natural History Study mortality and accounting for CNS Hb mortality benefit of belzutifan.
	SOC: Maximum of background mortality and VHL Natural History Study mortality	*SOC:* Maximum of background mortality and VHL Natural History Study mortality	*SOC:* Maximum of background mortality and VHL Natural History Study mortality
Event-free after surgery → event	*Belzutifan:* Hazard ratio of pre-surgery → metastatic disease (for belzutifan vs. VHL Natural History Study) is multiplied to the hazard rate of event-free after surgery → metastatic disease respectively in the SOC arm. For event-free after surgery → death: Maximum of background mortality and VHL Natural History Study mortality (for patients with ≥ 1 renal surgery) and accounting for CNS Hb mortality benefit of belzutifan.	*Belzutifan:* Event-free after surgery → event assumed equal to pre-surgery → event	*Belzutifan:* Event-free after surgery → event assumed equal to pre-surgery → event
	SOC: TTE from the re-weighted VHL Natural History for patients with ≥ 1 renal surgery applying exponential distribution. For event-free after surgery → death: Maximum of	*SOC:* Event-free after surgery → event assumed equal to pre-surgery → event	*SOC:* Event-free after surgery → event assumed equal to pre-surgery → event

	background mortality and VHL Natural History Study mortality (for patients with ≥ 1 renal surgery).		

*Assumes metastatic disease from non-primary tumour sites.

Surgery

Transitions from the surgery tunnel state

For both the belzutifan and SOC arms, patients in each tumour-site cohort who undergo surgery enter a tunnel state and where they remain for one cycle (1 week). All patients then transition to either event-free after surgery or death. Transitions for surgery → death therefore represent the risk of death as an immediate consequence of a primary tumour surgery (i.e., perioperative mortality: see the *Perioperative mortality* section for further details).

Incidences of non-primary VHL-related surgical procedures

Non-primary tumour surgeries were defined as any surgical procedure with therapeutic intent for VHL-associated manifestations other than for the primary tumour type in each tumour site cohort (i.e., non-RCC surgeries in the VHL-RCC cohort, non-CNS Hb surgeries in the VHL-CNS Hb cohort, and non-pNET surgeries in the VHL-pNET cohort).

For non-primary tumour surgeries in the VHL-RCC cohort in the belzutifan arm, the incidence rate was calculated as the number of observed non-primary tumour surgeries in the MK-6482-004 trial, divided by the follow-up time as of the 01 April 2022 data cut-off (49). For SOC, the same calculation was used for the incidence rate of non-primary surgeries in the VHL-RCC cohort using the VHL Natural History Study population subgroup. In the VHL-RCC cohort for both arms, the overall incidence rate of non-primary tumour surgeries was proportionally attributed to specific non-RCC VHL manifestations (i.e., CNS Hb, pNET, adrenal lesion, endolymphatic sac tumour, epididymal cystadenoma, or retinal Hb) based on the observed percentage breakdown of non-primary tumour surgery events in the reweighted VHL Natural History Study population subgroup.

For the VHL-CNS Hb and VHL-pNET cohorts, the incidence rates of non-primary tumour surgeries were derived from the respective subgroups of patients in the MK-6482-004 trial from the period after treatment initiation (in the belzutifan arm) or the period before treatment initiation (in the SOC arm). In the MK-6482-004 trial, all patients in the VHL-CNS Hb and VHL-pNET subgroups had RCC tumours at baseline. Because the corresponding tumour-site cohorts are defined to include patients with or without concurrent RCC tumours, the observed incidence rates of non-primary tumour surgeries for these cohorts were adjusted downward to account for the subset of patients without concurrent RCC. This was done based on data from the VHL RW QoL Disease Burden Study, which found that the proportion of patients without RCC was 36.8% (70/190) among patients with VHL-CNS Hb and 28.7% (35/122) among those with VHL-pNET. The incidence of RCC surgeries in the VHL-CNS Hb and VHL-pNET cohorts were thus reduced based on the assumption of zero RCC surgeries for the proportions of patients without baseline RCC. Specifically, the original rate of non-primary tumour surgeries in each arm (i.e., the as-observed rate during the post-treatment period of the MK-6482-004 trial for belzutifan and the pre-treatment period of the MK-6482-004 trial for SOC) was multiplied by the following correction factor:

*100% - (Observed % of non-primary tumour surgeries that were due to RCC during pre-treatment period of the MK-6482-004 trial) * (Observed % of patients without RCC in VHL RW QoL Disease Burden Study (44))*

In all three cohorts, following discontinuation of belzutifan, patients in the belzutifan arm were assumed to face the same incidence risk of non-primary tumour surgeries as the SOC arm after a period of waning of treatment effect (described in *Time to treatment discontinuation*). This approach is conservative given that the incidence rate of non-primary tumour surgeries for belzutifan was calculated using all available post-baseline follow-up in the MK-6482-004 trial (including person-time after discontinuation) and therefore already incorporates the impact of belzutifan discontinuation on surgery rates.

In each weekly cycle i of the model, the number of non-primary tumour surgeries occurring in that cycle was calculated by multiplying the weekly incidence of these

surgeries by the percentage of patients still alive in cycle i. As implied above, for the belzutifan arm, a different incidence rate was applied to the percentage of patients alive and on treatment versus alive and off treatment in each cycle. Risks of complications from non-primary tumour surgeries were accordingly calculated in each weekly cycle. Costs of non-primary tumour surgeries, as well as costs and QALY decrements due to non-primary tumour surgery complications, were then calculated in each cycle, and were layered (additively) onto the costs and QALYs estimated based on patients' distribution across primary tumour-related Markov health states.

Table 49 Incidence rates of surgeries (events/person-year) for non-primary VHL-related tumours, by cohort and treatment arm

Regimen	VHL-RCC cohort *Rate of non-RCC surgeries per year*	VHL-CNS Hb cohort *Rate of non-CNS Hb surgeries per year*	VHL-pNET cohort *Rate of non-pNET surgeries per year*	Sources & notes
Belzutifan	0.021187	0.029388	0.036498	*For RCC cohort:* MK-6482-004 trial (Data cut-off date: 1 April 2022) *For CNS Hb and pNET cohorts:* MK-6482-004 trial (Data cut-off date: 1 April 2022), with adjustment of RCC surgery incidence based on VHL RW QoL Disease Burden Study.
SOC *(also used in the belzutifan arm after treatment effect waning)*	0.178984	0.195923	0.340651	*For RCC cohort:* VHL Natural History Study (2021), reweighted to match MK-6482-004 population *For CNS Hb and pNET cohorts:* Analysis of pre-treatment period data from the MK-6482-004 trial, with adjustment of RCC surgery incidence based on VHL RW QoL Disease Burden Study

CNS: Central Nervous System; Hb: Haemangioblastoma; NHS: Natural History Study; pNET: pancreatic Neuroendocrine Tumour; SD: Standard Deviation; VHL: Von Hippel Lindau

Table 50 Distribution of surgeries for non-primary VHL-related tumours, by cohort

Surgery type	VHL-RCC cohort	VHL-CNS Hb cohort	VHL-pNET cohort	Sources & notes
RCC surgery		56.7%	38.2%	

CNS Hb surgery	52.4%		41.2%	*For RCC cohort:* VHL Natural History Study (2021), reweighted to match the MK-6482-004 population *For CNS Hb and pNET cohorts:* Analysis of pre-treatment period data from the MK-6482-004 trial, with adjustment of RCC surgery incidence based on VHL RW QoL Disease Burden Study (2022)
pNET surgery	3.4%	6.7%		
Adrenal lesion surgery	25.1%	3.3%	0.0%	
Endolymphatic sac tumor surgery	4.5%	0.0%	0.0%	
Epididymal cystadenoma surgery	0.1%	3.3%	0.0%	
Retinal Hb surgery	14.5%	29.9%	20.6%	
Total:	*100%*	*100%*	*100%*	

Surgical complications

Surgery places patients at risk of death as well as short- and long-term complications arising from surgery. Some of these long-term complications in the VHL-RCC and VHL-pNET cohorts are metabolic consequences due to a significant reduction or complete loss of organ function. The estimation approach and data sources for the risk of morbidity and mortality as a result of surgery is described below.

- *Perioperative mortality*

TPs for the surgery tunnel states → death represent the risk of death as an immediate consequence of surgical complications (i.e. perioperative mortality). Perioperative mortality risk in each cohort is assumed to be the same for both arms and is summarised in Table 51. For each cohort, the perioperative mortality risks were obtained from the following published literature sources:

- In the VHL-RCC cohort, the risk of death as an immediate consequence of surgery is sourced from Johnson et al. (2008) (50), a retrospective centre-based chart review of 51 repeat partial nephrectomies performed in patients with hereditary renal cancer, which reported a perioperative mortality rate of 1.96% (1/51). The perioperative mortality risk during subsequent surgeries is assumed to be equal to that in the first surgery.

- In the VHL-CNS Hb cohort, the risk of death as an immediate consequence of surgery is sourced from the Lonser et al. (2003) (51), a retrospective chart

review of 55 resections of spinal cord Hbs performed at the US NIH in patients with VHL disease, which reported a perioperative mortality rate of 1.82% (1/55).

- In the VHL-pNET cohort, the risk of death as an immediate consequence of first surgery is sourced from Krauss et al. (2018) (52), a multicentre international registry study of 273 patients with VHL-pNET, which reported a perioperative mortality of 1.7% (2 perioperative deaths out of 117 patients who underwent surgical procedures to remove pNETs).

Table 51 Perioperative mortality risk by VHL cohort

Cohort (equivalent for both model arms)	Surgery (tunnel state) → death		Sources & notes
	Risk	SE	
VHL-RCC	0.0196	(0.01941)	Johnson et al. (2008) (50)
VHL-CNS-Hb	0.0182	(0.01802)	Lonser et al. (2003) (51)
VHL-pNET	0.0171	(0.01198)	Krauss et al. (2018) (52)

CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumours; SE: standard error; RCC: renal cell carcinoma; VHL: Von Hippel Lindau

To align with the surgery-unsuitable or -undesirable population, the perioperative mortality risks were adjusted by a factor of 2.0 (i.e. doubled) for each cohort to reflect the increased risk of perioperative mortality as surgical procedures are a 'last resort' option in the MHRA label population in line with clinical expert opinion.

- *Risks of short-term and long-term surgical complications by type of surgery*

Rates of short- and long-term surgical complications associated with primary and non-primary VHL-related tumours were derived from data reported in the Optum study (34). Complications were identified based on diagnosis and/or procedure codes recorded within a specified time window following surgery for a given tumour type. The list of relevant surgical complications for each VHL-related tumour type, as well as the specific diagnosis and procedure codes used to identify each complication, were confirmed by input from clinical experts.

Short-term surgical complications

A 28-day window post-surgery in the available Optum study follow-up was used to identify short-term/acute complications. These were then applied to the economic

analysis by applying the risks of short-term complications from each type of surgery to the number of patients undergoing that surgery type within each weekly cycle. Risks of short-term surgical complications were assumed to be equivalent between belzutifan and SOC arms. Risk of short-term complications of surgery for each tumour type are presented in Table 52, Table 53, Table 54 and Table 55.

Given that belzutifan is licenced for patients for whom localised procedures are unsuitable or undesirable, the risk of short-term surgical complications observed in the Optum database study underestimates these risks in patients currently treated with SOC in the UK who would be eligible for belzutifan. These risks have therefore been adjusted upwards to better align with the expected risk of all surgical complications for the population eligible for belzutifan as per clinical expert input.

Table 52 Risks of short-term surgical complications per surgery for VHL-RCC

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
Acute renal failure	8.0%	16.0%
Cardiac complications	4.0%	8.0%
Erythroderma	0.8%	1.6%
Kidney infection	1.6%	3.2%
Other genitourinary complications	9.6%	19.2%
Postoperative infection (RCC-related)	6.4%	12.8%
Respiratory complications	20.8%	41.6%
Thrombosis and/or embolism	4.8%	9.6%
Vascular injury or anaemia	13.6%	27.2%

RCC: renal cell carcinoma; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 28-day period following the tumour reduction procedure
**Applies to primary tumour surgeries only

Table 53 Risks of short-term surgical complications per surgery for VHL-CNS Hb

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
Acute renal failure	7.7%	15.4%
CNS haemorrhage	12.8%	25.6%
Nerve palsy related to anaesthesia	5.1%	10.3%
Respiratory complications	20.5%	41.0%
Thrombosis and/or embolism	15.4%	30.8%
Vascular injury or anaemia	15.4%	30.8%

CNS: central nervous system; VHL: Von Hippel Lindau

*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 28-day period following the tumour reduction procedure
**Applies to primary tumour surgeries only

Table 54 Risks of short-term surgical complications per surgery for VHL-pNET

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
Abdominal abscess	10.0%	20.0%
Postoperative infection (pNET-related)	20.0%	40.0%
Respiratory complications	40.0%	80.0%
Thrombosis and/or embolism	10.0%	20.0%
Urinary tract infection	10.0%	20.0%
Vascular injury or anaemia	10.0%	20.0%

pNET: pancreatic neuroendocrine tumour; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 28-day period following the tumour reduction procedure
**Applies to primary tumour surgeries only

Table 55 Risks of short-term surgical complications per surgery for non-primary VHL-associated tumours

Complication	Risk of complication*
Complications of adrenal lesion surgery	
Acute renal failure	6.3%
Respiratory complications	31.3%
Thrombosis or embolism	12.5%
Vascular injury or anaemia	18.8%
Complications of endolymphatic sac tumour surgery	
Acoustic impairment	100.0%
Complications of retinal Hb surgery	
Vitreous haemorrhage	15.4%

Hb: haemangioblastoma; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 28-day period following the tumour reduction procedure

Long-term surgical complications

A 180-day post-surgery follow-up window in the Optum study was used to identify long-term complications including metabolic consequences (e.g., end-stage renal disease (ESRD) and/or dialysis following nephrectomy). Patients who develop long-term complications from any given type of surgery will have these complications for the remainder of their lifetime. This was estimated by tracking the cumulative proportion of patients who have developed each long-term complication and who are still alive in each weekly cycle. It is important to note that although long-term

complications are assumed to continue for the remainder of a patient's life, the increased mortality risk due to these complications, particularly metabolic consequences such as ESRD, chronic kidney disease (CKD) and pancreatogenic diabetes, is not modelled due to the limits of Markov cohort approach. Risks of long-term surgical complications were assumed to be equivalent between belzutifan and SOC arms (as with short-term complications). Risks of long-term complications following surgery are reported in Table 56, Table 57, Table 58 and Table 59.

Consistent with the approach for short-term complications, the risk of long-term surgical complications was aligned with the population eligible for belzutifan. The risks of long-term metabolic consequences resulting from surgery were adjusted further to reflect the limited/absent organ function following surgery in the licensed population, consistent with clinical expert feedback. These included risks of ESRD and CKD in the VHL-RCC cohort and secondary diabetes and immunocompromisation in the VHL-pNET cohort. Risk of cerebral vascular occlusion/stroke was also adjusted for the VHL-CNS Hb cohort to reflect the heightened risk associated with surgery in the licensed population.

Table 56 Risks of long-term complications following surgery for VHL-RCC

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
End stage renal disease and/or dialysis***	4.0%	80.0%
Chronic kidney disease***	24.0%	20.0%
Hernia surgery	1.6%	3.2%
Chronic pain	8.8%	17.6%
Cerebral vasculature occlusion or stroke	3.2%	6.4%

RCC: renal cell carcinoma; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 180-day period following the tumour reduction procedure
**Applies to primary tumour surgeries only.
***This is a metabolic complication resulting from limited/absent organ function following surgery. Note: all patients will have some form of renal impairment following VHL-RCC surgery, majority will have ESRD and the remainder will have CKD as per clinical expert opinion.

Table 57 Risks of long-term surgical complications per surgery for VHL-CNS Hb

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
Chronic pain (in CNS Hb population)	15.4%	30.8%
Cerebral vasculature occlusion or stroke	7.7%	85.0%
Seizure	10.3%	20.5%
Neurological complications	43.6%	87.2%

CNS Hb: central nervous system haemangioblastoma; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 180-day period following the tumour reduction procedure
**Applies to primary tumour surgeries only.

Table 58 Risks of long-term surgical complications per surgery for VHL-pNET

Complication	Risk of complication*	Risk of complication adjusted for MHRA label population**
Chronic pain (in pNET population)	10.0%	20.0%
Secondary diabetes or exocrine pancreatic insufficiency***	20.0%	100.0%
Immunocompromisation***	0.0%	100.0%

pNET: pancreatic neuroendocrine tumour; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 180-day period following the tumour reduction procedure.
**Applies to primary tumour surgeries only.
***This is a metabolic complication resulting from limited/absent organ function following surgery. For the VHL-pNET cohort this reflects loss of pancreatic function.

Table 59 Risks of long-term surgical complications per surgery for non-primary VHL-associated tumours

Complication	Risk of complication*
Complications of adrenal lesion surgery	
Adrenal insufficiency	31.3%
Chronic pain	25.0%
Complications of retinal Hb surgery	
Chronic pain	3.8%

Hb: haemangioblastoma; VHL: Von Hippel Lindau
*Risks of complications were measured in the Optum Clinformatics Data Mart claims study (2000–- 2020) during the 180-day period following the tumour reduction procedure

Aligning risk of surgery and metastatic disease to real-world SOC

In UK clinical practice, clinical geneticists lead on VHL services and co-ordination of care can be complex (3). This can result in delays in surgery which, in turn, lead to higher rates of metastatic disease. A limitation of the VHL Natural History Study is that patients in the study received an elevated SOC compared with UK clinical

practice, according to a UK clinical expert. The VHL Natural History Study focussed on patients treated at the US NCI, which is a Centre of Excellence. In this setting, rates of surgeries were considered to be higher (and hence rates of metastatic disease lower) than would be expected at other centres without this designation where most patients in the UK receive care (see *Optum Clinformatics Data Mart claims study* section above for further details). In addition, the frequency of routine imaging procedures and surgical interventions in the MK-6482-004 trial also reflects an elevated SOC. To reflect the expected level of care in clinical practice in England, TPs were adjusted in both arms using real-world data from the Optum Clinformatics Data Mart (Optum database) (41), which is more reflective of real-world clinical practice. The assumption that surgery rates observed in the VHL Natural History Study are elevated and result in lower rates of metastases compared to usual care in clinical practice was validated by clinical experts, as was the use of the real-world data from the Optum database to reflect SOC in the UK.

The Optum database is used to adjust the TPs for (i) pre-surgery → surgery (with the exception of those receiving immediate surgery in the SOC arm), (ii) pre-surgery → metastatic disease, and (iii) event-free after surgery → metastatic disease. To align with outcomes expected in the real world in the SOC arm, the difference in the cause-specific hazards of event-free after surgery → next surgery between the Optum database and the VHL Natural History Study population is added to the cause-specific hazards of pre-surgery → surgery (with the exception of those receiving immediate surgery in the SOC arm). Similarly, the difference in the observed cause-specific hazards of event-free after surgery → metastatic disease between the Optum database and the VHL Natural History Study population is added to the cause-specific hazard rates of pre-surgery → metastatic disease and event-free after surgery → metastatic disease in the SOC arm of the model. Adjustments were based on differences in the event-free after surgery → surgery/metastatic disease transition rather than pre-surgery → surgery/metastatic disease transition due to different index dates in the VHL Natural History Study and the Optum Clinformatics Data Mart study. In the belzutifan arm, the cause-specific hazard rates of these transitions are accordingly adjusted by applying the ratios of

the exponential rates of these transitions with belzutifan versus SOC (as estimated based on the VHL Natural History Study or pre-treatment period of the MK-6482-004 trial). Thus, the relative treatment effects of belzutifan (versus SOC) on these transitions are not affected when the adjustment is applied. In the VHL-CNS Hb and VHL-pNET cohorts, no adjustments based on real-world SOC are applied to the cause-specific hazard rates of pre-surgery → surgery, as these transitions were modelled based on data from the pre-treatment period of MK-6482-004 for those who did not receive immediate surgery. Data on risk of surgery during the pre-treatment period were collected retrospectively, and patients therefore may not have received an elevated SOC during this time period. The magnitude of differences in hazard rates for the event-free after surgery to next surgery and event-free after surgery to metastatic disease transitions between the Optum database and the VHL Natural History Study is presented below in Table 60.

Table 60 Adjustment factors applied to the hazard rates of surgery and metastatic disease between the Optum database and VHL Natural History Study (to account for real-world SOC)

Difference in hazard rates	VHL-RCC cohort	VHL-CNS Hb cohort	VHL-pNET cohort
Event-free after surgery to next surgery	-0.00109	N/A	N/A
Event-free after surgery to metastatic disease	0.00115	0.00107	0.00255

CNS Hb: central nervous system haemangioblastoma; N/A: not applicable; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Transitions from metastatic disease to death

The TPs from metastatic disease to death are assumed to be equal between the belzutifan and SOC arms. The transition from metastatic disease to death is dependent on the origin tumour of the metastases and, linked to the origin tumour, the associated survival with each metastatic disease treatment.

In each cohort, metastases were assumed to originate from either RCC or pNET, as data from the VHL Natural History Study show that RCC or pNET were the origin tumours for the vast majority of patients who developed metastatic disease. The rate of metastases by origin tumour by cohort is shown in Table 61.

Table 61 Metastases by origin tumour

Indication	% of metastases by origin tumour:	
	RCC	pNET
VHL-RCC	97%	3%
VHL-CNS Hb*	78%	22%
VHL-pNET	66%	34%

*For patients in the CNS Hb cohort, all cases of metastatic disease are due to metastases of non-primary RCC or pNET, as CNS Hb does not metastasise.
CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; VHL: Von Hippel Lindau

VHL-RCC and VHL-pNET cohorts

The origin tumour for the metastasis determines the possible metastatic treatment options. Estimated OS (starting from metastatic disease) in each treatment arm is calculated as a weighted average of estimated OS associated with the first-line metastatic disease treatments. Estimated OS is then used to derive a per-cycle HR. This approach was previously used in the NICE appraisal of pembrolizumab for adjuvant treatment of RCC (TA830) and was accepted by NICE (53).

The market shares of first-line treatments were assumed equal between the belzutifan and SOC arms, and were estimated based on the subsequent treatment market shares used in the NICE appraisal of pembrolizumab as an adjuvant treatment of RCC post-nephrectomy (TA830) in the VHL-RCC cohort (45), and European Society for Medical Oncology (ESMO) clinical practice guidelines for gastroenteropancreatic neuroendocrine neoplasms and input from clinical experts in the VHL-pNET cohort (54). Market share data for first-line regimens for metastatic disease in RCC and pNET are presented in Table 62 and Note: This market share data is as of November 2021.

Table 63, respectively.

Table 62 Market shares of first-line regimens for metastatic disease by treatment arm with RCC as the origin tumour

First-line regimens in advanced setting	First-line market shares by treatment arm (%)	
	Belzutifan	SOC
Sunitinib	30.0%	30.0%
Tivozanib	14.0%	14.0%
Pazopanib	29.0%	29.0%
Cabozantinib	13.0%	13.0%

Nivolumab / ipilimumab	14.0%	14.0%
Avelumab / axitinib	0.0%	0.0%
Pembrolizumab / lenvatinib	0.0%	0.0%

Note: This market share data is as of November 2021.

Table 63 Market shares of first-line regimens for metastatic disease by treatment arm with pNET as the origin tumour

First-line regimens in advanced setting	First-line market shares by treatment arm (%)	
	Belzutifan	SOC
Streptozocin / 5-fluorouracil	0.0%	0.0%
Streptozocin / doxorubicin	0.0%	0.0%
Temozolomide / capecitabine	0.0%	0.0%
Everolimus	0.0%	0.0%
Sunitinib	0.0%	0.0%
Interferon a2B	0.0%	0.0%
Lanreotide	50.0%	50.0%
Octreotide	50.0%	50.0%
No active treatment	0.0%	0.0%

- *Effectiveness of first-line treatments for metastatic RCC:*

For each advanced RCC treatment regimen, exponential models of OS and PFS were estimated based on the approach described below:

- For sunitinib in the advanced RCC setting, exponential rates of OS and PFS failure were computed based on the observed median OS and PFS in the sunitinib arm of KEYNOTE-426, a phase III, randomised, open-label, multicenter, global trial to evaluate the efficacy and safety of pembrolizumab in combination with axitinib versus sunitinib monotherapy as a first-line treatment for advanced RCC (55). The resulting exponential curves are plotted in Figure 19 alongside digitised KM curves for sunitinib in the KEYNOTE-426 trial to illustrate visual fit. The use of exponential distributions to model OS and PFS for sunitinib in the advanced RCC setting is also consistent with a previously published cost-effectiveness analysis based on the KEYNOTE-426 trial (56).

- For other advanced treatment regimens, HRs for OS and PFS versus sunitinib were each obtained from a NMA of trials conducted in advanced RCC. Trials included in the NMA were identified through a SLR of RCTs of first-line treatments in patients with locally advanced unresectable or metastatic RCC with clear-cell histology (57). For pembrolizumab/lenvatinib, the HRs for OS and

PFS versus sunitinib were obtained from a clinical trial publication by Motzer et al. (2021) in the advanced RCC setting (58), because pembrolizumab/lenvatinib was not included in this NMA. For each comparator, the model applied time-constant HRs estimated through fixed-effects NMAs of OS and PFS.

Table 64 presents the exponential rates of OS and PFS failure estimated for sunitinib and the no treatment option in the advanced setting. Table 65 summarises the HRs of OS and PFS failure with other treatment regimens versus sunitinib obtained from the NMA and resulting estimates of mean OS and PFS (in weeks) for each regimen.

Table 64 Exponential models of OS and PFS with sunitinib in advanced RCC

Advanced regimen	Exponential model of OS		Exponential model of PFS		Source
	Rate	SE	Rate	SE	
Sunitinib	0.0040	(0.0003)	0.0144	(0.0013)	KEYNOTE-426 (Rini et al. 2021)*

* For sunitinib in the advanced RCC setting, exponential models of OS and PFS were calculated based on median PFS and OS reported from the KEYNOTE-426 trial (55).
OS: overall survival; PFS: progression-free survival; SE: standard error

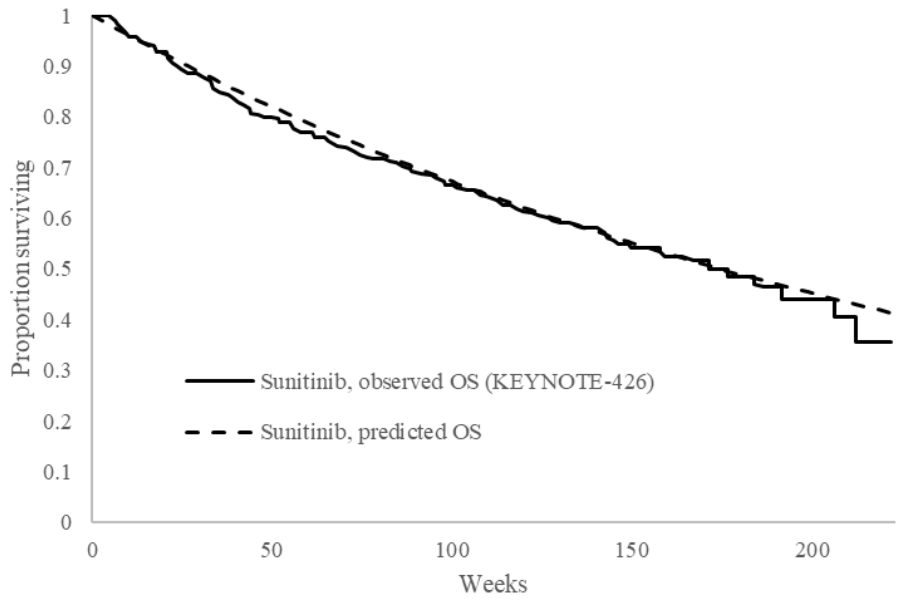
Table 65 HRs of OS and PFS failure with other treatments vs. sunitinib in advanced RCC

Advanced regimen	HR of death vs. sunitinib		HR of progression or death vs. sunitinib		Expected survival in metastatic state (weeks)	
	HR	SE of ln(HR)	HR	SE of ln(HR)	OS	PFS
Sunitinib	1.00		1.00		252	70
Tivozanib	1.33	0.27	1.19	0.26	189	59
Pazopanib	0.92	0.08	1.05	0.08	273	66
Cabozantinib	0.80	0.21	0.48	0.22	314	145
Nivolumab/ipilimumab	0.72	0.08	0.89	0.08	349	78
Avelumab/axitinib	0.80	0.13	0.69	0.09	314	101
Pembrolizumab/lenvatinib	0.66	0.15	0.39	0.11	381	179

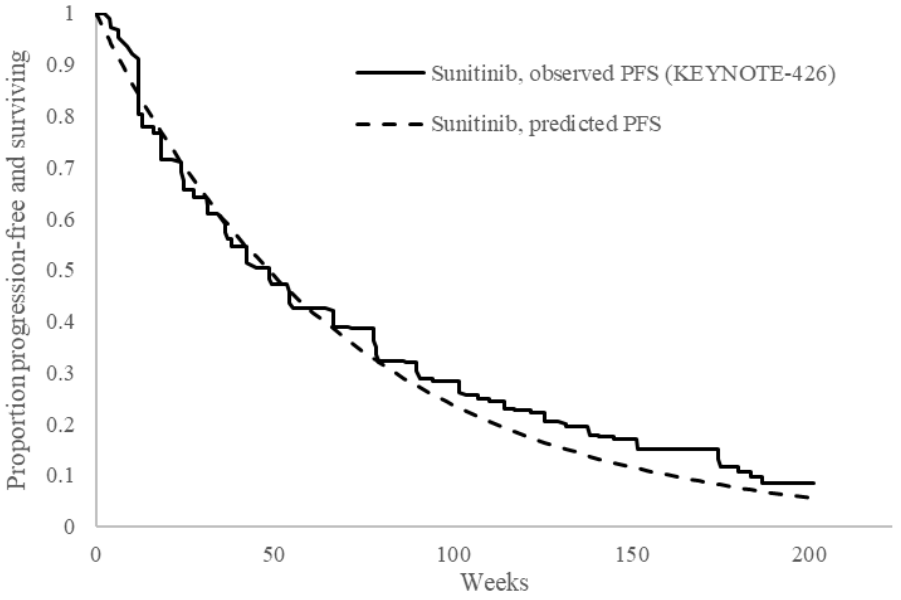
HR: hazard ratio; ln: natural logarithm; NMA: network meta-analysis; OS: overall survival; PFS: progression-free survival; SE: standard error.

Figure 19 Exponential models of OS and PFS compared with Kaplan-Meier curve extractions for sunitinib in the advanced RCC setting

OS



PFS



OS: overall survival; PFS: progression-free survival; RCC: renal cell carcinoma

- *Effectiveness of first-line treatments for metastatic pNET*

The approach used to obtain exponential models of OS and PFS for treatments of metastatic pNET was similar to that used to estimate the effectiveness of first-line treatments for metastatic RCC. The method was as follows:

- For streptozocin/5-fluorouracil and no active treatment in the advanced pNET setting, exponential rates of OS and PFS failure were estimated based on OS and PFS data extracted from Study E1281 (59) and NCT00428597 trial publications (60), respectively (see Table 66).

- Based on data availability, OS and PFS for streptozocin/doxorubicin and temozolide/capecitabine were assumed equal to that of streptozocin/5-fluorouracil, as these three combination regimens are indicated for higher-grade pNET.

- For other advanced treatment regimens, HRs for OS and PFS versus no active treatment were each obtained from an SLR and NMA of trials conducted in advanced pNET (61) (see Table 67). The HR of OS for each treatment versus no active treatment was assumed equal to the NMA-based HR of PFS for that treatment versus no active treatment.

Table 66 Exponential models of OS and PFS with streptozocin/5-fluorouracil and no active treatment in the advanced pNET setting

Advanced regimen	Exponential model of OS		Exponential model of PFS		Source
	Rate	SE	Rate	SE	
Streptozocin / 5-fluorouracil	0.0066	(0.0016)	0.0301	(0.0055)	Sun et al. (2005) [Study E1281]*
No active treatment	0.0055	(0.0014)	0.0275	(0.0051)	Faivre et al. (2017) [NCT00428597]**

*For streptozocin/5-fluorouracil in the advanced pNET setting, exponential models of OS and PFS were calculated based on median OS and PFS reported from Study E1281 (Sun et al. 2005) (59). Of note, the rates of OS and PFS failure are higher for streptozocin/5-fluorouracil than for no active treatment, as streptozocin/5-fluorouracil is indicated for higher-grade pNET.
** For patients who receive no active treatment in the advanced pNET setting, exponential models of OS and PFS were calculated based on median OS and PFS reported from NCT00428597 (Faivre et al. 2017) (60).
OS: overall survival; PFS: progression-free survival; pNET: pancreatic neuroendocrine tumour; SE: standard error

Table 67 HRs of OS and PFS failure with other treatment regimens vs. no active treatment in the advanced pNET setting

Advanced regimen	HR of death vs. no active treatment	HR of progression or death vs. no active treatment	Expected survival in metastatic state (weeks)

	HR	SE of ln(HR)	HR	SE of ln(HR)	OS	PFS
Streptozocin / 5-fluorouracil	1.20	-	1.09	-	152	33
Streptozocin / doxorubicin	1.20	-	1.09	-	152	33
Temozolomide / capecitabine	1.20	-	1.09	-	152	33
Everolimus	0.35	0.12	0.35	0.12	522	104
Sunitinib	0.42	0.24	0.42	0.24	435	87
Interferon a2B	0.37	0.42	0.37	0.42	493	98
Lanreotide	0.46	0.18	0.46	0.18	397	79
Octreotide	0.46	0.18	0.46	0.18	397	79
No active treatment	1.00	-	1.00	-	183	36

Based on data availability, OS and PFS for streptozocin/doxorubicin and temozolide/capecitabine were assumed equal to that of streptozocin/5-fluorouracil, as these three combination regimens are indicated for higher-grade pNET. Due to their indication for higher-grade pNET, mean OS and PFS for these three treatment regimens are lower than for no active treatment. For other advanced treatment regimens, HRs for OS and PFS vs. no active treatment were each obtained from a SLR and NMA of trials conducted in advanced pNET (Kaderli et al. 2019 (61)). The HR of OS for each treatment versus no active treatment was assumed equal to the NMA-based HR of PFS for that treatment versus no active treatment.
HR: hazard ratio; ln: natural logarithm; NMA: network meta-analysis; OS: overall survival; PFS: progression-free survival; SE: standard error.

VHL-CNS Hb cohort

In the CNS Hb cohort, the origin of all metastases is assumed to be secondary VHL-RCC or VHL-pNET tumours because the VHL Natural History Study showed that metastases did not arise from VHL-CNS Hb tumours, as described above. As such, the transition rate of metastasis to death was calculated similarly to the VHL-RCC and VHL-pNET cohorts, using the distribution of origin tumours in the VHL-CNS Hb population, the market shares of first-line treatments for VHL-RCC and VHL-pNET, and the efficacy of the first-line treatments for VHL-RCC and VHL-pNET.

Estimation of the HR of death from metastatic disease by treatment arm

Based on the above parameters, mean OS (in weeks) within the metastatic disease state was calculated in each target population as a weighted average of estimated OS associated with different first-line treatments for advanced RCC and pNET, based on the origin tumour distribution and market shares of first-line advanced treatments (see Table 68). Mean OS was then translated into a weekly exponential HR.

Table 68 Hazards of death from metastatic disease by target population and treatment arm

Indication / model arm	Expected survival in metastatic state (weeks): Weighted average based on origin tumour and first-line advanced treatment market shares			Metastatic disease to death: Exponential hazard rate based on estimated OS
	OS	PFS	Ratio of PFS:OS	
RCC population				
Belzutifan	275	78	0.28	0.0036
SOC	275	78	0.28	0.0036
CNS Hb population				
Belzutifan	299	78	0.26	0.0033
SOC	299	78	0.26	0.0033
pNET population				
Belzutifan	314	78	0.25	0.0032
SOC	314	78	0.25	0.0032

CNS: central nervous system; Hb: haemangioblastoma; OS: overall survival; PFS: progression free survival; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; SOC: standard of care

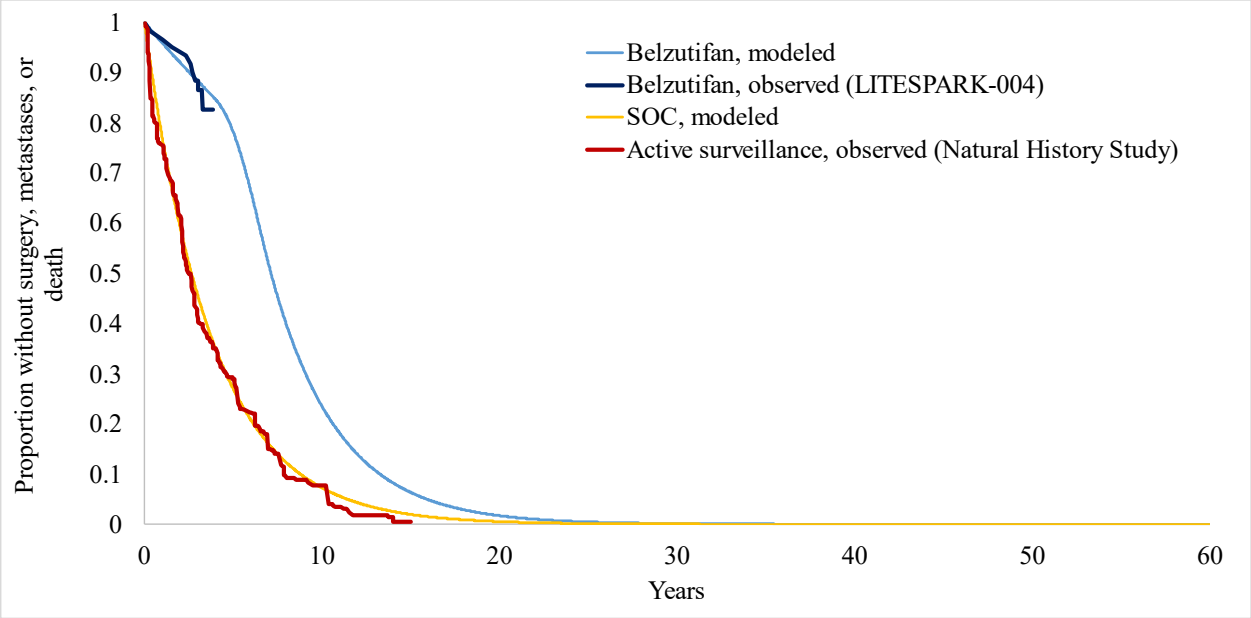
Validation of transition probabilities

Validation of the modelled transition from pre-surgery to surgery, metastases, or death

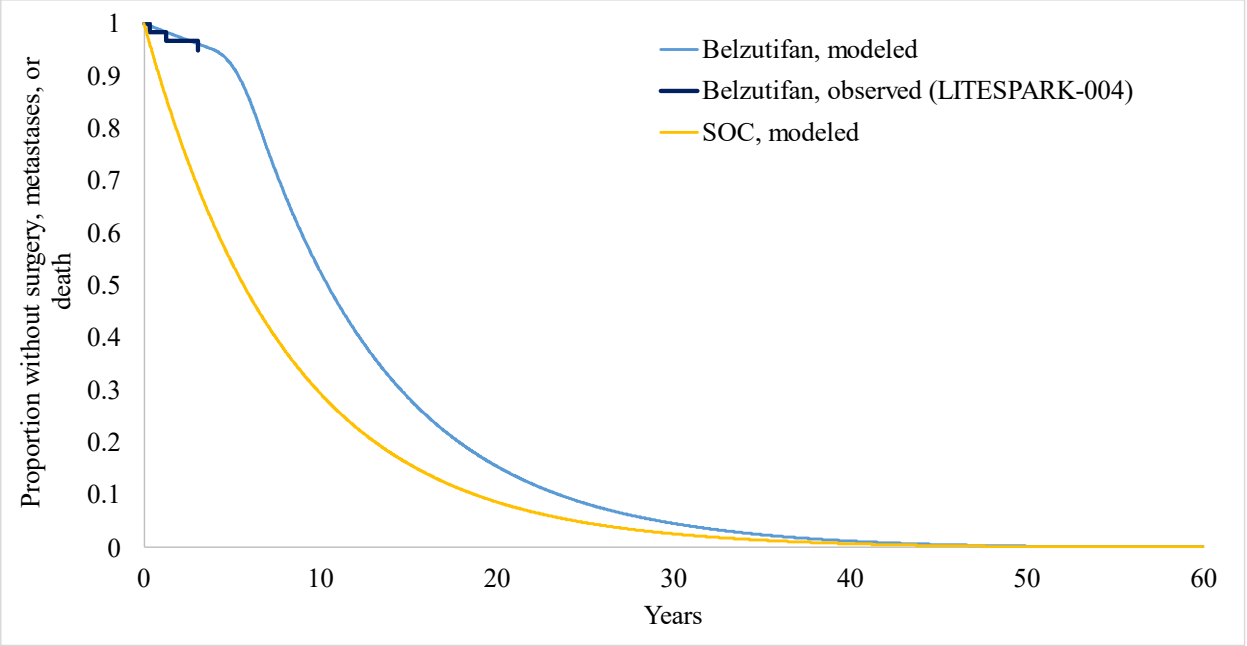
There is no available data for the target population stipulated by the MHRA label to allow for validation of the efficacy inputs modelled (see section B.2.1 Identification and selection of relevant studies and B.3.1 Published cost-effectiveness studies), therefore validation of the exact modelled outcomes could not be assessed. In order to include validation of efficacy outcomes against original sources, adjustments to account for real-world SOC and the assumption of immediate surgery for 90% (VHL-RCC and VHL-pNET cohorts) or 100% (VHL-CNS Hb) cohorts were removed to obtain interpretable comparisons. These are presented in Figure 20 and Table 69 provides a summary of the long-term extrapolation of time to first surgery, metastases, or death, for belzutifan and SOC in each VHL cohort.

Figure 20 Internal validations of time to surgery, metastases, or death against original data sources in each model arm and cohort

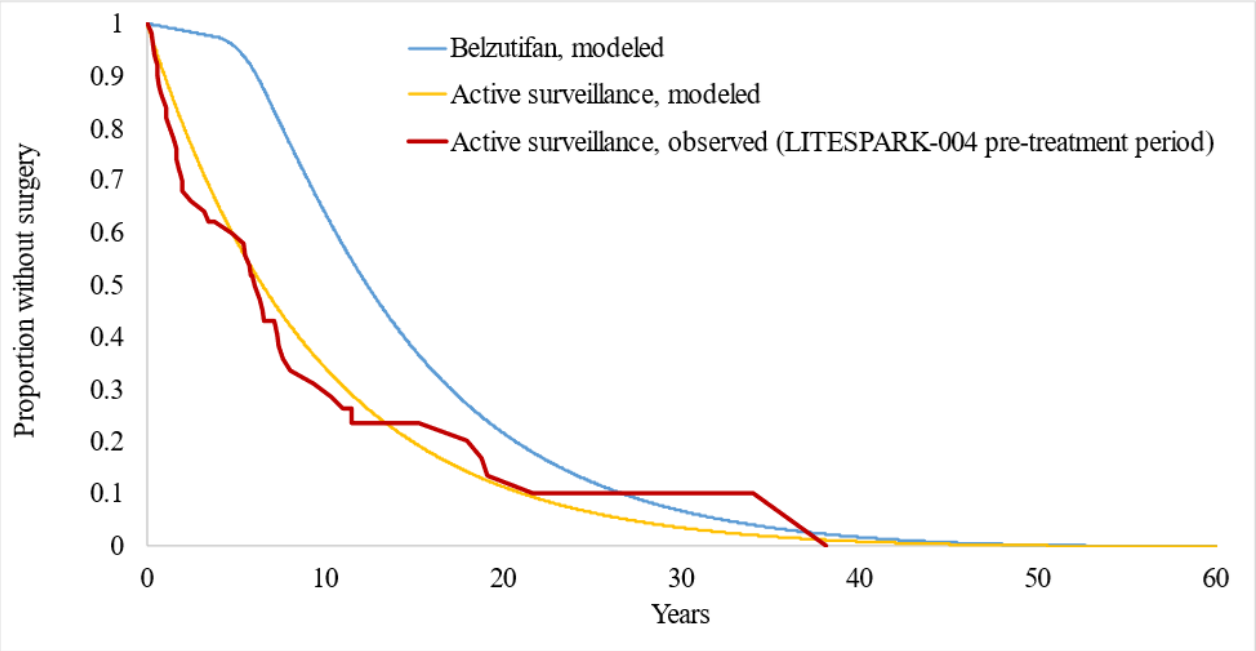
(a) VHL-RCC cohort – *Validation of both arms*



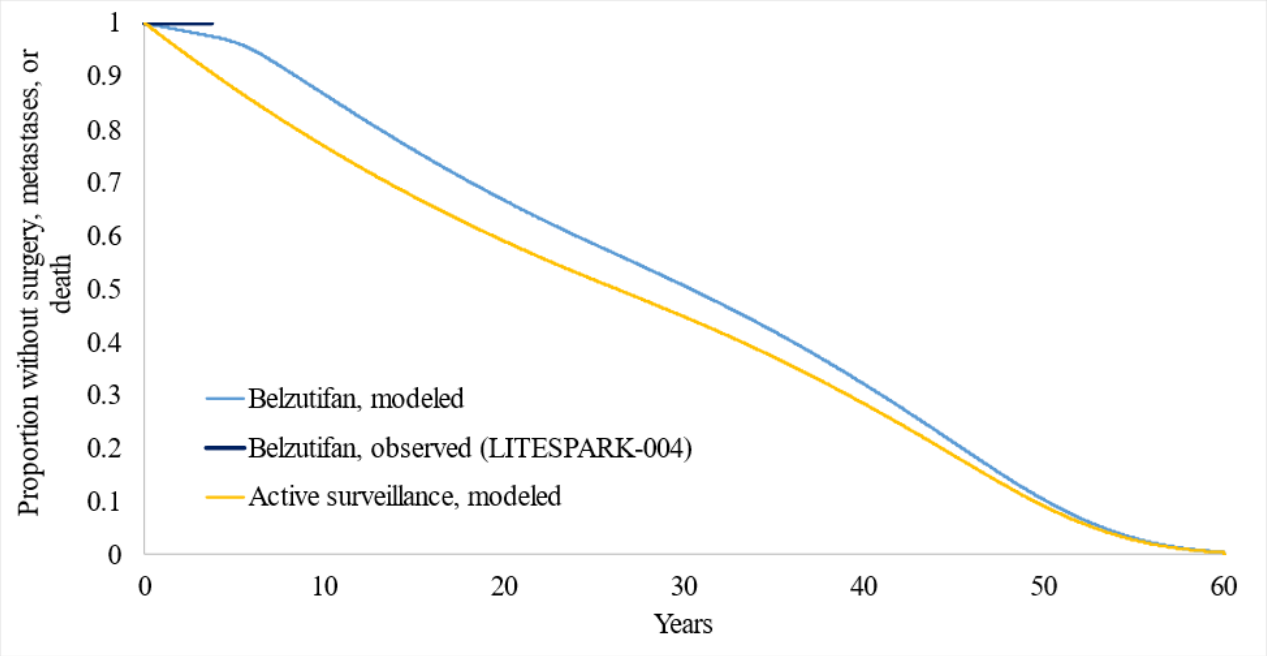
(b) VHL-CNS Hb cohort – *Validation of belzutifan arm*



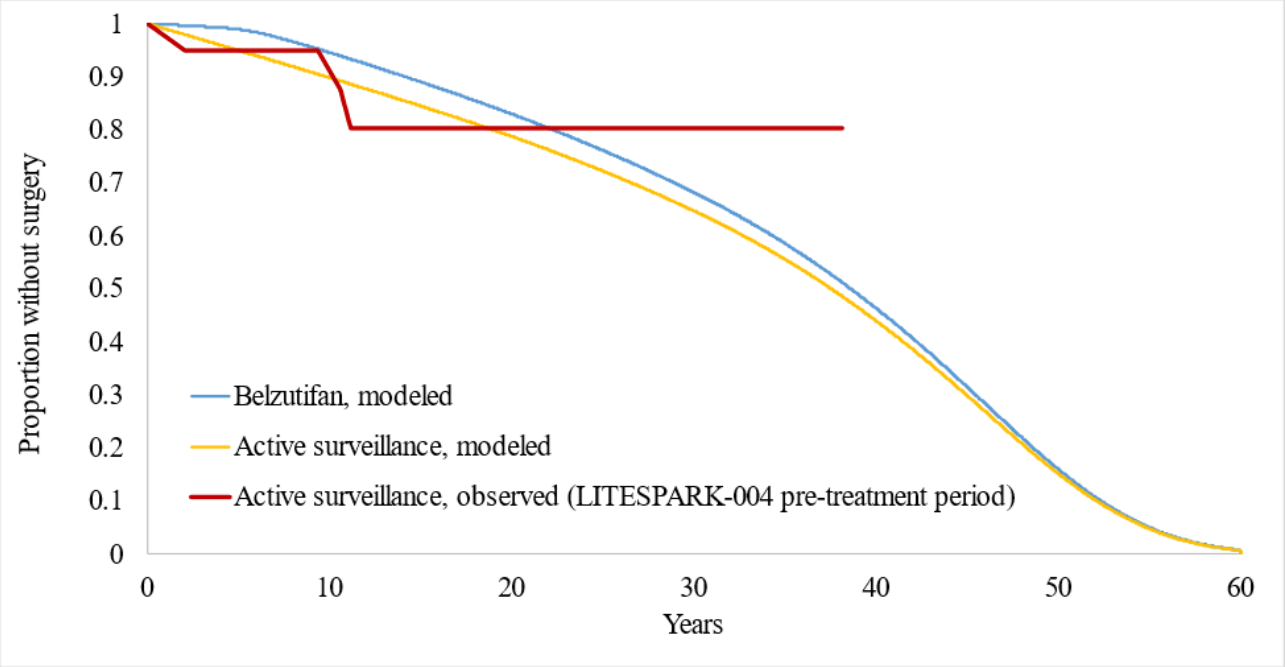
- **VHL-CNS Hb cohort –** *Validation of SOC arm†*



- **VHL-pNET cohort –** *Validation of belzutifan arm*



- **VHL-pNET cohort –** *Validation of SOC arm†*



Notes:
To obtain interpretable comparisons between modelled vs. observed curves, the modelled curves shown in figures (a) –l) above were generated without performing any adjustments to align with the GB label population or account for real-world standard of care.
†In figul (land (e), the observed KM curve for SOC is from the MK-6482-004 pre-treatment period, in which there were no metastases or deaths by definition (i.e., because patients had to be alive and metastases-free in order to enrol in the MK-6482-004 trial). Thus, to obtain interpretable comparisons between the modelled curve for SOC vs. the observed curve from the MK-6482-004 pre-treatment period, the cause-specific hazards of pre-surgery → metastases and pre-surgery → death were temporarily set to 0 when generating the modelled curves for SOC and belzutifan iligIs (c) and (e). The modelled curves in both arms are therefore hIer in figure (c) than figure (b), anligher in figure (e) than figure (d).
LITESPARK-004 refers to MK-6482-004 trial.
CNS Hb: central nervous system haemangioblastoma; KM: Kaplan-Meier; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; SOC: standard of care; TP: transition probability; VHL: Von Hippel Lindau

Table 69 Landmark estimates: time to first surgery, metastases, or death

Time (years)	VHL-RCC		VHL-CNS Hb		VHL-pNET	
	Belzutifan	SOC	Belzutifan	SOC	Belzutifan	SOC
0.5	98.0%	89.0%	99.0%	93.2%	98.9%	85.9%
1	96.1%	79.1%	98.0%	87.0%	97.7%	73.7%
1.5	94.2%	70.4%	97.0%	81.1%	96.6%	63.3%
1.75	93.3%	66.4%	96.5%	78.3%	96.1%	58.7%
2	92.3%	62.6%	96.0%	75.6%	95.5%	54.4%
5	79.8%	31.0%	88.8%	49.7%	85.9%	21.8%
10	33.0%	9.6%	52.0%	24.6%	29.5%	4.7%
20	3.4%	0.9%	13.1%	6.1%	1.6%	0.2%
30	0.3%	0.1%	3.2%	1.5%	0.1%	0.0%
40	0.0%	0.0%	0.8%	0.3%	0.0%	0.0%
45	0.0%	0.0%	0.3%	0.1%	0.0%	0.0%
50	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%

55	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
60	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau
Note: These landmark estimates reflect modelled inputs without the assumption of immediate surgery to reflect the target population stipulated by the MHRA label (therefore reflecting the proportion who do not get immediate surgery) or adjustments to account for real-world SOC. This was done to obtain an interpretable comparison due to the lack of an available data source reflecting the target population of interest.

Validation of the long-term extrapolation of OS

Similar to the validation of modelling time-to-surgery, -metastases, or -death, there are no available data sources to validate the modelling of OS in the target population stipulated by the MHRA label. To obtain comparisons between modelled OS and RWE available, adjustments to align with the GB label population or account for real-world SOC were removed.

In the VHL-RCC cohort, the predicted OS curve in the SOC arm was plotted alongside the observed KM OS curve from the VHL Natural History Study cohort, reweighted to match key baseline characteristics in the MK-6482-004 trial population. In the VHL-CNS Hb and VHL-pNET cohorts, the predicted OS curves for SOC were plotted alongside the observed KM OS curves from the VHL Natural History Study cohorts with a pre-index history of CNS Hb and pNET, respectively, reweighted to match key baseline characteristics in the corresponding subgroups of the MK-6482-004 trial population. These comparisons of predicted OS for SOC versus observed OS in the VHL Natural History Study can, to an extent, be interpreted as external validations: The predicted OS curve for SOC is dependent on all transition probabilities in this arm, including transition probabilities that were not estimated using Natural History Study data (i.e., metastatic disease → death in all cohorts and pre-surgery → surgery in the VHL-CNS Hb and VHL-pNET cohorts).

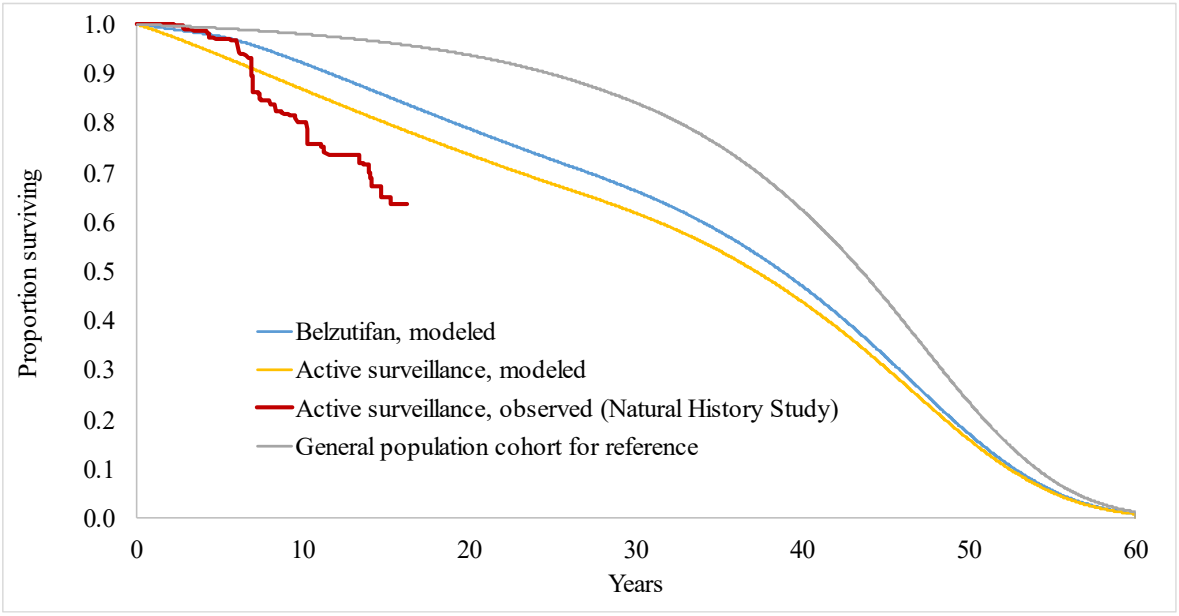
The resulting OS validations are presented in Figure 21. Table 70 provides the landmark estimates for OS in each cohort. As shown, modelled OS for SOC in the VHL-RCC cohort was higher in the long term than the observed OS curve from the VHL Natural History Study, suggesting that the modelled effectiveness results in the VHL-RCC cohort may be conservative with respect to belzutifan. Because modelled

OS is dependent on all transition probabilities and was not directly fitted to the OS curve from the VHL Natural History Study, some divergence between the modelled vs. observed OS curves is to be expected. Modelled OS for SOC was aligned with observed OS from the VHL Natural History Study in the VHL-CNS Hb and VHL-pNET cohorts. For the VHL-RCC cohort, the divergence was more pronounced and this may be due to the difference in the available treatments for advanced RCC in the cost-effectiveness model versus the VHL Natural History Study. Advanced RCC therapies in the cost-effectiveness model are based on the set of regimens recommended by NICE or listed as a preferred or recommended first-line regimen according to the National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines) and this may not be reflective of advanced RCC treatments historically available to patients in the VHL Natural History Study. Therefore, the modelled OS and observed OS are less likely to be aligned. This divergence was less pronounced for the VHL-CNS Hb and VHL-pNET cohorts. For the VHL-CNS Hb cohort there are no metastatic disease therapies since CNS Hb cannot metastasize. For the VHL-pNET cohort, there has been fewer developments in metastatic disease treatments in recent years so those included in the cost-effectiveness model are likely to be similar to those available to patients in the VHL Natural History Study.

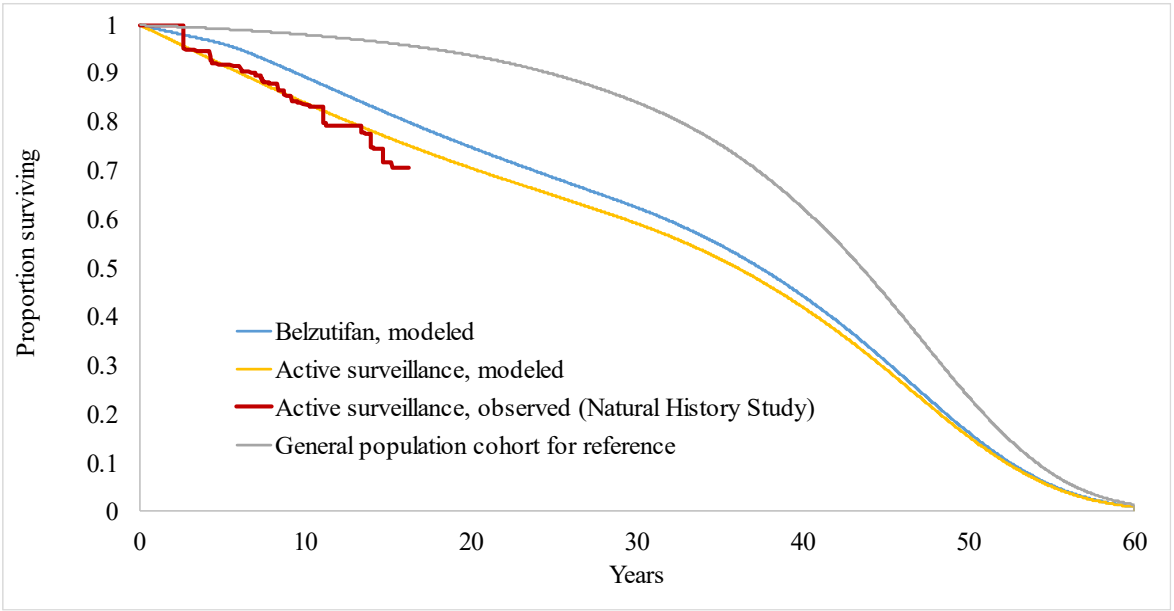
Clinical experts were also consulted to validate the efficacy inputs and other key model decisions (see B.3.14 Validation section below).

Figure 21 Validation of OS in the SOC arm against the observed OS from the VHL Natural History Study

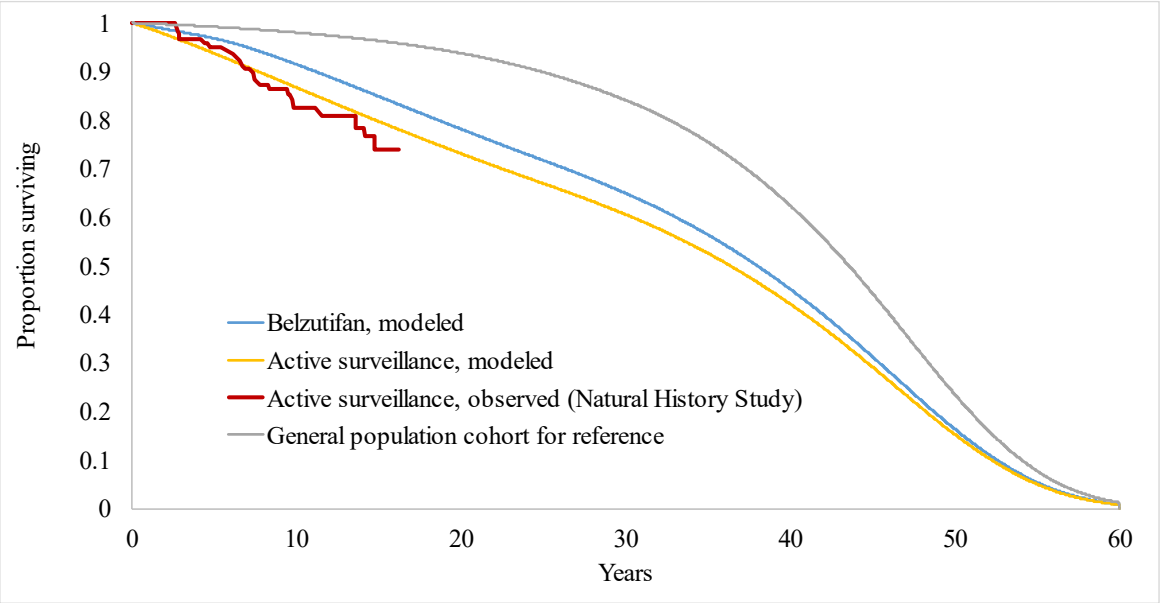
(a) VHL-RCC cohort



(b) VHL-CNS Hb cohort



(c) VHL-pNET cohort



Note: To obtain interpretable comparisons between modelled vs. observed curves, the modelled curves shown in figures (a) – (c) above were generated without performing any adjustments to align with the GB label population or account for real-world standard of care. Of note, the modelled OS curves for SOC are dependent upon all transition probabilities in that arm and were not directly fitted to the OS curves from the VHL Natural History Study. Consequently, even when not applying adjustments for real-world standard of care or the GB label population, some divergence is expected between the observed vs. modelled OS curves for active surveillance.
CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Table 70 Landmark estimates: OS by cohort

Time (years)	VHL-RCC		VHL-CNS Hb		VHL-pNET	
	Belzutifan	SOC	Belzutifan	SOC	Belzutifan	SOC
0.5	99.6%	99.0%	99.2%	98.4%	99.4%	98.5%
1	99.1%	97.5%	98.5%	96.6%	98.7%	96.2%
1.5	98.5%	95.6%	97.7%	94.7%	98.0%	93.3%
1.75	98.2%	94.5%	97.3%	93.7%	97.6%	91.7%
2	97.8%	93.4%	96.9%	92.7%	97.3%	90.0%
5	93.0%	76.5%	92.1%	79.4%	92.2%	66.6%
10	72.8%	47.8%	75.7%	57.2%	66.5%	34.1%
20	27.1%	14.6%	38.8%	26.4%	17.6%	7.2%
30	7.7%	3.8%	17.4%	11.5%	3.6%	1.4%
40	1.9%	0.9%	7.2%	4.7%	0.7%	0.3%
45	0.9%	0.4%	4.1%	2.6%	0.3%	0.1%
50	0.4%	0.2%	1.9%	1.2%	0.1%	0.1%
55	0.1%	0.1%	0.6%	0.4%	0.0%	0.0%
60	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%

CNS Hb: central nervous system haemangioblastoma; pNET: pancreatic neuroendocrine tumours; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Adverse events

The model considers Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan, or Grade ≥3 TRAEs occurring in >0% of patients treated with belzutifan. Grade ≥3 AEs were incorporated into the model due to their expected impact on resource utilisation and quality of life.

AE rates for patients treated with belzutifan were sourced from the MK-6482-004 trial, based on the proportions of patients with AEs reported for the all-subjects-as-treated population (see Table 71 Risks and durations of modelled AEs). In the SOC arm, it is assumed that the risk of all AEs is zero. Therefore, the model uses risks of drug-related (rather than all-cause) grade 3 to 5 AEs for the belzutifan arm as approximations of the incremental AE risks associated with belzutifan versus SOC.

Mean durations of the included AEs were sourced from the MK-6482-004 trial, and were used in the model to estimate the duration of the disutility and cost impact of each AE (described in Adverse reactions and Adverse reaction unit costs and resource use sections, respectively).

Table 71 Risks and durations of modelled AEs

Grade ≥3 AEs	Grades	Risks		Mean duration of AE (weeks)
		Belzutifan (%)	SOC (%)	
Anaemia	≥3	11.5	0.0	7.90
Fatigue	≥3	4.9	0.0	2.29

Mean duration of each AE type was obtained from the MK-6482-004 trial and reflects the average weeks per event multiplied by the number of events per patient who had the particular AE type.
AE: adverse event; SOC: standard of care

Time to treatment discontinuation

As specified in the MK-6482-004 trial protocol and the MHRA label, patients can remain on belzutifan treatment until unacceptable treatment-related toxicity or unequivocal disease progression. Belzutifan time on treatment (ToT) was modelled by fitting a Gompertz curve to patient-level data on time-to-treatment discontinuation in the MK-6482-004 trial. The Gompertz curve was selected in the base case as it was the best fitting of the 7 distributions investigated (exponential, Weibull, Gompertz, log-logistic, log-normal, gamma, and generalised gamma) according to

both AIC and BIC and yielded plausible ToT projections relative to other distributions. A comparison of fit statistics for the parametric models fitted to the belzutifan ToT data is presented in Table 72.

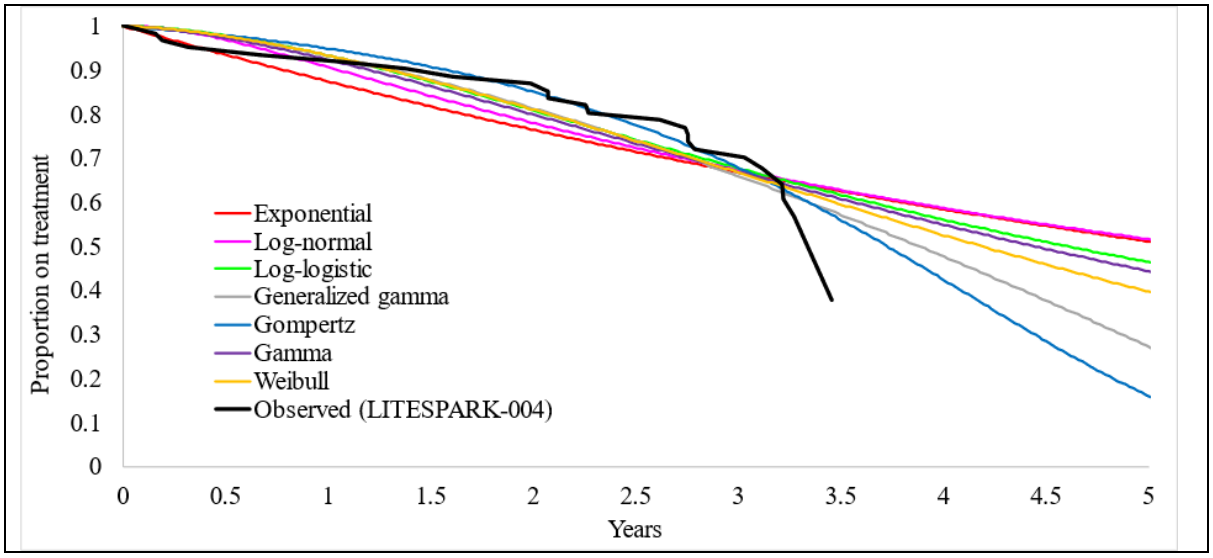
Table 72 Statistical fit for parametric models fitted to belzutifan ToT

Distributions fitted to ToT	AIC	BIC	Rank by AIC	Rank by BIC
Exponential	322.2	324.4	6	3
Weibull	319.2	323.5	2	2
Log-normal	325.2	329.5	7	7
Log-logistic	321.3	325.5	5	5
Gompertz	315.1	319.3	1	1
Gamma	320.5	324.7	4	4
Generalized gamma	319.3	325.6	3	6

AIC: Akaike information criterion; BIC: Bayesian information criterion; ToT: time on treatment.

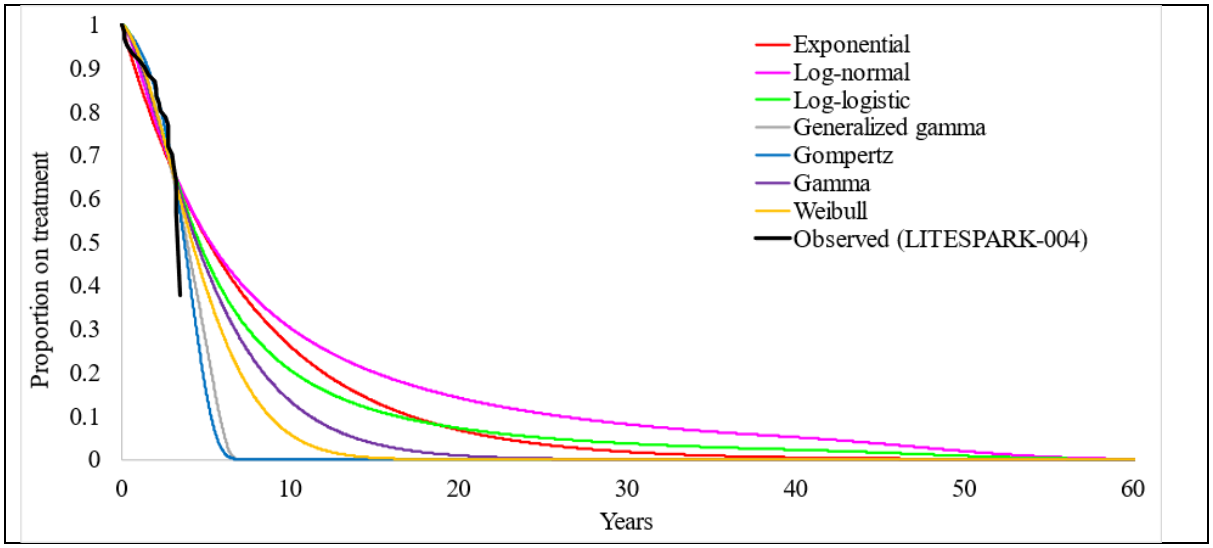
Gompertz was selected as the most plausible distribution given the shape of the KM curve and clinical plausibility of long-term projections of ToT. In the KM analysis of ToT, the number of at risk was 45 patients at 36 months and decreased to only 16 patients by 40 months and 3 patients by 42 months. A Weibull distribution fitted to ToT was explored in scenario analyses as the second best-fitting curve. The observed ToT and modelled ToT for belzutifan for the available trial period and for longer time horizons are presented in Figure 22 and Figure 23, respectively, and the base case.

Figure 22 Modelled vs. observed ToT for belzutifan in MK-6482-004 trial under different parametric distributions during the MK-6482-004 trial



Notes: As of the 1 April 2022 data cutoff date, 62.3% of patients (38/61) were continuing to receive belzutifan.
LITESPARK-004 refers to the MK-6482-004 trial
KM: Kaplan-Meier; ToT: time on treatment

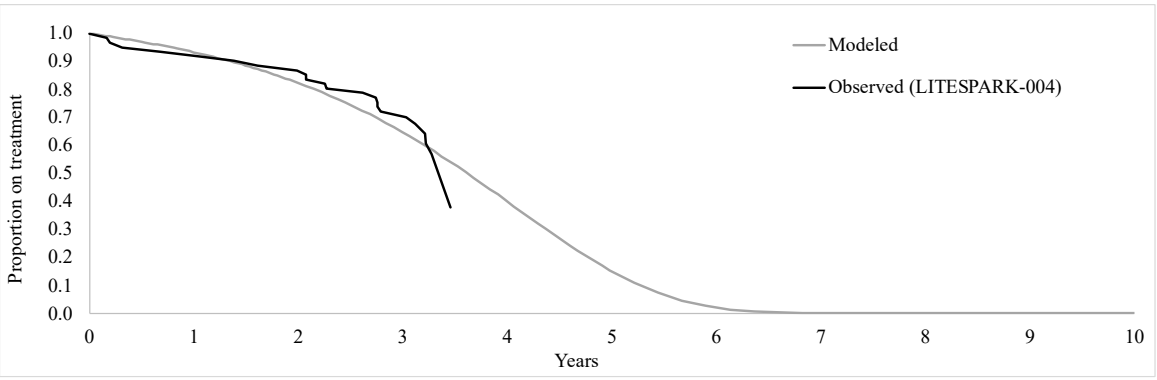
Figure 23 Modelled vs. observed ToT for belzutifan in MK-6482-004 trial under different parametric distributions over long-term extrapolation



Note: LITESPARK-004 refers to the MK-6482-004 trial
ToT: time on treatment

The modelled ToT used in the base case is presented in Figure 24.

Figure 24 Belzutifan ToT modelled using a Gompertz distribution



Note: LITESPARK-004 refers to the MK-6482-004 trial
ToT: time on treatment

Treatment effect waning following belzutifan discontinuation

Patients can remain on belzutifan treatment until unacceptable treatment-related toxicity or unequivocal disease progression as per the MHRA label. As of the 01-April-2022 data-cut, median follow-up of 37.7 months (3.14 years), the majority of patients in the MK-6482-004 trial (61%) remained on treatment. No other clinical trial assessing belzutifan currently provides longer follow-up data. There is therefore limited data on the impact of treatment discontinuation on the treatment effect of belzutifan in shrinking tumours, reducing the risks of surgery and metastases in VHL patients.

The modelled risk of transitioning from the pre-surgery health state (to either surgery, death or metastatic disease) in the belzutifan arm is based on data collected in MK-6482-004, which reflects the discontinuations that did occur during the trial follow-up. In the long term, the proportion of patients discontinuing treatment increases based on the modelled ToT curve. The potential impact of this reduction in treatment effective is included in the base-case analysis.

When the impact of discontinuation on treatment effect is explored in health economic analyses, four features are commonly specified:

1) *data source informing efficacy following cessation of treatment effect:* assuming cessation of treatment effect beyond a certain time point, long-term efficacy is informed by an alternative data source, often the standard of care or trial comparator arm

2) *how waning of efficacy parameters is implemented in the economic model:* As treatment effect gradually wanes, efficacy data is often reflected as a blend of efficacy data from the treatment and comparator arms from the pivotal clinical trial or other data sources

3) *time point post discontinuation when waning is initiated:* any plausible waning of treatment effect would take effect gradually following treatment discontinuation.

4) *time point post discontinuation when waning is completed:* treatment effect from this time point onwards is fully informed by an alternative data source

To account for the uncertainty around the impact of treatment discontinuation on long-term efficacy, an assumption of treatment effect waning is reflected in the base case, the features of which are detailed below.

Data source informing efficacy following cessation of treatment effect

When treatment effect begins to gradually wane, the hazards of transitioning from the pre-surgery health state are estimated to converge to those associated with SOC, based on data from the VHL Natural History Study. Treatment with belzutifan was associated with high response and disease control rates in the MK-6482-004 trial, providing potentially transformative clinical benefits to patients. Given the uncertainty surrounding long-term outcomes post discontinuation, waning to a data source which reflects an ongoing risk of surgery was considered to strike an appropriate balance between the favourable clinical results from the MK-6482-004 trial at more than three years' follow-up, and the uncertainty around the impact of discontinuation potentially taking effect many years after treatment initiation.

How waning of treatment effect is implemented in the economic model

When treatment effect waning is assumed, it is implemented in the model by using a corresponding "off-treatment" health state for all health states except metastatic disease (given patients receive subsequent therapies treating metastatic disease) and death. State membership in an "off-treatment" state is based on the discontinuation rate in a given cycle based on the modelled ToT for belzutifan. In these off-treatment health states, the clinical efficacy parameters of patients in the

belzutifan arm were assumed to gradually converge over time towards those of SOC. The following belzutifan clinical efficacy parameters are subject to treatment effect waning following treatment discontinuation:

- All transition probabilities starting from the pre-surgery and event-free after surgery state
- Overall response rates which inform composite utility values
- Incidence of non-primary tumour surgeries

Time point post-discontinuation when waning is initiated

Following discontinuation based on modelled ToT, treatment effect waning is applied after a time period equivalent to the maximum follow-up in the MK-6482-004 trial following discontinuation (up to 3.84 years, as detailed in Table 73). As described above, belzutifan efficacy estimated from the MK-6482-004 trial data during this time period already accounts for any impact of belzutifan discontinuation. Therefore, it would be implausible to consider treatment effect waning to occur before an equivalent time period has elapsed since treatment discontinuation. The time period between treatment discontinuation and the initiation is referred to as the period of residual treatment benefit.

Time point post-discontinuation when waning is completed

At the time of discontinuation, the size of a patient's largest tumour will have been considerably reduced compared to what it would have been in the absence of belzutifan treatment, as evidenced by the high response rates reported in MK-6482-004. Starting from the discontinuation time point, the time it would subsequently take for the largest tumour to reach the size warranting surgery would be longer than what it would be for patient who had been treated with SOC rather at a similar time point. The risk of surgery for a belzutifan-treated patient at the time of discontinuation would therefore not be equivalent to a patient having previously received only SOC, due to these patients having different size tumours at this time point. This would be true even if the tumour growth rate immediately reverted post discontinuation to the pre-treatment growth rate. Therefore, assuming an instantaneous switch to SOC hazards following treatment discontinuation would be implausible.

To reflect this, treatment effect waning was assumed to initiate from the end of the period of residual treatment benefit and take effect gradually over a 2.71-year period. This time period is based on the magnitude of tumour shrinkage observed in MK-6482-004 compared to the tumour growth rate in the time period prior to belzutifan initiation. In the MK-6482-004 trial, at the tumour measurement that occurred closest to the time of discontinuing belzutifan, the average size of patients' largest RCC tumour was smaller than at baseline. Based on a conservative assumption that the rate of tumour growth immediately following discontinuation reverts to its pre-treatment growth rate, 2.71 years represents the amount of time until the largest RCC tumour resets to its baseline size. This duration is based on an estimate of annual tumour growth rate among patients treated with belzutifan in the MK-6482-004 trial. (See Table 73 below for further details on the calculations of treatment effect waning parameters.) A similar assumption was made for the pNET and CNS Hb cohorts, due to the small sample size of discontinued patients in the CNS Hb and pNET subgroups who had an available CNS Hb and pNET measurement near to the time of treatment discontinuation.

The impact of discontinuation on treatment effect in the outcomes reflected in the economic model is a source of uncertainty. Therefore, alternative treatment effect waning assumptions were tested in scenario analysis in order to quantify the effect of this uncertainty on cost-effectiveness results.

Table 73 Base-case assumptions for treatment effect waning time period

Treatment waning period	Value (years)	Rationale
Time point to initiate treatment waning	3.84	• In the base-case analysis, treatment effect waning was initiated at 3.84 years (46.1 months), the maximum follow-up duration available from MK-6482-004 as of the 01 Apr 2022 data cut-off date. • There is no evidence to support initiating treatment effect waning before the end of the trial period, as the estimation of transition probabilities used all available follow-up from the trial (rather than only the portion of follow-up before discontinuation). Consequently, applying treatment effect waning during the observed trial period would worsen the alignment between observed vs. predicted curves for time to surgery, metastatic disease, or death in the belzutifan arm.
Time point to complete treatment waning	6.55	• Treatment waning was assumed to occur gradually over a 2.71-year period from the end of the maximum follow-up in the (i.e., 6.55 = 3.84 + 2.71 years). This 2.71-year period approximates the period of residual benefit that patients are expected to receive from belzutifan beyond the time of discontinuation: At the tumour measurement that occurred closest to the time of discontinuing belzutifan (average of 25.2 days from discontinuation), the size of patients' largest RCC tumour was smaller on average (15.36 mm) than the average size of the largest RCC tumour at baseline (24.9 mm). Assuming that the tumour growth rate immediately reverts to pre-treatment levels (average of 3.52 mm/year before treatment) after discontinuation, 2.71 years represents the amount of time until the largest RCC tumour resets to baselin– levels (i.e., (24.9 - 15.36)/3.52 = 2.71 years) (48, 62). • The same residual benefit period was assumed for the pNET and CNS Hb cohorts, due to the small sample size of discontinued patients in the CNS Hb and pNET subgroups who had an available CNS Hb and pNET measurement (respectively) near the time of discontinuation.

B.3.4 Measurement and valuation of health effects

Health-related quality-of-life data from clinical trials

HRQoL data (including EQ-5D) were not collected as part of the MK-6482-004 trial. In the absence of data from the trial, the economic analysis incorporates EQ-5D data from the VHL RW QoL Disease Burden Study and KEYNOTE-564 trial (in the CR level of the pre-surgery, surgery, and event-free after surgery states) as the primary sources of HRQoL data (44).

The VHL RW QoL Disease Burden Study is a 2021 cross-sectional patient survey conducted in 220 adult patients in the UK, USA, Canada, France and Germany with a VHL diagnosis and manifestations in the kidneys, pancreas, or CNS. This study could not be used to estimate the utility associated with CR (see *Health state utility values*) which was sourced instead from the KEYNOTE-564 trial. The KEYNOTE-564 trial was a phase 3 placebo-controlled clinical trial of pembrolizumab in adult patients with adjuvant treatment of RCC post-nephrectomy. The utility associated with CR in this trial is used for all three VHL cohorts since all patients in the MK-6482-004 trial had at least one measurable solid RCC tumour.

Mapping

No HRQoL data was collected from the MK-6482-004 trial. For the EQ-5D data collected from the VHL RW QoL Disease Burden Study and KEYNOTE-564, the EQ-5D-5L score was mapped onto the UK EQ-5D-3L value set as per the NICE reference case (63). The 3L value set was then used to derive utility values for the economic model.

Health-related quality-of-life studies

An SLR to identify relevant HRQoL studies was conducted in July 2020, and subsequently updated in July 2022 to cover the period from July 2020 to July 2022. Appendix H: Health-related quality-of-life studies provides full details of the methods, and an overview and results of the identified studies, as well as quality assessments. The SLR identified four studies that reported HRQoL data in patients, of which two presented HRQoL data in a tumour-specific population (Siller et al. 2017 and Rochette et al. 2018 both reported HRQoL data from patients with VHL-CNS Hb (64, 65)). The remaining two studies were conducted in non-tumour specific VHL populations, therefore no HRQoL studies were identified specifically in the VHL-RCC or VHL-pNET populations. There have been no previous NICE appraisals in VHL, therefore no HRQoL data could be identified from these sources.

In the absence of published HRQoL studies in the three primary tumour site populations, the VHL RW QoL Disease Burden Study was the most relevant source of HRQoL data for patients with VHL in each health state. In line with NICE guidance, this study collected EQ-5D scores and also collected patients' disease

status, time since most recent surgery (less than/greater than 6 months) and those with one vs. multiple tumour manifestations. The survey included patients who range from relatively well, to severely unwell due to VHL. The mean utility score (EQ-5D Crosswalk from 5L to 3L) was 0.699 (SD: 0.27) using the UK value set. Importantly, the range was 0.240 to 0.988 reflecting the variation in impact VHL can have on patients.

As with the MK-6482-004 trial and VHL Natural History Study, the VHL RW QoL Disease Burden Study is not fully generalisable to the MHRA license population, as the survey was not limited to only those patients who require therapy and for whom localised procedures are unsuitable or undesirable. As a result, the HRQoL estimates from the patients in the VHL RW QoL Disease Burden Study is likely an overestimate of HRQoL versus the licensed population.

Adverse reactions

The impact of adverse reactions to belzutifan on HRQoL was explored in the economic analysis. In line with standard practice for NICE appraisals, the model considers Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan or Grade ≥3 TRAEs occurring in >0% of patients treated with belzutifan. These AEs are relevant to the economic model as they are expected to have an impact in terms of resource use or HRQoL. AE rates are sourced from the MK-6482-004 trial.

AE-related disutility is applied as a one-time QALY decrement in the first cycle following initiation with belzutifan. The disutility associated with AEs was calculated in each treatment arm as a function of:

- Treatment-specific AE risk
- Mean duration of AEs per affected patient in the MK-6482-004 trial
- Estimated disutility associated with AE based on regression analyses of EQ-5D data from the KEYNOTE-564 study

The disutility values of AEs used in the base case are presented in the Health-related quality-of-life data used in the cost-effectiveness analysis section.

Health-related quality-of-life data used in the cost-effectiveness analysis

In line with NICE's preference for the use of EQ-5D data to inform QALY estimates, the economic analysis uses EQ-5D data to inform utilities for each health state and disutilities related to surgical complications, disutilities due to AEs, and disutilities due to VHL-associated tumours at non-primary tumour sites. The economic analysis also considers background disutility due to ageing, based on Ara and Brazier (2010) (66). It should be noted that the EQ-5D is limited in its ability to reflect impact of belzutifan this patient population. A study exploring the HRQoL aspects that are not captured by the EQ-5D-5L in chronic conditions found that that the most common aspects not captured for patients with rare diseases were fatigue and impact on relationships and social life (67). This is even more relevant for people with VHL disease due to the frequency of monitoring and the impact of this chronic disease on loved ones. The benefit of belzutifan in reducing the size of tumour manifestations and alleviating fear for patient's loved ones is not adequately captured by the EQ-5D.

While efforts have been made to ensure the utility values used in the economic analysis reflect the target population of belzutifan, the sources of the utility values used in the economic analysis (described further in the subsequent sections) did not measure utility only in patients who require therapy and for whom localised procedures are unsuitable or undesirable, as per the licenced indication of belzutifan. The licensed population has more severe disease (and hence would be expected to have worse utility scores) than the population informing the utility data in the economic analysis. This means that the effect of belzutifan on HRQoL is underestimated in the economic analysis.

The complex nature of VHL disease leads to complexities in modelling patient trajectories and patient and carer HRQoL. As previously described, clinical management of VHL requires the balance of:

1. Prevention of metastatic disease originating from RCC or pNET tumours
2. Maintenance of organ function

3. Minimisation of burdensome symptoms, particularly for patients with VHL-associated CNS Hb.

The interaction of these three components of clinical management differs between patients and over time. The dominant influence on a patient's quality of life is therefore driven by the 'worst' burden they are experiencing. Hence, the model includes utilities (and disutilities) representing each component.

Health state utility values

Utility in the non-metastatic health states

Health state utilities were estimated based on data collected in the VHL RW QoL Disease Burden Study, as described in Health-related quality-of-life studies. This was a large (n=220) and recent (2022) survey, which collected EQ-5D data from patients who were aligned with the MK-6482-004 trial population (adult patients with a VHL diagnosis with manifestations in the kidney, CNS, or pancreas).

For the pre-surgery, surgery, and event-free after surgery health states, a better response is associated with a higher utility value, as a better response avoids the complications associated with tumour growth and the greater risk of metastases resulting from disease progression, which would reduce patients' HRQoL. A better response is more likely to delay disease progression and hence maintain higher HRQoL for longer. The economic analysis therefore uses response-adjusted utility values for each primary tumour site population in the pre-surgery, surgery, and event-free after surgery health states, derived using a weighted average of the utility values from the VHL RW QoL Disease Burden Study. This was calculated by applying the distributions of best OR level (i.e. CR, PR/SD, or PD) from the MK-6482-004 trial and VHL Natural History Study for the belzutifan and SOC arms, respectively, to the utility scores by best response that were calculated based on data from the VHL RW QoL Disease Burden Study.

The pre-surgery, surgery and event-free after surgery health state utility values were based on self-reported tumour response status and EQ-5D-5L responses from patients in the VHL RW QoL Disease Burden Study who reported having non-

metastatic disease. Patients were asked to report the response level of their VHL-associated tumours according to their doctor and latest imaging results (response status for specific tumour manifestations was not requested). Only one patient reported having achieved a CR or and 8 patients reported a PR. The patient who reported CR and four patients who reported PR were not currently being treated with any medication for VHL-related cancer; therefore, these patients were assumed to have SD. This was based on clinical expert feedback indicating that a spontaneous reduction in the size of VHL-related tumours was very unlikely in the absence of active treatment. Hence, the utility value associated with CR estimated from this data set was not considered plausible. Given the small number of patients with PR and the high potential for misclassification amongst the PR and SD categories based on patient responses, a utility value pooled across the PR and SD categories was calculated (combined n=116, including 4 PR and 112 SD after reclassifying 5 untreated patients as described) and calculated separately for PD (N=49). The EQ-5D-5L score was mapped onto the UK EQ-5D-3L value set. Table 74 presents the base-case utility inputs by specific health state.

HRQoL in patients in the VHL-CNS Hb cohort who have PD, and for whom localised procedures are unsuitable or undesirable, was expected to be particularly poor. The growth of CNS Hb can result in severe neurological disability. Because patients with CNS Hb in the VHL RW QoL Disease Burden Study were not selected for being unsuitable or undesirable for localised procedures, the utility value estimated for VHL CNs Hb patients in this study was not considered representative of the population eligible to receive belzutifan per the MHRA label. Based on input 193ymptominicians on the symptoms experienced by VHL CNS Hb patients with PD who are unsuitable or undesirable for localised procedures, the utility associated with motor neurone disease is considered an appropriate proxy. Therefore, the utility value associated with PD in patients with CNS Hb was obtained from a study reporting in patients with motor neuron disease conducted by Kiebert et al. (2001) (68), which estimated a mean utility value of 0.550 based on structured interviews of 77 patients with amyotrophic lateral sclerosis.

The utility associated with CR (used in the non-metastatic health states) could not be estimated from the VHL RW QoL Disease Burden Study due to the implausibility described above. Therefore, the utility value associated with CR was approximated using the utility value estimated from data collected in the KEYNOTE-564 trial for the in patients post nephrectomy at higher risk of disease recurrence who were treated with pembrolizumab in the adjuvant setting. Patients in this trial were considered representative of VHL patients with the most favourable prognosis and HRQoL. The disease-free state among patients in KEYNOTE-564 state was considered suitably representative of CR status based on the definition of CR according to RECIST v1.1 criteria (i.e., disappearance of all target lesions, with any pathological target or non-target lymph nodes reduced in short axis to <10 mm). The disease-free (without toxicity) utility was previously estimated for NICE TA830 based on EQ-5D-5L data from KEYNOTE-564 during patient-visits in which patients remained disease free (45). (The AE-related disutility presented in the *Disutility due to adverse events* section was obtained from the same analysis.)

A weighted average of response status utilities was used to estimate non-metastatic health state utilities to capture the distribution across response levels in each health state. When calculating the weighted average of the CR, PR, SD, and PD utilities, the relative weight of each utility in the non-metastatic health states was based on: ORR results from the MK-6482-004 trial for belzutifan; and patient-reported response level among untreated VHL RW QoL Disease Burden Study respondents for SOC (Table 75). No achievement of CR/PR is expected under the SOC arm, as in the absence of systemic treatment prior to surgery, a spontaneous reduction in tumour is highly unlikely based on clinicians' feedback. The resultant weighted averages of the CR, PR, SD, and PD utilities (shown in Table 76) were used in all non-metastatic health states, rather than just the pre-surgery state, as patients can continue to receive belzutifan and achieve/maintain CR, PR, or SD following surgery. (Differences in health-related quality of life between the pre-surgery, surgery, and event-free after surgery states were captured through the application of surgical complication-related disutilities, described in the *Disutility due to surgical complications* section.) However, despite the absence of evidence of patients losing response in the MK-6482-004 trial, it is assumed that after discontinuing belzutifan,

the distribution across the CR, PR, SD, and PD response levels would linearly converge towards those associated with the SOC arm in line with how efficacy inputs are estimated following discontinuation (see *Treatment effect waning following belzutifan discontinuation*).

Utility in the metastatic disease state

In the metastatic disease state, utility values were assigned to the pre- and post-progression metastatic disease sub-states based on average EQ-5D-5L utility by self-reported progression status among patients with metastatic disease in the VHL RW QoL Disease Burden Study. Overall utility in each treatment arm in the metastatic disease health state is calculated as a weighted average of the utilities associated with pre-progression and post-progression metastatic disease, based on the estimated proportion of time spent progression-free in the metastatic disease state, as determined by the ratio of PFS to OS in the metastatic disease state (estimated using an NMA as described in *Transitions from metastatic disease to death* section). The ratio of PFS to OS is based on a weighted average of expected PFS and OS for each first-line metastatic disease treatment and the market shares of first-line metastatic disease treatments in each origin tumour. Because patients in the belzutifan and SOC arms were expected to receive the same mix of first-line treatments upon developing metastatic disease, overall utility in the metastatic disease state was the same in both arms.

Health state utility values by response status (for non-metastatic states) and by progression status (for the metastatic disease state) used in the base-case analysis are summarised in Table 74.

Table 74 Health state utilities in the base case by response / progression status

Health state	Utility		Sources
	Value	SE	
Non-metastatic states (pre-surgery, surgery, and event-free after surgery states), by objective response*			
CR	0.868	(0.005)	KEYNOTE-564 (data cutoff date: 14 Dec 2020)
PR	0.754	(0.019)	VHL RW QoL Disease Burden Study
SD	0.754	(0.019)	VHL RW QoL Disease Burden Study

PD *(in the VHL-RCC and -pNET cohorts)*	0.665	(0.036)	VHL RW QoL Disease Burden Study
PD *(in the VHL-CNS Hb cohort)*	0.550	(0.025)	Kiebert et al. (2001)
Metastatic disease state			
Metastatic disease (pre-progression)	0.525	(0.205)	VHL RW QoL Disease Burden Study
Metastatic disease (post-progression)	0.412	(0.176)	VHL RW QoL Disease Burden Study

*Disutilities associated with surgical complications are applied separately
CR: Complete Response; PD: Progressive Disease; PR: Partial Response; SD: Stable Disease; VHL: Von Hippel Lindau

Table 75 Distribution of objective response level by VHL cohort and treatment arm used to calculate utility values in the pre-surgery, surgery, and event-free after surgery states

Cohort / Treatment arm	Objective response level*				
	Complete response	*Partial response*	*Stable disease*	*Progressive disease*	*Not evaluable***
RCC population					
Belzutifan	6.6%	57.4%	34.4%	0.0%	1.6%
SOC	0.0%	0.0%	57.9%	23.3%	18.9%
CNS Hb population					
Belzutifan	8.0%	36.0%	46.0%	6.0%	4.0%
SOC	0.0%	0.0%	57.9%	23.3%	18.9%
pNET population					
Belzutifan	31.8%	59.1%	9.1%	0.0%	0.0%
SOC	0.0%	0.0%	57.9%	23.3%	18.’%

*For SOC, patients' distribution across response categories was approximated based on self-reported response status among patients in the VHL RW QoL Disease Burden Study (2022) who reported receiving no prescribed medication for VHL-related cancer (N=159). Untreated patients in the QoL study who reported complete response (1 patient) or partial response (4 patients) were assumed to have stable disease, based on clinical expert feedback that a spontaneous reduction in the size of VHL-related tumours is very unlikely in the absence of treatment.
**When calculating the weighted average of utility in each non-metastatic health state, patients in the “not evaluable” category are proportionally redistributed to the other categories.
CNS Hb: central nervous system haemangioblastoma; CR: Complete Response; PD: Progressive Disease; pNET: pancreatic neuroendocrine tumours; PR: Partial Response; RCC: renal cell carcinoma; SD: Stable Disease; SOC: standard of care; VHL: Von Hippel Lindau

Table 76 Response-adjusted overall utility in non-metastatic health states (pre-surgery, surgery, and event-free after surgery)

Cohort / Treatment arm	Response-adjusted utility in each non-metastatic health state *(calculation)*
VHL-RCC	
Belzutifan	0.762

SoC	0.728
VHL-CNS Hb	
Belzutifan	0.751
SoC	0.695
VHL-pNET	
Belzutifan	0.790
SoC	0.728

Disutility due to surgery and surgical complications

Patients in the model cohort are also subject to disutility due to surgeries and surgical complications for VHL-associated tumours. The disutility associated with the perioperative recovery from VHL-pNET surgery is considered along with the risks of both short-term and long-term surgical complications for each VHL-associated surgery type. The disutility associated with the perioperative recovery period following VHL-pNET surgery is derived from a cost-effectiveness study comparing laparoscopic versus open distal pancreatectomy for pancreatic cancer (69). The rates of short-term complications were derived from real-world data from the Optum Clinformatics Data Mart claims study during the 28-day period following each surgery. The rates of long-term complications were similarly derived from the Optum Clinformatics Data Mart claims study, measured over a 180-day period. The following long-term complications were considered: chronic pain, cerebral vasculature occlusion/stroke, seizure, neurological complications, and secondary diabetes or exocrine pancreatic insufficiency).

Disutility due to the perioperative recovery associated with surgery

It is expected that there would be significant disutility associated with the perioperative recovery VHL-associated tumour procedures, particularly with primary tumour surgeries modelled in this analysis that are a ‘last-resort’ surgery. We conservatively assume no disutility associated with the perioperative recovery period for VHL-RCC and VHL-CNS Hb associated surgeries due to a lack of available data for this input. For the disutility associated with perioperative recovery from VHL-pNET associated surgery, a disutility of -0.186 is assumed based on a cost-effectiveness study assessing pancreatectomy for pancreatic cancer (69). This disutility was calculated by subtracting the subsequent stable period following distal

pancreatectomy from the utility associated with complicated open distal pancreatectomy in the first 3 months. This disutility is applied for a 28-day period to reflect the perioperative recovery period.

Disutility due to surgical complications

The disutility of each short-term surgical complication and the surgery itself is applied to the 28-day period (or 180-day period for the long-term complications described above) following the surgery, in accordance with the timeframe in which the risks of these complications were measured. The disutility of long-term complications is applied to the proportion of patients who experienced each complication in the Optum Clinformatics Data Mart claims study in all cycles starting from the first surgery until death or the end of the modelled time horizon.

Disutility values for each surgical complication were identified from published literature sources and estimated from data collected in the VHL RW QoL Disease Burden Study (see Table 77 and Table 78). Where available, disutilities of long-term complications were derived from the VHL RW QoL Disease Burden Study, based on the difference in average utility between patient with versus without specific co-morbidities. For other, less-common long-term complications that were not assessed in the VHL RW QoL Disease Burden Study, disutilities were obtained from published literature sources. The presence/absence of short-term complications and there disutilities were not feasible to measure within the cross-sectional VHL RW QoL Disease Burden Study, as these disutilities need to be measured while a patient is actively experiencing the complication. The acute nature of short-term surgical complications limits the timeframe in which their disutility can be measured; moreover, patients may be unlikely to complete the survey while experiencing an acute surgical complication. All disutilities of short-term surgical complications were therefore obtained from published sources. The incremental disutility resulting from long- and short-term complications of surgery for other VHL tumour is also determined by the incidence rate of surgeries and surgical complications for these tumours per cycle and sourced using the same methods described above.

Disutility values applied to long-term and short-term surgical complications in the base case are summarised in Table 77 and Table 78, respectively.

Table 77 Long-term surgical complication disutility values in the base case

Complication	Disutility	Source
VHL-associated RCC surgery		
End stage renal disease and/or dialysis*	-0.527	Lee et al. (2005) (weighted average of hemodialysis and peritoneal dialysis disutilities) (70)
Chronic kidney disease*	-0.136	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic kidney disease) (44)
Hernia surgery	-0.200	Simianu et al. (2020) (difference in utility with hernia complication vs. baseline state) (71)
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain) (44)
Cerebral vasculature occlusion or stroke	-0.370	Gandhi et al. (2012) (non-fatal stroke disutility) (72)
VHL-associated CNS Hb surgery		
Chronic pain (in CNS Hb population)	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain) (44)
Cerebral vasculature occlusion or stroke	-0.370	Gandhi et al. (2012) (non-fatal stroke disutility) (72)
Seizure	-0.270	Assumed equal to neurological complications
Neurological complications	-0.270	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without motor loss/ataxia) (44)
VHL-associated pNET surgery		
Chronic pain (in pNET population)	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain) (44)
Secondary diabetes or exocrine pancreatic insufficiency*	-0.042	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without diabetes) (44)
Immunocompromisation*	-0.081	NICE Committee Papers for GID-TA10024 (everolimus in neuroendocrine tumors), based on difference between the mean utility values for (stable disease without AE) minus (stable disease with leukopenia AE) (73)
Disutility due to surgery for other VHL-associated manifestations		
Complications of adrenal lesion surgery		
Adrenal insufficiency	-0.042	Assumed equal to diabetes
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain) (44)
Complications of retinal Hb surgery		
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain) (44)

CNS: central nervous system; Hb: hemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; VHL: Von Hippel Lindau

*This is a metabolic complication resulting from limited/absent organ function following last-resort surgery

Complication	Disutility	Source

Table 78 Short-term surgical complication disutility values in the base case

Complication	Disutility	Source
VHL-associated RCC surgery		
Acute renal failure	-0.150	Nisula et al. (2013) (74), as cited in Zargar et al. (2018) (acute kidney injury disutility) (75)
Cardiac complications	-0.240	Gandhi et al. (2012) (non-fatal myocardial infarction disutility) (72)
Erythroderma	-0.335	Poole et al. (2010) (severe atopic dermatitis) (76)
Kidney infection	-0.340	Stevenson et al. (2014) (kidney infection disutility) (77)
Other genitourinary complications	-0.255	Stevenson et al. (2014) (simple average of urinary obstruction and incontinence disutilities) (77)
Postoperative infection (RCC-related)	-0.360	Stevenson et al. (2014) (abscess disutility) (77)
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility) (78)
Thrombosis and/or embolism	-0.330	Stevenson et al. (2014) (deep vein thrombosis disutility) (77)
Vascular injury or anaemia	-0.073	Nafees et al. (2008) (approximated by disutility of fatigue) (79)
VHL-associated CNS Hb surgery		
Acute renal failure	-0.150	Nisula et al. (2013) (74), as cited in Zargar et al. (2018) (acute kidney injury disutility) (75)
CNS hemorrhage	-0.240	Wang et al. (2020) (minor intracranial hemorrhage) (80)
Nerve palsy related to anesthesia	-0.120	Memeh et al. (2020) (temporary unilateral laryngeal nerve injury) (81)
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility) (78)
Thrombosis and/or embolism	-0.330	Stevenson et al. (2014) (deep vein thrombosis disutility) (77)
Vascular injury or anaemia	-0.073	Nafees et al. (2008) (approximated by disutility of fatigue) (79)
VHL-associated pNET		
Abdominal abscess	-0.360	Stevenson et al. (2014) (abscess disutility) (77)
Postoperative infection (pNET-related)	-0.360	Stevenson et al. (2014) (abscess disutility) (77)
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility) (78)
Thrombosis and/or embolism	-0.330	Stevenson et al. (2014) (deep vein thrombosis disutility) (77)
Urinary tract infection	-0.270	Stevenson et al. (2014) (urinary tract infection) (77)
Vascular injury or anaemia	-0.073	Nafees et al. (2008) (approximated by disutility of fatigue) (79)

Complication	Disutility	Source
Perioperative recovery after pNET surgery	-0.186	Gurusamy et al. (2017) (utility of complicated open distal pancreatectomy in first 3 months minus utility of subsequent stable period) (69)
Disutility due to surgery for other VHL-associated manifestations		
Complications of adrenal lesion surgery		
Acute renal failure	-0.150	Nisula et al. (2013) (74), as cited in Zargar et al. (2018) (acute kidney injury disutility) (75)
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility) (78)
Thrombosis or embolism	-0.330	Stevenson et al. (2014) (deep vein thrombosis disutility) (77)
Vascular injury or anaemia	-0.073	Nafees et al. (2008) (approximated by disutility of fatigue) (79)
Complications of endolymphatic sac tumour surgery		
Acoustic impairment	-0.150	Verkleij et al. (2021) (moderate unilateral hearing loss) (82)
Complications of retinal Hb surgery		
Vitreous haemorrhage	-0.223	Assumed equal to vision loss disutility derived from Ament et al. (2018) (neurological complication: visual loss at 2 months)

CNS: central nervous system; Hb: haemangioblastoma; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; VHL: Von Hippel Lindau

Disutility due to adverse events and ageing

Disutility due to AEs is also included in the model, based on the rates of Grade ≥3 TRAEs occurring in ≥5% of patients observed in the MK-6482-004 trial, as described in the Adverse reactions section. The AE-related disutility was applied as a lump-sum QALY decrement at model entry. The QALY decrement associated with AEs was calculated in each treatment arm as a function of: treatment-specific AE risks (Adverse reactions section); the mean durations of these AEs per affected patient in the MK-6482-004 trial (Adverse reactions section); and the estimated disutility associated with an active grade 3+ AE (Table 79). The disutility of grade 3+ AEs was obtained from an analysis of EQ-5D-5L data from KEYNOTE-564 that was previously conducted for NICE TA830 (45), and represents the difference in utility associated with disease-free (without toxicity) vs. disease-free (during any grade 3+ AE) in KEYNOTE-564.

Additionally, a background disutility related to aging of the cohort over time is applied within the model. This is based on a published linear regression model (presented in Table 80) from Ara and Brazier (2010) (66), which predicts mean utility values for

individuals within the general population, conditional on age (in years), age-squared, and gender.

Table 79 Summary of adverse event disutility values in the base case

AE	Disutility	Source
Anaemia	-0.06417	KEYNOTE-564 (data cut-off date: 14 Dec 2020) (45)
Fatigue	-0.06417	

AE: adverse event

Table 80 Regression coefficients used for the estimation of age-related disutility

Parameter	Coefficient	Source
Age (years)	-0.0002587	Ara et al. (2010) (66)
Age2	-0.0000332	
Male	0.0212126	
Intercept	0.9508566	

Disutility of caregivers

VHL disease is a severe condition with a profound impact on the health status and well-being of patients' caregivers and close family members, who are likely to experience anxiety for the patient, fear of tumour recurrences, and bereavement in the event of the patient's premature death (see B.1.3 Health condition and position of the technology in the treatment pathway). Additionally, during the patient's lifetime, caregivers may carry multiple responsibilities such as providing physical care and emotional support to the patient, scheduling and coordinating healthcare services, and managing disease-related finances; therefore, the disutility of caregivers was considered in scenario analysis. This was modelled according to patients' health state distribution in each cycle, based on published studies conducted among family members and caregivers of cancer patients. Caregiver disutility values were identified from a targeted review of published literature sources. The review identified no publications that examined caregiver HRQoL impairment for VHL-RCC, VHL-pNET or VHL-CNS Hb, so estimates from studies in other oncology indications were used as proxies. In the scenario analysis, the caregiver disutility value was assumed to conservatively apply to one caregiver only, despite patients potentially requiring more than one caregiver. Caregiver disutility is modelled by health state and applied to all cohorts, and a caregiver bereavement disutility is applied as a one-time QALY decrement upon patient death. It is important to note that the caregiver disutility in

relatively severe health states before the patient's death is conceivably worse than the disutility associated with bereavement as while the patient is alive, the disutility includes both the disutility due to emotional distress/worry over the patient's condition, as well as the disutility due to the burden of providing care for the patient. In line with the approach taken for the estimate for utility in patients in the VHL-CNS Hb cohort who have PD (see *Utility in the non-metastatic health states* section), caring for a patient with severe neurological disability due to CNS Hb tumour growth is expected to be particularly burdensome. Therefore, a caregiver disutility is applied specifically for these patients using multiple sclerosis as a proxy. This disutility is taken from the NICE technology appraisal for ocrelizumab for treating relapsing–remitting multiple sclerosis [TA533] (83) taking the disutility associated with an Expanded Disability Status Scale (EDSS) score of 9. Caregiver disutility values are summarised below in Table 81.

Table 81 Summary of caregiver disutility values in the base case

Health state / Response status	Disutility	Source
Pre-surgery	-0.030	Turner et al. (2013) [breast, colorectal, or prostate cancer survivor] (84)
Surgery and event-free after surgery	-0.065	Based on the midpoint between caregiver disutility in the pre-surgery state versus the metastatic disease state.
Metastatic disease	-0.100	Sjolander et al. (2012) [lung or GI cancer]
After death of patient	-0.050	Song et al. (2012) [terminal cancer] (85)
PD patients in the VHL-CNS Hb cohort	-0.140	Ocrelizumab for treating relapsing–remitting multiple sclerosis [TA533] (caregiver disutility associated with EDSS score of 9) (83)

GI: Gastrointestinal, NSCLC: non-small cell lung cancer

B.3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR was conducted in July 2020, and subsequently updated in July 2022 to identify cost and resource use data for VHL-associated RCC, pNET, and CNS Hb tumours for use in the economic analysis. The SLR identified two relevant healthcare

resource use and cost publications (Jonasch et al, 2022 and Sundaram et al, 2022 (86, 87)). Full details of the SLR methods, identified studies and results are presented in Appendix I: Cost and healthcare resource identification, measurement and valuation section.

Intervention and comparators' costs and resource use

Belzutifan and SOC

The model considers the acquisition and administration costs associated with belzutifan. As SOC is defined as established clinical management without belzutifan, assumed to be surgery for the vast majority of patients, the model does not consider drug acquisition or administration costs for SOC. For costs associated with surgery and its complications see *Costs of surgery and complications*.

Table 82 presents the treatment acquisition and administration costs used in the economic analysis for belzutifan and SOC. The belzutifan acquisition cost at list price is £11,936.70 for a pack of 90 oral 40 mg tablets.

The dosing schedule of belzutifan is consistent with the treatment protocol used in the MK-6482-004 trial and the MHRA license. Three (40mg) tablets are taken daily to form a daily dose of 120 mg. In the base case, the mean relative dose intensity (RDI) observed in the MK-6482-004 trial ([redacted]%) is applied to the drug acquisition cost per 90-tablet pack of belzutifan to account for any delays or interruptions in treatment.

In the belzutifan arm, an oral drug dispensing cost of £245.23 once every 4 weeks (i.e., assuming a 4-week fill at each dispensing), sourced from NHS 2020/21 Reference Costs. No drug administration or acquisition costs are included for the SOC arm. Costs in this arm are analogous to the costs per health state presented in the *Health state unit costs and resource use* section.

Table 82 Belzutifan treatment acquisition and administration costs per 28-day cycle

Item	Cost	Source
Belzutifan 40 mg – 90 tablet pack (list price)	£11,936.70	BNF
Administration cost – oral drug dispensing (per pharmacy dispensing)	£10.80	Band 6 Hospital Pharmacist based on 12

Item	Cost	Source
		minutes of time for each dispensing, PSSRU 2021

Belzutifan is administered at a dose of 120 mg daily.
Mg: milligrams; NHS: National Health Service

Metastatic disease therapies

The model also considers the drug acquisition and administration cost associated with metastatic disease therapies (including both first-line and second-line options). Most patients who enter the metastatic disease state are assumed to receive an active first-line treatment for advanced RCC or pNET (depending on the origin of the tumour). Based on data from the VHL Natural History Study, where no cases of metastatic disease originated from CNS Hb, no patients receive metastatic therapy for CNS Hb. In addition, a subset of patients with advanced RCC or pNET are assumed to receive no active metastatic disease treatment, as not all patients with metastatic disease receive active treatment.

The first-line metastatic treatments considered in the economic analysis are based on the set of regimens recommended by NICE or listed as a preferred or recommended first-line regimen according to the National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines) for the treatment of advanced RCC and pNET. The approach to modelling metastatic therapies is in line with previous NICE appraisals, including pembrolizumab for adjuvant treatment of renal cell carcinoma (TA830) and pembrolizumab for adjuvant treatment of completely resected stage 3 melanoma (TA766) (53, 88).

The economic analysis also captures the costs associated with second- and later-line metastatic therapies in patients who progress on first-line metastatic therapies. For both first- and second-line therapies, the market share of metastatic therapies is assumed to be equal between the belzutifan and SOC arms.

Drug acquisition costs for first- and second- line therapies for advanced RCC and pNET are determined in the model as a function of the unit drug cost, PAS discount (if non-confidential), defined dosing schedule, and RDI. The pricing of sunitinib in the base case reflects the PAS discount for this therapy, in which the first treatment cycle of sunitinib is free to the NHS (89). The pricing of pazopanib reflects the price

subject to a 12.5% PAS discount (90). Otherwise, for treatments subject to a confidential PAS discount, the model uses list prices in the base case. Unit drug costs per vial or capsule are sourced from the British National Formulary (BNF) for branded agents and the electronic market information tool (eMIT) for generic agents (91, 92), and are presented in Table 83.

Table 83 Unit drug costs for first- and second-line therapies for advanced RCC and pNET

Drug	Strength per unit (mg or MU)	Cost per unit (£)
Pembrolizumab	100	2,630.00
Sunitinib 12.5 mg	12.5	28.03
Sunitinib 37.5 mg	37.5	84.00
Sunitinib 50 mg	50	112.10
Axitinib	5	62.80
Tivozanib	1.34	97.71
Pazopanib	400	37.37
Cabozantinib	60	171.43
Nivolumab	40	439.00
Ipilimumab	50	3,750.00
Avelumab	200	768.00
Lenvatinib	10	47.90
Everolimus 5 mg	5	75.00
Everolimus 10 mg	10	89.10
Temsirolimus	30	620.00
Interferon a2B	25	103.94
Cisplatin	50	6.03
Etoposide	100	3.84
Irinotecan	500	15.51
Leucovorin	350	5.50
Oxaliplatin	100	7.28
Streptozocin	1000	570.00
5-fluorouracil	2500	4.21
Doxorubicin	200	20.02
Temozolomide	180	3.47
Capecitabine	300	0.13
Lanreotide	120	937.00
Octreotide	30	656.88

Mg: milligrams; MU: million units; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma

For each first-line and second-line treatment of advanced RCC, the mean RDI (i.e., proportion of planned dose consumed) was applied to the drug acquisition cost to account for potential dose interruptions or reductions due to AEs or non-compliance.

Relative dose intensities for each of the first- and second-line therapies were obtained from pivotal clinical trials and HTA appraisals in advanced RCC settings.

For treatments of advanced pNET, dosing schedules were obtained from prescribing information, relevant trial publications, or clinical expert opinion. Based on data availability, the RDI of advanced pNET treatments was assumed to be 100%.

For IV drugs with weight-based or BSA-based dosing, the base-case analysis assumed that vial-sharing is allowed. Under this assumption, the number of vials required per infusion was calculated based on the mean body weight or BSA of patients in the MK-6482-004 trial population. For example, for a weight-based therapy, number of vials was calculated as average patient weight in kilograms multiplied by the required dose per kg (i.e., mg/kg) divided by the strength per vial (i.e., mg/vial, based on the vial strength associated with the lowest cost per mg).

A scenario analysis tested the alternative assumption where vial-sharing was not allowed. Under this scenario, the number of vials required per infusion was estimated based on a log-normal distribution of patient weight or BSA, using the mean and standard deviation reported for patients in the MK-6482-004 trial. This approach calculated the proportion of patients requiring different numbers of vials based on the estimated percentage of patients who fall into the corresponding weight or BSA interval. Recommended dosing schedules and RDI for advanced RCC treatments and advanced pNET treatments are shown in Table 84 and Table 85, respectively.

Table 84 Dosing schedules and RDI for first- and second-line therapies for advanced RCC

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
First-line therapies					
Sunitinib	Sunitinib	50 mg QD orally for 4 weeks, then 2 weeks off treatment	No max	74.7%	KEYNOTE-426 (data cut-off date: 24 Aug 2018) (93)
Tivozanib	Tivozanib	1.34 mg QD orally for 3 weeks followed	No max	94.0%	NICE TA512 (94)

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
		by 1 week without treatment			
Pazopanib	Pazopanib	800 mg QD orally	No max	86.0%	NICE TA215 (90)
Cabozantinib	Cabozantinib	20/40/60 mg QD orally	No max	94.3%	NICE TA542 (95)
Nivolumab / ipilimumab	Nivolumab (in combination)	3 mg/kg IV Q3W for up to 4 doses	12	94.8%	Assumed equal to that of pembrolizumab
	Ipilimumab	1 mg/kg IV Q3W for up to 4 doses	12	94.8%	
	Nivolumab (maintenance)	480 mg IV Q4W starting 6 weeks after the last combination dose	No max	94.8%	
Avelumab/ axitinib	Avelumab	800 mg Q2W	No max	91.5%	Motzer et al. (2019) [JAVELIN Renal 101] (96)
	Axitinib	5 mg BID orally	No max	89.4%	
Pembrolizumab / lenvatinib	Pembrolizumab	200 mg IV Q3W	104	94.8%	Assumed equal to that of pembrolizumab
	Lenvatinib	20 mg orally QD	No max	69.9%	Motzer et al. (2021) [KEYNOTE-581] (58)
Second-line therapies					
Nivolumab	Nivolumab	480 mg IV Q4W or 240 mg IV Q2W	No max	92.0%	NICE TA417 (97)
Pembrolizumab	Pembrolizumab	200 mg IV Q3W	No max	94.8%	Assume same as in first line
Axitinib	Axitinib	5 mg orally BID	No max	102.0%	NICE TA333/TA417 (97, 98)
Cabozantinib	Cabozantinib	60 mg orally QD	No max	100.0%	NICE TA463 (99)
Lenvatinib / everolimus	Lenvatinib	18 mg orally QD	No max	75.0%	Motzer et al. (2015) [NCT01136733] (100)
	Everolimus	5 mg orally QD	No max	85.0%	
Pazopanib	Pazopanib	800 mg orally QD	No max	86.0%	Assume same as in first line
Sunitinib	Sunitinib	50 mg orally QD for 4 weeks, then 2 weeks off treatment	No max	74.7%	Assume same as in first line

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
Everolimus	Everolimus	10 mg orally QD	No max	91.8%	NICE TA219/TA432 (101, 102)
Temsirolimus	Temsirolimus	25 mg IV QW	No max	92.4%	Hudes et al. (2007) [NCT00065468] (103)
Cytokines (interferon)	Interferon a2B	10 MU SC three days per week	No max	100.0%	Assumption

BID: twice a day; IV: intravenous; MU: million units; PAS: patient access scheme; Q#W: once every # weeks; QD: once a day, QW: once weekly; RCC: renal cell carcinoma; SC: subcutaneous; ToT: time on treatment

Table 85 Dosing schedules for first- and second- line therapies for advanced pNET

Regimen	Drug component	Dosing schedule description	Maximum ToT	Sources for dosing schedule
First-line treatments				
Streptozocin / 5-fluorouracil	Streptozocin	500 mg/m^2 IV on days 1 to 5 Q10W	No max	Sun et al. (2005) (59)
	5-fluorouracil	400 mg/m^2 IV on days 1 to 5 and days 36 to 40 Q10W	No max	
Streptozocin / doxorubicin	Streptozocin	500 mg/m^2 IV on days 1 to 5 Q10W	No max	Sun et al. (2005) (59)
	Doxorubicin	40 mg/m^2 IV Q5W	No max	
Temozolomide / capecitabine	Temozolomide	200 mg/m2 orally daily for 5 days Q4W	No max	Strosberg et al. (2011) (104)
	Capecitabine	750 mg/m2 orally twice daily for 14 days Q4W	No max	
Everolimus	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, Afinitor (everolimus)
Sunitinib	Sunitinib 12.5 mg	37.5 mg orally QD	No max	Prescribing information, Sutent (sunitinib)
Interferon a2B	Interferon a2B	5 MU SC three days per week	No max	Faiss et al. (2003) (105)
Lanreotide	Lanreotide	120 mg SC Q4W	No max	Prescribing information, Lanreotide
Octreotide	Octreotide	20 mg SC Q4W	No max	Clinical expert input
Second-line treatments				
Cisplatin / etoposide	Cisplatin	80 mg/m2 IV Q3W for up to 6 cycles	18	Iwasa et al. (2010) (106)

Regimen	Drug component	Dosing schedule description	Maximum ToT	Sources for dosing schedule
	Etoposide	100 mg/m2 IV on days 1-3 Q3W for up to 6 cycles	18	
Everolimus	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, Afinitor (everolimus)
FOLFIRI	5-fluorouracil	400 mg/m2 IV (in a 10-min bolus) + 1,200 mg/m2 (in a 44-h infusion) Q2W	No max	Hentic et al. (2012) (107)
	Irinotecan	180 mg/m2 (on day 1) Q2W	No max	
	Leucovorin	400 mg/m2 (in a 2-h infusion) Q2W	No max	
FOLFOX	5-fluorouracil	400 mg/m2 (bolus) + 2,400 mg/m2 (as a 46-h continuous infusion) Q2W	No max	Faure et al. (2017) (108)
	Oxaliplatin	85 mg/m2 IV infusion (over 120 minutes) Q2W	No max	
	Leucovorin	100 mg/m2 IV infusion (over 120 minutes on day 1) Q2W	No max	
Streptozocin / 5-fluorouracil	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Sun et al. (2005) (59)
	5-fluorouracil	400 mg/m2 IV on days 1 to 5 and days 36 to 40 Q10W	No max	
Streptozocin / doxorubicin	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Sun et al. (2005) (59)
	Doxorubicin	40 mg/m2 IV Q5W	No max	
Sunitinib	Sunitinib 12.5 mg	37.5 mg orally QD	No max	Prescribing information, Sutent (sunitinib)
Temozolomide / capecitabine	Temozolomide	200 mg/m2 orally daily for 5 days Q4W	No max	Strosberg et al. (2011) (104)
	Capecitabine	750 mg/m2 orally twice daily for 14 days Q4W	No max	
Interferon a2B	Interferon a2B	5 MU SC three days per week	No max	Faiss et al. (2003) (105)
Lanreotide / everolimus	Lanreotide	120 mg SC Q4W	No max	Prescribing information, Lanreotide
	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, Afinitor (everolimus)
	Octreotide	20 mg SC Q4W	No max	Clinical expert input

Regimen	Drug component	Dosing schedule description	Maximum ToT	Sources for dosing schedule
Octreotide / everolimus	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, Afinitor (everolimus)
Lanreotide	Lanreotide	120 mg SC Q4W	No max	Prescribing information, Lanreotide
Octreotide	Octreotide	20 mg SC Q4W	No max	Clinical expert input

BID: twice a day; IV: intravenous; MU: million units; PAS: patient access scheme; pNET: pancreatic neuroendocrine tumour; Q#W: once every # weeks; QD: once a day, QW: once weekly; SC: subcutaneous; ToT: time on treatment

Unit costs of IV and oral drug administration were based on the 2019/2020 NHS Reference Costs (see Table 86). For each treatment in the advanced RCC and pNET settings, unit costs of drug administration were applied based on the following principles:

- For IV- and SC-administered single-agent regimens, the unit cost per infusion was based on SB12Z (simple parenteral chemotherapy).
- For IV-administered combination regimens that either do not contain cisplatin or do not require multiple infusions per chemotherapy cycle, the unit cost per chemotherapy cycle (covered all drug components) was based on SB13Z (complex parenteral chemotherapy).
- For IV-administered combination regimens in which one or more drug components is administered more than once per chemotherapy cycle, the unit cost per chemotherapy cycle was based on the sum of SB13Z (complex parenteral chemotherapy) and SB15Z (subsequent elements of a chemotherapy cycle). For the combination of cisplatin with etoposide, the unit cost per chemotherapy cycle was based on the sum of SB14Z (complex parenteral chemotherapy, including prolonged infusion) and SB15Z (subsequent elements of a chemotherapy cycle), given the prolonged 6- to 8-hour infusion time required for cisplatin and the multiple days of administration required for etoposide per 3-week chemotherapy cycle.

- Orally administered single-agent or combination regimens were assumed to require one oral drug dispensing cost based on SB11Z (deliver exclusively oral chemotherapy) once every 4 weeks (or once every 6 weeks for sunitinib in the advanced RCC setting). (For combination regimens that include both orally administered and IV-administered drug components, the administration cost associated with the oral component was assumed to be covered by the IV administration costs.)

Table 86 Unit costs of drug administration in the advanced RCC setting

Route	Type of administration	Unit cost per administration or pharmacy dispensing (£)	Source
IV or SC	Simple parenteral chemotherapy	361.53	NHS Refe–ence Costs 2020/2021 - SB12Z (deliver simple parenteral chemotherapy at first attendance) (109)
IV	Complex parenteral chemotherapy	426.80	NHS Refe–ence Costs 2020/2021 - SB13Z (deliver more complex parenteral chemotherapy at first attendance)
IV	Complex parenteral chemotherapy with prolonged infusion	526.52	NHS Refe–ence Costs 2020/2021 - SB13Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance)
IV	Complex parenteral chemotherapy with subsequent infusion(s)	897.42	NHS Refe–ence Costs 2020/2021 - SB13Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance) + SB15Z (deliver subsequent elements of a chemotherapy cycle)
IV	Complex parenteral chemotherapy with prolonged infusion and subsequent infusion(s)	997.14	NHS Refe–ence Costs 2020/2021 - SB14Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance) + SB15Z (deliver subsequent elements of a chemotherapy cycle)
Oral	Oral drug dispensing	245.23	NHS Refe–ence Costs 2020/2021 - SB11Z (deliver exclusively oral chemotherapy)

IV: intravenous; NHS: national health service; RCC: renal cell carcinoma; SC: subcutaneous

Discontinuation rates for first-line metastatic treatments for advanced RCC and advanced pNET are approximated from exponential rates of PFS failure (as described in the *Transitions from metastatic disease to death* section). Some regimens are also subject to a maximum treatment duration according to dosing

schedules recommended by NICE (Table 84 and Table 85). Median ToT for each subsequent-line metastatic treatment of RCC is sourced from relevant second-line clinical trials conducted in advanced RCC populations, or for interferon a2B extracted from a first-line trial in the absence of second-line setting clinical data. For all second-line treatments of metastatic pNET, median ToT was assumed to be 4 months (17.4) weeks, based on median PFS reported by Hentic et al. (2012) in a clinical trial of folic acid, fluorouracil, and irinotecan (FOLFIRI) as second-line treatment of pNET (107). Mean ToT for each subsequent therapy is calculated as a function of median ToT, based on an assumption of constant hazards. The estimated discontinuation rate and (where applicable) the maximum duration of each drug component in a regimen is used to estimate the mean total cost of each treatment regimen in the first- and second-line setting. ToT for each second-line treatment regimen of advanced RCC is presented below in Table 87.

Table 87 ToT for second-line treatment regimens in the advanced RCC setting

Second-line treatment regimen	Component	ToT (months)		Source
		Median	*Mean*	
Nivolumab	Nivolumab	23.9	34.5	Motzer et al. (2015) [CheckMate 025] (110)
Axitinib	Axitinib	35.7	51.4	Motzer et al. (2013) [AXIS] (111)
Cabozantinib	Cabozantinib	36.5	52.7	Motzer et al. (2018) [METEOR] (112)
Lenvatinib/everolimus	Lenvatinib	33.0	47.7	Motzer et al. (2015) [NCT01136733]
	Everolimus	33.0	47.7	
Pazopanib	Pazopanib	32.2	46.4	Sternberg et al. (2013) [VEG105192] (113)
Sunitinib	Sunitinib	32.2	46.4	Assume same median ToT as pazopanib
Everolimus	Everolimus	19.1	27.6	Motzer et al. (2018) [METEOR] (112)
Temsirolimus	Temsirolimus	19.1	27.6	Hutson et al. (2014) [INTORSECT] (114)
Cytokines (interferon)	Interferon a2B	12.0	17.3	Rini et al. (2008) [CALGB 90206] (115)

RCC: renal cell carcinoma; ToT: time on treatment

The mean cost of first- and second-line treatments is calculated for each treatment arm as a weighted average based on the first- and second-line market shares within

each treatment arm. A summary of the metastatic therapies, market shares, and total costs applied in the model is provided in Table 88.

Table 88 Metastatic treatment market shares and costs

Metastatic therapy regimens	Market share by treatment arm		Total cost of regimen (£)	
	Belzutifan	SOC	Acquisition	Administration
First-line metastatic therapy (metastatic RCC)				
Sunitinib	30.0%	30.0%	24,866	2,846
Tivozanib	14.0%	14.0%	28,217	3,587
Pazopanib	29.0%	29.0%	26,106	4,066
Cabozantinib	13.0%	13.0%	164,162	8,894
Nivolumab / ipilimumab	14.0%	14.0%	110,894	7,420
Avelumab / axitinib	0.0%	0.0%	221,156	18,242
No active treatment	0.0%	0.0%	0	0
Second-line metastatic therapy (metastatic RCC)				
Nivolumab	0.0%	0.0%	41,804	3,118
Axitinib	7.0%	7.0%	46,133	3,154
Cabozantinib	32.0%	32.0%	63,235	3,231
Lenvatinib / everolimus	0.0%	0.0%	42,856	2,923
Pazopanib	4.0%	4.0%	18,274	2,846
Sunitinib	0.0%	0.0%	15,796	1,897
Tivozanib	0.0%	0.0%	22,083	2,807
Everolimus	7.0%	7.0%	15,804	1,692
Sorafenib	0.0%	0.0%	19,385	1,385
Cytokines (interferon)	0.0%	0.0%	2,159	6,259
No active treatment	50.0%	50.0%	0	0
First-line metastatic therapy (metastatic pNET)				
Streptozocin / 5-fluorouracil	0.0%	0.0%	9,044	2,984
Streptozocin / doxorubicin	0.0%	0.0%	9,052	2,984
Temozolomide / capecitabine	0.0%	0.0%	448	2,038
Everolimus	0.0%	0.0%	64,837	6,373
Sunitinib	0.0%	0.0%	48,629	5,311
Interferon a2B	0.0%	0.0%	6,133	35,551
Lanreotide	50.0%	50.0%	18,528	7,149
Octreotide	50.0%	50.0%	12,989	7,149
No active treatment	0.0%	0.0%	0	0
Second-line metastatic therapy (metastatic pNET)				
Cisplatin / etoposide	0.0%	0.0%	172	4,270
Everolimus	25.0%	25.0%	15,650	1,538
FOLFIRI	0.0%	0.0%	347	11,259
FOLFOX	0.0%	0.0%	297	11,259
Streptozocin / 5-fluorouracil	25.0%	25.0%	6,826	2,252
Streptozocin / doxorubicin	25.0%	25.0%	6,832	2,252
Sunitinib	25.0%	25.0%	12,414	1,538
Temozolomide / capecitabine	0.0%	0.0%	338	1,538
Interferon a2B	0.0%	0.0%	1,565	9,072
Lanreotide / everolimus	0.0%	0.0%	21,528	2,268
Octreotide / everolimus	0.0%	0.0%	19,771	2,268
Lanreotide	0.0%	0.0%	5,878	2,268

Octreotide	0.0%	0.0%	4,121	2,268
No active treatment	0.0%	0.0%	0	0

FOLFIRI: folic acid, fluorouracil, and irinotecan; FOLFOX: folic acid, fluorouracil, and oxaliplatin; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; SOC: standard of care

Health state unit costs and resource use

Healthcare resource use costs are accrued by health state, as described in *Treatment decision point*.

Health state unit costs in the pre-surgery, surgery, event-free after surgery, and metastatic disease states are sourced from NHS Reference Costs 2020/21 and the PSSRU Unit Costs of Health and Social Care 2021. Due to the severity of disease in this target population, the highest complication and comorbidity (CC) score is used for costs where applicable. Additionally, where appropriate, social care costs have been included for patients with sequalae/complications that require social care adaptations (see *Miscellaneous unit costs and resource use* section below).

In the base case, health state costs for non-metastatic health states include costs for outpatient visits, laboratory tests and radiologic exams. In the metastatic disease health state, costs include those for salvage surgery, outpatient visits, laboratory tests and radiologic exams. As the metastatic state encompasses both pre- and post-progression metastases, health state costs are calculated as a weighted average of resource use associated with pre- versus post-progression metastases based on the proportion of time spent progression-free within the metastatic disease state.

Resource use costs were sourced from relevant literature sources and clinical surveillance guidelines in each tumour-specific cohort. Resource use frequencies in the pre-surgery and event-free after surgery states were sourced from Maher et al. (2011), and Kanno et al. (2014) in the event-free after surgery state for the CNS Hb cohort (116, 117). Resource use in the metastatic disease state was sourced from KEYNOTE-564 (data cut-off date: 14 Jun 2021), and the previous NICE appraisal TA542 (cabozantinib in the untreated locally advanced or metastatic RCC setting) for RCC origin tumours, and TA476 (nab-paclitaxel with gemcitabine for untreated metastatic pancreatic cancer) for pNET origin tumours (95, 118). Additionally, for a

proportion of patients with metastatic RCC, one-time costs of salvage surgery were assumed to be incurred upon entering the metastatic disease state, based on the observed percentage of patients with surgery among those who experienced distant metastases as their first DFS failure type in the KEYNOTE-564 trial.

Patients who transition to death incur a one-time cost associated with palliative/terminal care of £7,220.05 (inflation-adjusted using the health component of the Consumer Price Index from the Office of National Statistics), based on costs during the last 90 days before death as reported by Georghiou & Bardsley (2014) (119). This source was accepted by NICE to inform terminal care costs in the NICE submission for cabozantinib in the untreated locally advanced or metastatic RCC setting (TA542) (95).

A summary of health state unit costs, units per month, and percentage of patients utilising the resource is provided in Table 89.

Table 89 Health state costs applied in the base case

Health states	Items	Unit cost (£)	% patients	Resource use per month	Cost source
Pre-surgery and event-free after surgery states*	General practitioner visit	39.00	100%	0.08	Gener–l pract–tioner costs - PSSRU - Unit Costs of Health and Social Care 2021
	Ophthalmologist visit	1–6.35	100%	0.08	WF01– - Service Code– 130 - Ophthalmology - Non-Admitted Face-to-Face – ttendance, Follow-up - NHS Reference Costs 2020/21
	Complete blood count test	–.63	100%	0.08–DAPS05 - Haematology - Directly access–d pathology services - NHS Reference Costs 2020/21	
	Urinalysis	–.85	100%	0.08	DAPS04 - –linical Biochemistry - Directly access–d pathology services - NHS Reference Costs 2020/21
	CT scan of abdomen/pelvis	133.80	71%	0.08	Weighted average of total HRG activity for RD–0A, RD21A, and RD22Z - NHS Reference Costs 2020/21
	MRI of brain *(in RCC and pNET cohort)*	230.62	100%	0.04	Weighted average of total HRG activity for RD–1A,

					RD02A, and RD03Z - NHS Reference Costs 2020/21
	MRI of brain *(in CNS Hb cohort)*	230.62	100%	0.08	Weighted average of total HRG activity for RD–1A, RD02A, and RD03Z - NHS Reference Costs 2020/21
	Ultrasound	230.62	58%	0.08	Weighted average of total HRG activity for RD–1A, RD02A, and RD03Z - NHS Reference Costs 2020/21
Metastatic disease state (RCC origin tumour)**	Salvage surgery	7,850.92	21%	1.00 (one-time upon entry only)	NHS. Robot-assisted nephrectomy: Evidence summary report (2014), inflation-adjusted to 2021 GBP
	Medical oncologist visit	2–4.55	100%	1.00	WF01– - Service Code: 370–- Medical Oncologist - Total O–tpatient Attendances - NHS Reference Costs 2020/21
	Complete blood count test	–.63	100%	1.00–DAPS05 - Haematology - Directly access–d pathology services - NHS Reference Costs 2020/21	
	CT scan of abdomen/pelvis	133.80	100%	1.00 (one-time upon entry) 0.33 thereafter	Weighted average of total HRG activity for RD–0A, RD21A, and RD22Z - NHS Reference Costs 2020/21
Metastatic disease state (pNET origin tumour)**	Medical oncologist visit	2–4.55	100%	1.00	WF01– - Service Code: 370–- Medical Oncologist - Total O–tpatient Attendances - NHS Reference Costs 2020/21
	Cancer specialist nurse–90.49	50%	1.00	N10AF - Specialist Nursing, Cancer Related– Adult, Face to face - NHS Reference Costs 2020/21	
	Complete blood count test	–.63	100%	6.00–DAPS05 - Haematology - Directly access–d pathology services - NHS Reference Costs 2020/21	
	CT scan of abdomen/pelvis	133.80	100%	1.00 (one-time upon entry) 0.33 thereafter	Weighted average of total HRG activity for RD–0A, RD21A, and RD22Z - NHS Reference Costs 2020/21
	MRI of abdomen/pelvis	230.62	10%	1.00 (one-time upon entry)	Weighted average of total HRG activity for RD–1A, RD02A, and RD03Z - NHS Reference Costs 2020/21
	Ultrasound	230.62	5%	1.00 (one-time upon entry)	Weighted average of total HRG activity for RD–1A, RD02A, and RD03Z - NHS Reference Costs 2020/21

Death	Terminal care	£7,220.05	100%	1.00 (one-time upon death)	Theo Georghiou and Martin Bardsley. Exploring the cost of care at the end of life. September 2014. Nuffield Trust. With inflation-adjustment to 2021 GBP.

*Pre-surgery and event-free after surgery resource use is assumed to be the same for all cohorts in the model with the exception of MRI scan of brain
**Unless stated, for all metastatic states (i.e., pre- and post-progression), resource use is assumed to be the same for all cohorts in the model
Notes: Frequencies of salvage surgery are based on observed percentages of patients with surgery among those who experienced distant metastases as their first DFS failure type in KEYNOTE-564. For the metastatic disease state in the pNET cohort, the 6 complete blood count tests include a total of 5 liver function tests and one blood test.
CT: Computed Tomography; DFS: Disease-Free Survival; GBP: Great British Pound; MRI: Magnetic Resonance Imaging; NHS: National Health Service; pNET: pancreatic neuroendocrine tumour; PSSRU: Personal Social Services Research Unit; RCC: renal cell carcinoma

Cost of surgery and complications

The model considers costs of surgery related to the primary tumour (i.e. RCC, pNET, CNS Hb) in addition to surgery not related to the primary tumour. The costs incurred by surgical complications is also accounted for. The model assumes that the risks of surgical complications are equal between the belzutifan and SOC arms of the model. As detailed in *Surgery* above, the model calculated the occurrence of tumour reduction surgeries for primary and non-primary VHL-related tumours and associated short- and long-term complications over time. Short-term of surgical complications costs are applied as a one-off cost per surgery. Annual costs of long-term surgical complications were converted into weekly costs, which were applied as a recurring per-cycle cost based on the cumulative proportion of patients who have developed each long-term complication and are still alive in each cycle.

The unit costs of all surgeries are shown in Table 90. The costs associated with each surgical complication is reported in Table 91 and Table 92.

Table 90 Unit costs per surgical procedure

Surgery	Unit cost per procedure (£)	Cost source
Surgical procedure for VHL-associated RCC	7,850.92	Solutions for Public Health (on behalf of NHS England) (2014), inflation-adjusted to 2021 GBP

Surgery	Unit cost per procedure (£)	Cost source
Surgical procedure for VHL-associated CNS Hb	20,573.29	AA82Z – Total HRGs, Intracranial Telemetry, with Cortical Mapping or Resection of Brain NHS Reference Costs 2020/21
Surgical procedure for VHL-associated–pNET	23,922.25	GA03C - Total HRGs, Very Complex, Hepatobiliary or Pancreatic Procedures, with CC Score 4+ - NHS Reference Costs 2020/21
Adrenal lesion surgery	10,369.44	KA04 – Total HRGs, Adrenal Procedures with CC Score 2+ - NHS reference costs 2020/21
Endolymphatic sac tumour surgery	5,029.35	Weighted average CB02A, CB02B–and CB02C Total HRGs - Non-Malignant, Ear, Nose, Mouth, Throat or Neck Disorder–, with Interventions - NHS reference costs 2020/21
Epididymal cystadenoma surgery	6,609.57	Weighted av–rage LB35C and LB35D - Scrotum, Testis or Vas Deferens Disorder–, with Interventions - NHS reference costs 2020/21
Retinal Hb surgery	3,970.02	Weighted average BZ80A, B–80B, BZ81A and BZ81B - Complex/Very Complex Vitreous Retinal Procedur–s, 19 years and over - NHS reference costs 2020/21

CNS: Central Nervous System; GBP: Great British pounds; Hb: Haemangioblastoma; NHS: national health service; pNET: Pancreatic Neuroendocrine Tumour; RCC: Renal Cell Carcinoma; VHL: Von Hippel Lindau

Table 91 Costs of short-term surgical complications

Complication	Cost per complication* (£)	Cost source
Short-term complications of surgical procedures for VHL-RCC		
Acute renal failure	7,534.29	NHS R–ference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21
Cardiac complications	3,685.32	NHS Reference Costs (Weighted costs of EB03A, EB05A, E–10A, EB13A and EB14A - Total HRGs) 2020/21
Erythroderma	8,559.58	NHS Reference Costs (JD07A Skin Disorders with Interventions, with CC Score 12+ - Total HRGs) 2020/21
Kidney infection	7,612.90	NHS Reference Costs (LA04H Kidney or Urinary Tract Infections, with Interventions, with CC Score 12+ - Total HRGs) 2020/21
Other genitourinary complications	1,375.85	NHS Reference Costs (LA09J General Renal Disorders with Interventions, with CC Score 6+ - Total HRGs) 2020/21

Complication	Cost per complication* (£)	Cost source
Postoperative infection (RCC-related)	13,139.64	NHS Reference Costs (WH07A Infections or Other Complications of Procedures, with Multiple Interventions, with CC Score 2+ - Total HRGs) 2020/21
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Short-term complications of surgical procedures for VHL-CNS Hb		
Acute renal failure	7,534.29	NHS R–ference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21
CNS hemorrhage	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21
Nerve palsy related to anesthesia	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Short-term complications of surgical procedures for VHL-pNET		
Abdominal abscess	10,881.28	NHS Reference Costs (FD10A Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions, with CC Score 8+ - Total HRGs) 2020/21
Postoperative infection (pNET-related)	13,139.64	NHS Reference Costs (WH07A Infections or Other Complications of Procedures, with Multiple Interventions, with CC Score 2+ - Total HRGs) 2020/21
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Urinary tract infection	1,715.45	NHS –eference Costs (LA04 - Total HRGs) 2019-2020

Complication	Cost per complication* (£)	Cost source
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Short-term complications of adrenal lesion surgery		
Acute renal failure	7,534.29	NHS R–ference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21
Thrombosis or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21
Short-term complications of retinal Hb surgery		
Vitreous haem–rrhage	1,009.15	BZ24 - Non-s–rgical ophthalmology - NHS reference costs 2020/21
Complications of endolymphatic sac tumour surgery		
Acoustic im–airment	956.29	CB02 - Non-Malignant, Ear, Nose, Mouth, Thr–at or Neck Disorders - NHS reference costs 2020/21

*Costs are applied as one-time costs for patients undergoing surgery in each cycle for short-term/acute complications
CNS: Central Nervous System; Hb: Haemangioblastoma; HRG: Healthcare Resource Group; N/A: not applicable; NHS: National Health Service; pNET: Pancreatic Neuroendocrine Tumour; RCC: Renal Cell Carcinoma; VHL: Von Hippel Lindau

Table 92 Annual costs of long-term surgical complications

Complication	Cost per complication* (£)	Cost source
Annual cost of long-term complications for VHL-RCC surgeries		
End stage renal disease and/or dialysis**	30,477.27	Kerr et al. (2012) (estimated expenditure on dialysis per patient requiring dialysis), inflation-adjusted to 2021 GBP (120)
Chronic kidney disease**	1,034.32	Kerr et al. (2012) (overall annual cost of CKD per patient diagnosed with CKD), inflation-adjusted to 2021 GBP (120)
Hernia surgery	2,021.79	Coronini-Cronberg et al. (2013), inflation-adjusted to 2021 GBP (121)
Chronic pain	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21
Cerebral vasculature occlusion or stroke†	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21
Annual cost of long-term complications for VHL-CNS Hb surgeries		

Complication	Cost per complication* (£)	Cost source
Chronic pain (CNS Hb)	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21
Cerebral vasculature occlusion or stroke†	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21
Seizure	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21
Neurological complications†	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21
Annual cost of long-term complications for VHL-pNET surgeries		
Chronic pain (pNET)	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21
Secondary diabetes or exocrine pancreatic insufficiency**	9,681.57	NHS Reference Costs (GC17A Non-Malignant, Hepatobiliary or Pancreatic Disorders, with Multiple Interventions, with CC Score 9+ - Total HRGs) 2020/21
Immunocompromisation**	794.74	NHS Reference Costs (WJ11Z Other Disorders Of Immunity- Total HRGs) 2020/21
Annual cost of long-term complications for adrenal lesion surgeries		
Adrenal insuf–iciency	876.28	KA08 - Othe– Endocrine Disorders - NHS reference costs 2020/21
Chro–ic pain	649.85	–D05 - Abdominal Pain - NHS reference costs 2020/21
Annual cost of long-term complications for Renal Hb surgeries		
Chroni– pain	1,009.15	BZ24 - Non-s–rgical ophthalmology - NHS reference costs 2020/21

*Costs are applied annually (recurring) from the time of surgery until death or the end of the modelled time horizon for long-term complications.
**This is a metabolic complication resulting from limited/absent organ function following last-resort surgery
†Social care costs are added to this unit cost in the cost-effectiveness model.
CNS: Central Nervous System; Hb: Haemangioblastoma; HRG: Healthcare Resource Group; NHS: National Health Service; pNET: Pancreatic Neuroendocrine Tumour; RCC: Renal Cell Carcinoma

Adverse reaction unit costs and resource use

The model considers the costs and resource use associated with Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan in the MK-6482-004 trial as well as all Grade ≥3 treatment-related AEs associated with belzutifan (see Adverse

reactions section). These AEs are expected to incur healthcare resource use, costs, and a utility decrement.

Risks of the included AEs for patients treated with belzutifan were sourced from the MK-6482-004 trial, based on the proportions of patients with AEs reported for the all-subjects-as-treated population. In the SOC arm, it is assumed that the risk of all AEs is zero. Therefore, the model uses risks of drug-related (rather than all-cause) grade 3 to 5 AEs for the belzutifan arm to approximate the incremental AE risk associated with belzutifan versus SOC. As described in the Adverse reactions section, the durations of AEs were also sourced from the MK-6482-004 trial.

Unit costs per AE were obtained from NHS 2020/21 reference costs and weighted by the risk of each AE among patients in the belzutifan arm. Costs of AE management are applied as one-time costs in the first cycle. A summary of costs associated with AEs is presented in Table 93. AE costs are only applied to the belzutifan arm. It is assumed that the SOC arm is associated with no AE costs as patients are not on active treatment and all costs would be captured in the health state cost.

Table 93 List of adverse reactions and summary of costs in the economic model

AEs	AE risk (%)	AE cost (£)	Source
Total cost of AEs	**N/A**	**46.62**	**Weighted average of rate and cost of individual AEs**
Anaemia	[redacted]	356.39	NHS Reference Cost 2020/21, WH13: Abnormal Findi–gs without Diagnosis - Regular Day or Night Admissions (weighted average)
Fatigue	[redacted]	116.45	NHS Reference Cost 2020/21, WH17: Admission Rela–ed to Social Factors - Regular Day or Night Admissions (weighted average)

AE: adverse event; N/A: not applicable

Miscellaneous unit costs and resource use

As detailed in *B.1.3 Health condition and position of the technology in the treatment pathway*, VHL-associated manifestations cause significant patient burden beyond surgical procedures and hospital admissions. Social care is required particularly for patients with neurological disability resulting from CNS Hb and from debilitating surgical complications. These patients are unable to perform standard activities of

daily living and therefore require social care support for everyday needs. The NICE reference case (63) stipulates costs should be considered from an NHS and PSS perspective. Social care costs associated with stroke and neurological dysfunction as a complication of surgery associated with VHL have therefore been included in the base-case analysis. Additionally, for PD patients in the VHL-CNS Hb cohort, social care costs associated with disease management have also been included to reflect the social care required for this patient population who experience significant morbidity. Motor neurone disease is again used as a proxy health condition (in line with utility and caregiver disutility estimation approaches). The social care costs and sources used are reported in Table 94.

Table 94 Social care costs included in this appraisal

Complication/Patient population	Annual cost (£)	Source & estimation method
Stroke	£3,232.45	Estimated societal costs of stroke in the UK based on a discrete event simulation (Patel et al., 2020) (122). Course of community rehabilitation reported in the supplementary material.
Neurological complications	£849.11	The size, burden and cost of disorders of the brain in the UK (Fineberg et al., 2013) (123). Proportion of estimated total UK per-subject cost of brain disorders attributed to direct non-medical costs converted to GBP and inflated to 2021.
Disease management of PD patients in VHL-CNS Hb cohort	£1,085.31	Health Utilities and Costs for Motor Neurone Disease (Moore et al., 2019) (124). Community services cost over a 3-month period reported in the supplementary material inflated to 2021.

B.3.6 Severity

VHL is a severe disease and its associated RCC, CNS Hb, and pNET tumours are associated with significant morbidity and mortality. VHL-associated tumours can cause damage to affected organs, such as end-stage renal failure in patients with RCC tumours. If left untreated, VHL-associated RCC and pNET tumours can metastasise throughout the body, ultimately leading to premature death (16). Patients with VHL-associated tumours also suffer a considerable psychological burden from the negative effects of the disease on HRQoL (29), and the need for constant monitoring and surgery to treat tumours. The psychological impact of VHL is highlighted by the two deaths which occurred in the MK-6482-004 trial (which were found to be unrelated to belzutifan treatment): one death was due to suicide (in a

patient who had 9 prior surgeries) and the other due to a fentanyl overdose (in a patient who had 17 prior surgeries) (37).

Until belzutifan was developed, no pharmacological interventions were available specifically for the treatment of VHL, with treatment options limited only to routine surveillance, surgery, and subsequent therapies once tumours progressed to metastatic disease. There is a significant unmet need for options to treat VHL. The Company investigated whether VHL-associated RCC, CNS Hb, and pNET qualifies for the severity modifier, as per the NICE 2022 Manual (63).

Under the severity modifier, the NICE committee may apply a greater weight to QALY gains if the technology is indicated for a condition with a high degree of severity (as determined based on proportional and absolute QALY shortfall). To understand the extent to which VHL deprives patients of their remaining QALYs, the total lifetime accrued QALY of patients receiving SOC (as estimated in the cost-effectiveness model) is measured as a proportion of the total lifetime QALY gain of healthy patients of the same age and sex distribution (63). The total QALYs associated with SOC were obtained from the results of the base-case analysis, and the estimated total QALYs for the general population reflected the baseline characteristics of the MK-6482-004 trial (see Table 95) as calculated within the economic model.

The results of the QALY shortfall analysis (see Table 96) indicate that VHL-associated RCC and CNS Hb cohorts both qualify for the 1.7 severity modifier and VHL-associated pNET cohort qualifies for the 1.2 severity modifier. As depicted in Figure 9 in Section B.2.7, however, in the real world these three VHL cohorts are not actually distinct cohorts and therefore the appropriate severity weighting is for the full GB-indicated population, not based on primary tumour. No patient in the MK-6482-004 trial had only one tumour manifestation. Given what is known about the significant burden of VHL-pNET on mortality and morbidity and the severe complications arising from surgeries of the pancreas, there is little doubt that VHL-pNET meets the definition of highly severe that the severity modifier has been designed to identify. Therefore, the cost-effectiveness results for all VHL cohorts

assessed in the current analysis should be all be weighted using the 1.7 QALY weight and this weighting has been incorporated in to the ICER calculations.

Table 95 Summary features of QALY shortfall analysis

Factor	Value (reference to appropriate table or figure in submission)	Reference to section in submission
Sex distribution	47.5% female	Patient population
Starting age	41.0 years	Patient population

QALY: quality adjusted life year

Table 96 Summary of QALY shortfall analysis

Cohort	Expected total QALYs for the general population	Total expected QALYs for people with VHL on current SOC	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
VHL- GB marketing authorisation population (weighted cohort)	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated RCC	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated CNS Hb	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated pNET	18.15	[redacted]	[redacted]	[redacted]	1.2

QALY: quality adjusted life year; SOC: standard of care; VHL: Von Hippel Lindau

The results of the QALY shortfall analysis are unsurprising given that the modelled outcomes are associated with patients who have run out of effective treatment options and have experienced loss of organ function, with deleterious impacts on HRQoL. The assessment of severity of VHL is further supported by the study by Wilding et al. (2012), described in greater detail in Error! Reference source not found., which reported that VHL disease reduces median life expectancy by nearly 19 years in men and by nearly 34 years in women. These estimates were based on data from the UK in a VHL cohort far broader than the population indicated per the MHRA label (i.e. not limited to patients requiring therapy and for whom localised procedures are unsuitable or undesirable), therefore it is highly likely that the foregone years of life in the assessed population would be far greater than these estimates. Given the known complications facing patients at the decision point evaluated in this appraisal as per the MHRA licensed indication, the results from this

real-world study leaves little doubt that the VHL population evaluated in this appraisal meets NICE's definition of the highest tier of severity, and the highest severity modifier should be taken into account when assessing the cost-effectiveness results in this appraisal. These severity assessments confirm the burden of this disease.

B.3.7 Uncertainty

Data gaps are common to any health technology assessment of treatments for very rare diseases. VHL is a very rare disease with a highly heterogeneous presentation, compounding the challenge of conducting an HTA.

What is not uncertain, is that belzutifan offers clinically meaningful benefits in VHL patients in terms of tumour shrinking and reduced risk of surgery. This is evidenced in MK-6482-004 by the ORRs (in RCC and also in pNET and CNS Hb), with most patients showing a degree of tumour shrinkage and very few patients experiencing disease progression. Belzutifan offers a step-change in the management of VHL, a lifelong condition with no effective therapies and high unmet need.

The challenge for this HTA is valuing this substantial clinical benefit given the data package available and the constraints in modelling a highly heterogenous population based on a trial with only 61 participants and no control arm. An additional challenge is the discrepancy between the trial patient population and the population defined in the MHRA marketing authorisation. The licensed population is more restricted – requiring therapy and not suitable or desirable for local procedures – and also broader – three primary tumours, not just RCC.

Real world studies were conducted to address known gaps in the trial data (46). However, these studies were commissioned prior to final GB-MA and so are not fully aligned with the GB-eligible patient population.

A further challenge is the low number of events post-belzutifan initiation. Few surgery events were observed in the study period (VHL-RCC: n=7; VHL-CNS: Hb n=2; VHL-pNET: n=0, as of 01 April 2022 data-cut) (37). While this low event incidence highlights the efficacy of belzutifan in shrinking tumours, controlling

disease, and preventing surgery compared to the pre-treatment period, it makes predicting life-time risk of surgery or metastases difficult. Small datasets and low event rates are common limitations of modelling in rare diseases such as VHL.

A number of simplifying assumptions have been used in the economic model that do not fully value the benefit of belzutifan: the true value of preventing blindness as a 'side-effect' of belzutifan treatment, for example, has not been included. The model isolates tumours, rather than presenting the whole system and whole-system tumour reduction, significantly underestimating the value of belzutifan. We note an absence of data on some of the worst outcomes associated with current SOC: suicidal ideation and completed suicides are frequently referenced by clinicians but we were not able to identify any data on rates in the eligible patient population.

To mitigate these challenges key assumptions and model inputs were extensively discussed and validated with UK clinical experts to ensure the modelled outcomes were generalisable to UK clinical practice and the licenced indication. Substantial numbers of sensitivity and scenario analyses have been undertaken.

Acknowledging limitations in the economic model and gaps in the ideal data package there is no uncertainty that this treatment is a good use of NHS resources in the treatment of this rare diseases across a range of plausible scenarios.

B.3.8 Managed access proposal

No managed access scheme has been proposed for belzutifan in its licensed indication.

B.3.9 Summary of base-case analysis inputs and assumptions

Summary of base-case analysis inputs

In line with the NICE reference case, the analysis was conducted from the NHS and PSS perspective using a lifetime horizon (59 years) and with costs and QALYs discounted at 3.5%. A summary of inputs used in the base-case is presented in Appendix J1.2 Summary of base-case analysis inputs.

Assumptions

This section provides a summary of the assumptions used in the model.

Table 97 Summary of overall model assumptions

Parameter	Assumption
Efficacy parameters	
Natural history of disease	The model assumes the origin of all metastases to be RCC or pNET in all cases, based on real-world evidence from the VHL Natural History Study and validation from clinical experts.
Number of surgeries	The model assumes that only one surgery related to the primary tumour can occur to reflect the MHRA label population for whom localised procedures unsuitable or undesirable, therefore reflecting a 'last resort' surgery resulting in loss of organ function and/or problematic sequelae.
Treatment efficacy of belzutifan on risk of metastatic disease	The risk reduction of metastatic disease in the belzutifan arm is equal to the risk reduction of surgery because belzutifan reduces both the risk of surgeries and the risk of metastatic disease by decreasing the size, or halting the growth, of tumours.
Treatment efficacy of belzutifan on risk of death	The risk reduction of death due to CNS Hb tumours in the belzutifan arm is assumed to be equal to the risk reduction of surgery because belzutifan reduces both the risk of death due CNS Hb tumours and the risk of surgeries by decreasing the size, or halting the growth, of CNS Hb tumours. The risk of death due to all other causes in the belzutifan arm is assumed to be equal to the SOC arm.
Metastatic disease treatments and risk of death on metastatic disease treatments	The market shares of first- and second-line metastatic disease treatments and the risk of death during the metastatic disease state are assumed to be equal between the belzutifan and SOC arms because prior use of belzutifan is assumed to have no influence on the treatment choice in metastatic disease or the risk of death on these metastatic disease treatments.
Treatment effect waning	MK-6482-004 does not provide conclusive evidence of treatment effect waning; however, to reflect this source of uncertainty, the potential impact of discontinuation of treatment effect was included in the base-case. Following belzutifan treatment discontinuation, clinical efficacy inputs are assumed to gradually converge towards those of the VHL Natural History Study. As belzutifan efficacy estimated from the MK-6482-004 trial data during this time period already accounts for any impact of belzutifan discontinuation, it would be implausible to consider treatment effect waning to occur before an equivalent time period has elapsed since treatment discontinuation. The time period between treatment discontinuation and the initiation is referred to as the period of residual treatment benefit. Treatment effect waning was assumed to initiate from the end of the period of residual treatment benefit and take effect gradually over a 2.71-year period. This time period is based on the magnitude of tumour shrinkage observed in MK-6482-004 compared to the tumour growth rate in the time period prior to belzutifan initiation. In the MK-6482-004 trial, at the tumour measurement that occurred closest to the time of discontinuing

Parameter	Assumption
	belzutifan, the average size of patients' largest RCC tumour was smaller than at baseline. Based on a conservative assumption that the rate of tumour growth immediately following discontinuation reverts to its pre-treatment growth rate, 2.71 years represents the amount of time until the largest RCC tumour resets to its baseline size. This duration is based on an estimate of annual tumour growth rate among patients treated with belzutifan in the MK-6482-004 trial. A similar assumption was made for the pNET and CNS Hb cohorts, due to the small sample size of discontinued patients in the CNS Hb and pNET subgroups who had an available CNS Hb and pNET measurement near to the time of treatment discontinuation.
Real-world SOC adjustment	Given that they were treated at the US National Cancer Institute (NCI, a Centre of Excellence), patients in the VHL Natural History Study are expected to have received an elevated SOC compared to normal UK clinical practice. Similarly, clinical trial conditions also reflect an elevated SOC. The model therefore adjusts the rate of surgeries and metastasis in both arms to account for this discrepancy (see Aligning risk of surgery and metastatic disease to real-world SOC for more details).
MHRA label adjustment	The MHRA label limits access to patients for whom localised procedures are unsuitable or undesirable, which is assumed to be a more severe population than the MK-6482-004 trial population and the VHL Natural History Study (which did not have this limitation). The following rates are therefore adjusted to reflect the MHRA label population versus the trial population in line with clinical expert opinion: • Perioperative mortality rates are adjusted upwards to reflect the more severe MHRA label population. • Surgical complication rates are adjusted upwards to reflect that complications are much more likely in a population for whom surgical procedures are undesirable or unsuitable. See *Surgical complications* for more details.
Safety	
Surgical complications	The model assumes that the risk of surgical complications is equal between the belzutifan and SOC arms because belzutifan is not expected to affect the rate of surgical complications. The surgical complications included in the model were validated by clinical experts. For non-primary tumours, patients are assumed only to be at risk of surgical complications from the first surgery because the model does not track the number of non-primary surgeries that patients undergo.
AEs	The model assumes no risk of AEs for patients managed by SOC alone. AEs for the belzutifan arm are sourced from MK-6482-004.
Treatments included in the model	
Intervention	All patients in the intervention arm are assumed to receive the recommended dose of belzutifan until discontinuation (120 mg administered orally once daily until disease progression or unacceptable toxicity) as per the SmPC (5). Dose adjustments and

Parameter	Assumption
	modifications have been reflected as per the MK-6482-004 trial data (see Appendix C1.1 SmPC for more information).
Comparator	The comparator in the model is SOC, defined as current established clinical management of VHL-RCC, VHL-pNET and VHL-CNS Hb as outlined in the NICE scope.
Utilities	
Health state utilities	Health state utility values are equal between the belzutifan and SOC arms. The utility associated with short-term and long-term complications is applied as a disutility.
Disutility of surgical complications	The disutility of short-term surgical complications was assumed to be restricted to the 28-day period following the tumour reduction procedure, in accordance with the timeframe in which the risks of complications were measured in the Optum Clinformatics Data Mart claims study. The disutility of long-term complications were similarly derived from the Optum Clinformatics Data Mart claims study, measured over a 180-day period.
Caregiver disutility	The disutility associated with caregivers is applied in a scenario analysis to reflect the burden from VHL disease. This disutility is sourced from published literature.
Number of caregivers	Each patient is assumed to have 1 caregiver.
Disutility of AEs	The disutilities of AEs for belzutifan are applied in the base-case.
Costs	
Drug costs	Intervention drug costs are at list price. All comparator drug costs are at list price unless PAS discount is available publicly. The mean relative dose intensity (RDI) is applied to all drug acquisition costs to account for any delays or interruptions in treatment.
Surgery and surgical complication costs	The model considers costs of surgery related to the primary tumour (i.e. RCC, pNET, CNS Hb) in addition to surgery not related to the primary tumour. The costs incurred by surgical complications is also accounted for. The model assumes that the risks of surgical complications are equal between the belzutifan and SOC arms of the model. Short-term of surgical complications costs are applied as a one-off cost per surgery. Annual costs of long-term surgical complications were converted into weekly costs, which were applied as a recurring per-cycle cost
Health care resource use	Resource use estimates are assumed equal for pre-surgery and event-free after surgery health states for all cohorts. Resource use estimates for metastatic disease are applied upon entering health state, pre-progression and post-progression.
Vial sharing	Vial sharing for metastatic disease therapies is assumed in the base case. A scenario analysis is explored where no vial sharing is allowed.

AE: adverse event; CNS: central nervous system; Hb: haemangioblastoma; NCI: National Cancer Institute; pNET: pancreatic neuroendocrine tumour; RCC: renal cell carcinoma; SOC: standard of care; TP: transition probability; US: United States; VHL: Von Hippel Lindau

B.3.10 Base-case results

Base-case incremental cost-effectiveness analysis results

VHL-RCC

The list price ICER for belzutifan versus SOC in the VHL-RCC cohort is £42,997 per QALY gained. For belzutifan, the total per patient costs are [redacted], the total per patient LYs gained are [redacted] and the total per patient QALYs gained are [redacted]. Compared with SOC, belzutifan is associated with incremental per patient costs of [redacted], [redacted] additional LYs, and [redacted] additional QALYs (with the severity modifier applied). The list price VHL-RCC cohort results are presented in Table 98.

Table 98 VHL-RCC results – List price

Technologies	Total costs (£)	Total QALYs	Total LYG	Inc. costs (£)	Inc. QALYs *	Inc. LYG	ICER of belzutifan versus SOC (£/QALY)
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	42,997

*Note: The x1.7 severity weight is applied to the incremental QALYs.
ICER: incremental cost effectiveness ratio; LYG: life year gain; QALY: quality-adjusted life year; SOC: standard of care

VHL-CNS Hb

The list price ICER for belzutifan versus SOC in the VHL-CNS Hb cohort is £56,933 per QALY gained. For belzutifan, the total per patient costs are [redacted], the total per patient LYs gained are [redacted] and the total per patient QALYs gained are [redacted]. Compared with SOC, belzutifan is associated with incremental per patient costs of [redacted], [redacted] additional LYs, and [redacted] additional QALYs (with the severity modifier applied). The list price VHL-CNS Hb cohort results are presented in Table 99.

Table 99 VHL-CNS Hb results – List price

Technologies	Total costs (£)	Total QALYs	Total LYG	Inc. costs (£)	Inc. QALYs*	Inc. LYG	ICER of belzutifan versus SOC (£/QALY)
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	33,490

*Note: The x1.7 severity weight is applied to the incremental QALYs.

ICER: incremental cost effectiveness ratio; LYG: life year gain; QALY: quality-adjusted life year; SOC: standard of care

VHL-pNET

The list price ICER for belzutifan versus SOC in the VHL-pNET cohort is £77,649 per QALY gained. For belzutifan, the total per patient costs are [redacted], the total per patient LYs gained are [redacted] and the total per patient QALYs gained are [redacted]. Compared with SOC, belzutifan is associated with incremental per patient costs of [redacted], [redacted] additional LYs, and [redacted] additional QALYs (with the severity modifier applied). The list price VHL-pNET cohort results are presented in Table 100.

Table 100 VHL-pNET results – List price

Technologies	Total costs (£)	Total QALYs	Total LYG	Inc. costs (£)	Inc. QALYs*	Inc. LYG	ICER of belzutifan versus SOC (£/QALY)
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	45,676

*Note: The x1.7 severity weight is applied to the incremental QALYs.
ICER: incremental cost effectiveness ratio; LYG: life year gain; QALY: quality-adjusted life year; SOC: standard of care

B.3.11 Exploring uncertainty

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was conducted to estimate the probability of belzutifan being cost-effective relative to SOC, based on different willingness-to-pay thresholds. A Monte-Carlo simulation with 1,000 iterations was conducted. In each iteration, the model inputs were randomly drawn from the specified distributions.

Table 101 summarizes the results from the PSA using the list price belzutifan. In the PSA using the list price, the ICER is £44,854 per QALY gained belzutifan versus SOC in the VHL-RCC population. The incremental per patient costs with belzutifan versus SOC are [redacted] in the VHL-RCC population and the incremental per patient QALYs gained are [redacted] in the VHL-RCC population with the severity modifier applied. The ICER is £34,352 per QALY gained belzutifan versus SOC in the VHL-CNS Hb population. The incremental per patient costs with belzutifan versus SOC are [redacted] in the VHL-CNS Hb population and the incremental per patient QALYs gained are [redacted] in the VHL-CNS Hb population with the severity modifier applied. The ICER is £46,966 per QALY gained belzutifan versus SOC in the VHL-pNET population. The incremental per patient costs with belzutifan versus SOC are [redacted] in the VHL-pNET population and the incremental per patient QALYs gained are [redacted] in the VHL-pNET population with the severity modifier applied.

Table 101 Total costs, QALYs and ICERs from the PSA (list price)

Population	Treatment	Total costs	Total QALYs	ICER*
VHL-RCC	**Belzutifan**	[redacted]	[redacted]	-
	SOC	[redacted]	[redacted]	£44,854
VHL-CNS Hb	**Belzutifan**	[redacted]	[redacted]	-
	SOC	[redacted]	[redacted]	£34,352
VHL-pNET	**Belzutifan**	[redacted]	[redacted]	-
	SOC	[redacted]	[redacted]	£46,966

*Note: The x1.7 severity weight is applied to the incremental QALYs in the ICER calculation.
CNS Hb: central nervous system haemangioblastoma; ICER: incremental cost effectiveness ratio; pNET: pancreatic neuroendocrine tumour; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life year; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

For belzutifan versus SOC, belzutifan list price scatterplots of simulated incremental cost and QALYs (with the severity modifier applied) are presented in Figure 25, Figure 26, and Figure 27 for the VHL-RCC, VHL-CNS Hb and VHL-pNET populations, respectively.

Figure 25 Scatterplots of incremental costs and effectiveness for belzutifan vs. SOC in the VHL-RCC population across 1,000 iterations of the PSA (list price)

PSA: probabilistic sensitivity analysis; RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Figure 26 Scatterplots of incremental costs and effectiveness for belzutifan vs. SOC in the VHL-CNS Hb population across 1,000 iterations of the PSA (list price)

CNS: central nervous system; Hb: haemangioblastoma; PSA: probabilistic sensitivity analysis; SOC: standard of care; VHL: Von Hippel Lindau

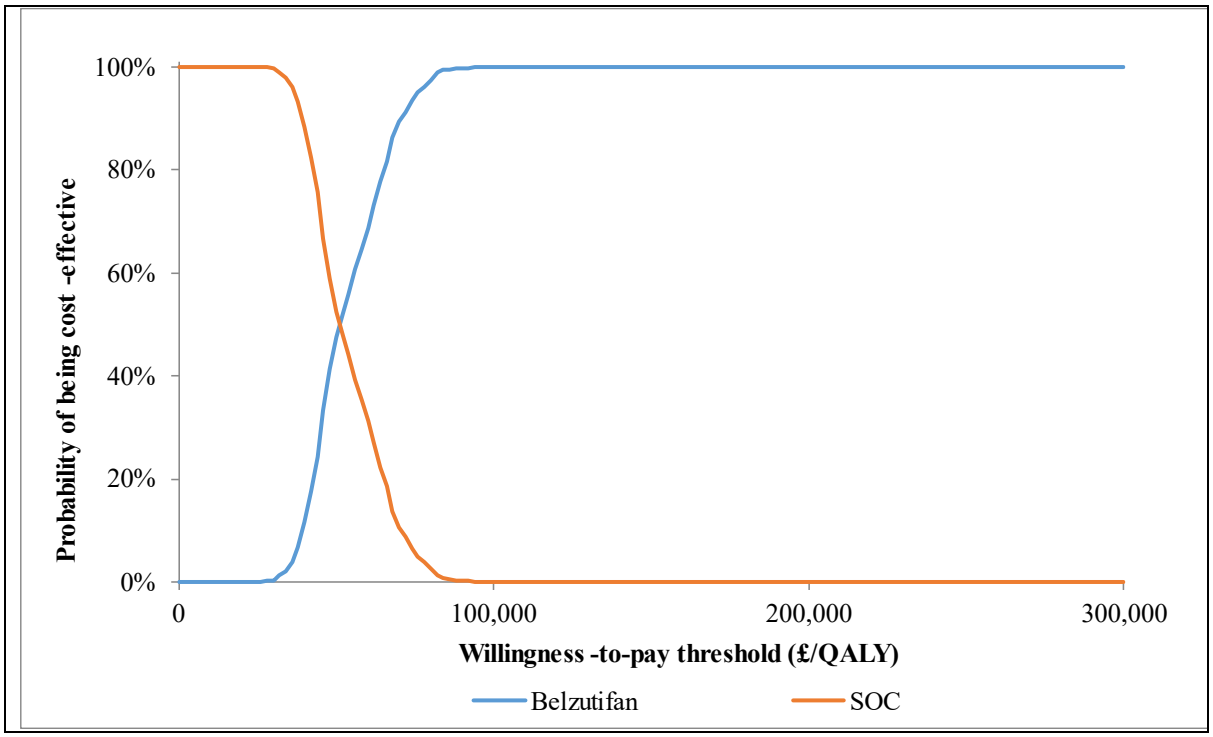
Figure 27 Scatterplots of incremental costs and effectiveness for belzutifan vs. SOC in the VHL-pNET population across 1,000 iterations of the PSA (list price)

pNET: pancreatic neuroendocrine tumour; PSA: probabilistic sensitivity analysis; SOC: standard of care; VHL: Von Hippel Lindau

Cost-effectiveness acceptability curves are presented in Figure 28, Figure 29, and Figure 30 for the VHL-RCC, VHL-CNS Hb and VHL-pNET populations, respectively,

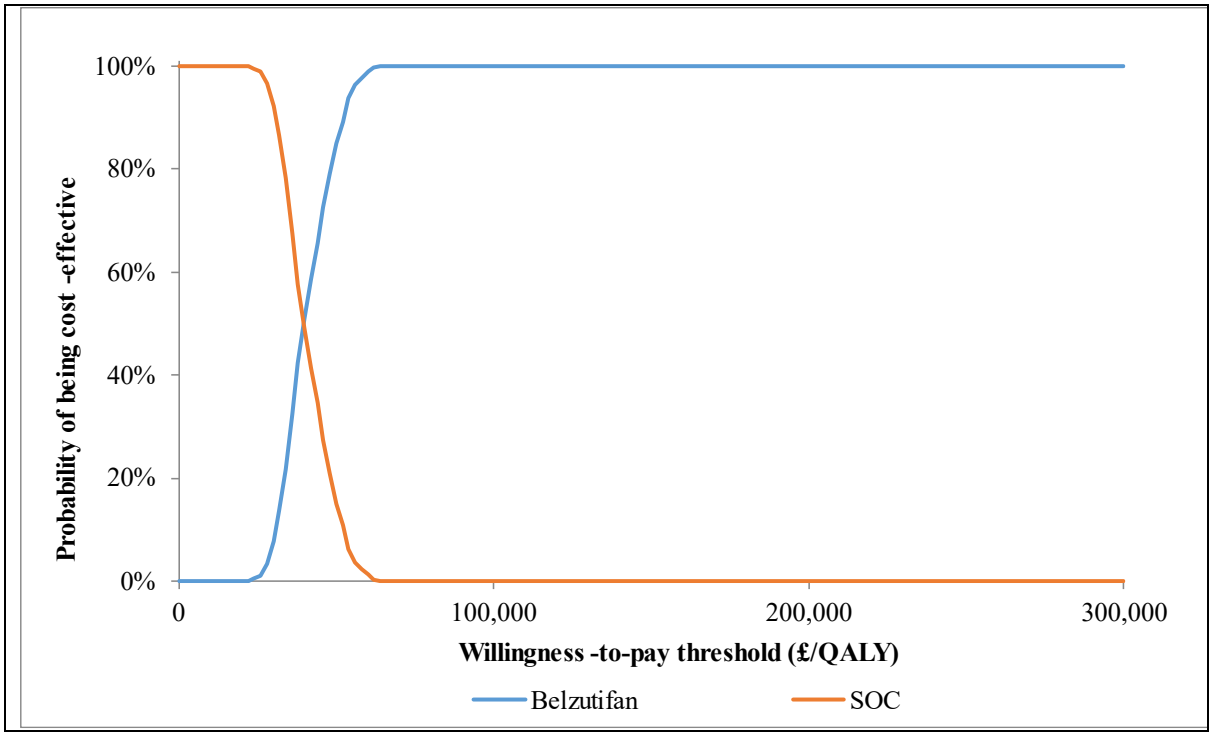
using the list price of belzutifan. The cost-effectiveness acceptability curves show the probability of belzutifan being cost-effective versus SOC of 0.3% for the VHL-RCC cohort, 7.8% for the VHL-CNS Hb cohort and 0.2% for the VHL-pNET cohort at a willingness-to-pay threshold of £30,000 per QALY gained.

Figure 28 Cost-effectiveness acceptability curves for belzutifan vs. SOC in the VHL-RCC population (list price)



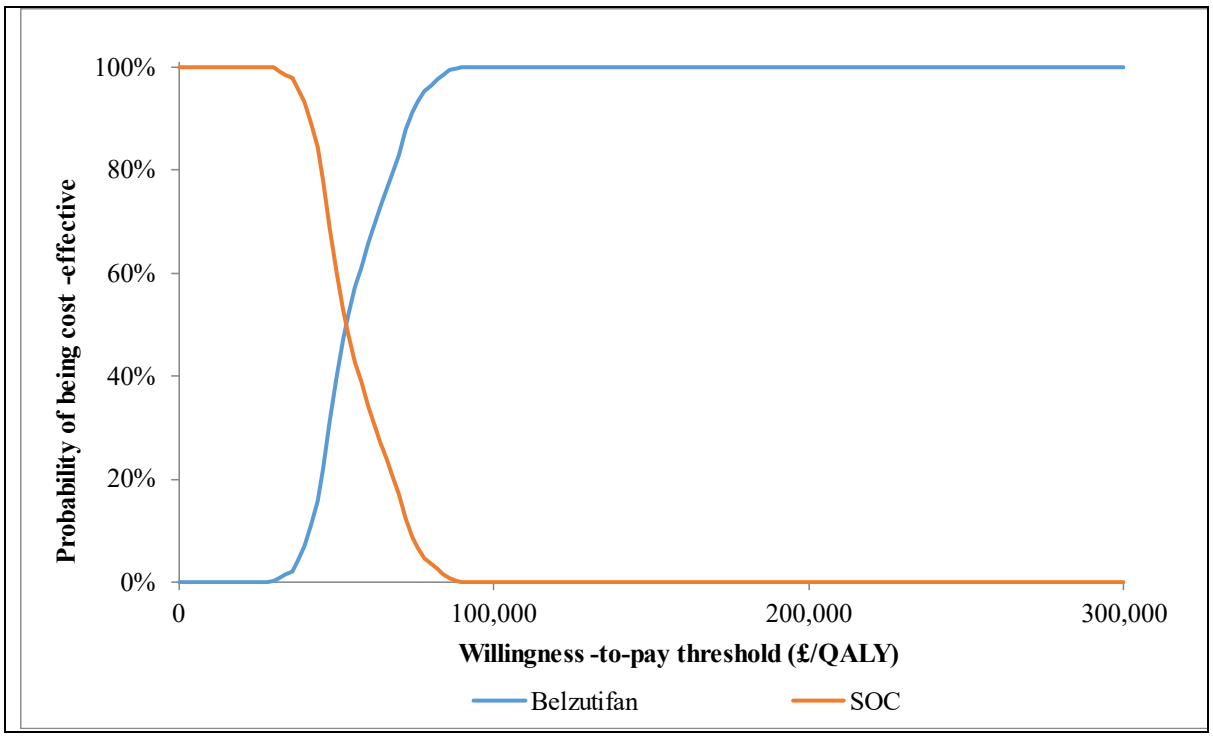
RCC: renal cell carcinoma; SOC: standard of care; VHL: Von Hippel Lindau

Figure 29 Cost-effectiveness acceptability curves for belzutifan vs. SOC in the VHL-CNS Hb population (list price)



CNS: central nervous system; Hb: haemangioblastoma; SOC: standard of care; VHL: Von Hippel Lindau

Figure 30 Cost-effectiveness acceptability curves for belzutifan vs. SOC in the VHL-pNET population (list price)



pNET: pancreatic neuroendocrine tumour; SOC: standard of care; VHL: Von Hippel Lindau

Deterministic sensitivity analysis

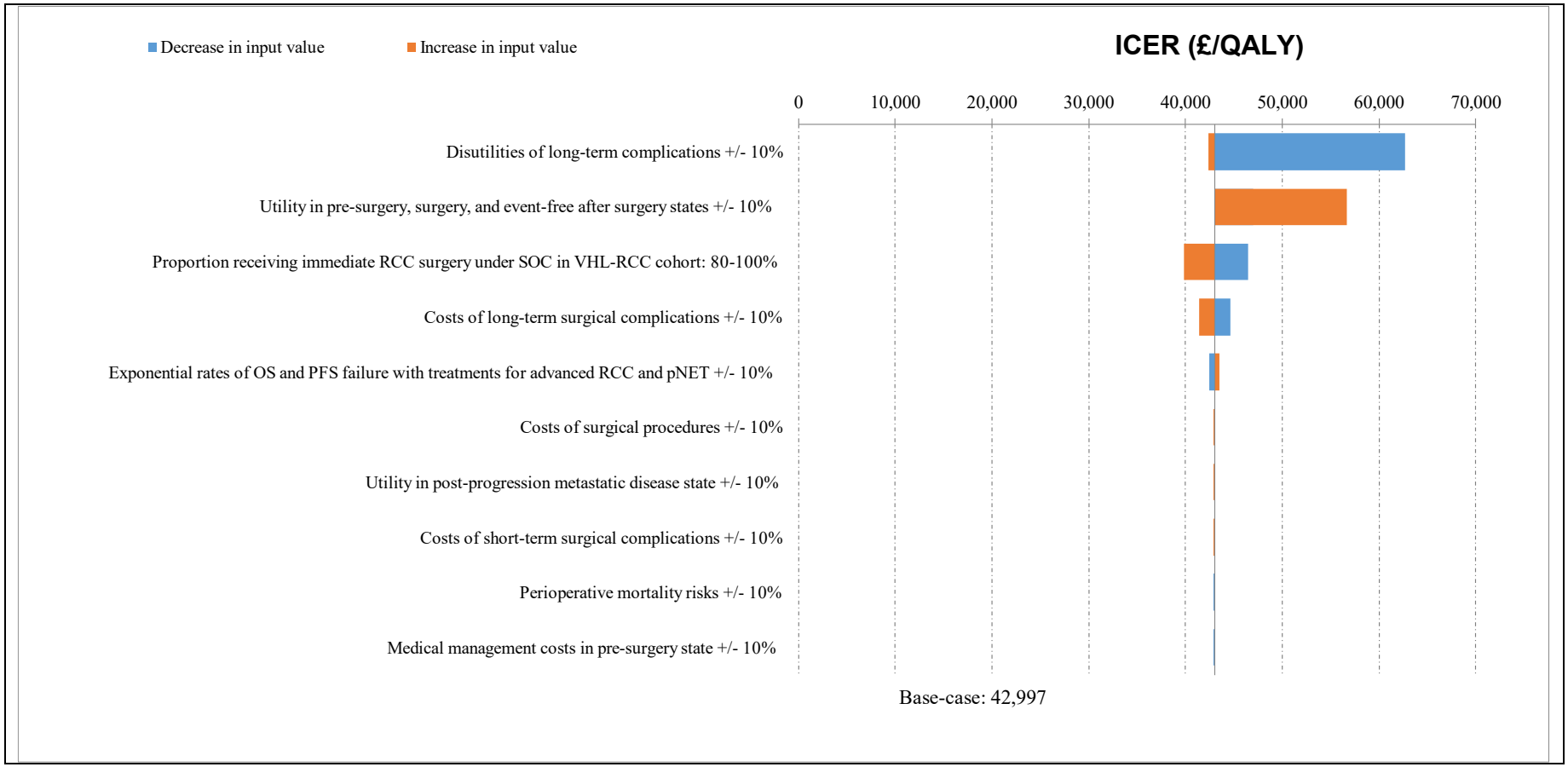
The results of the deterministic sensitivity analysis (DSA) are presented in the tornado diagram in

Figure 31-Figure 33, which illustrates the 15 parameters that have the greatest influence on the ICER for each cohort. These results have the severity modifier applied in the QALY calculation. Tabular results of the DSA are shown in *Appendix*

J1.4 Full DSA and scenario analysis results.

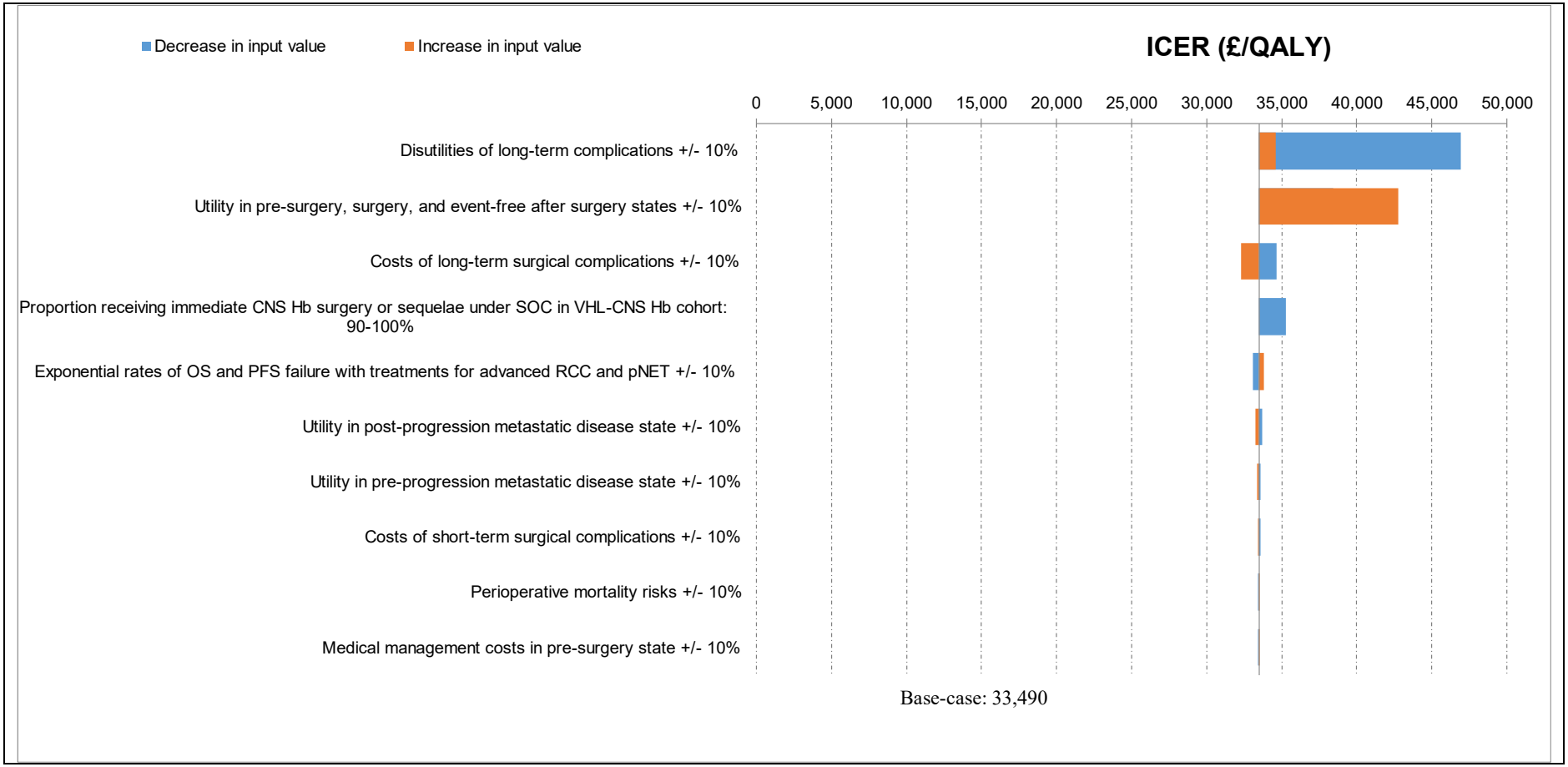
The disutility associated with long-term complications is the most sensitive parameter and the utility associated with the non-metastatic health states is the second most sensitive parameter across all three cohorts. The results should be interpreted with caution as both an increase and decrease in the parameter value can lead to an increase in the ICER due to a change in the severity modifier weighting which is not appropriate in this appraisal.

Figure 31 DSA results – tornado diagram for belzutifan vs. SOC in the VHL-RCC population (list price)



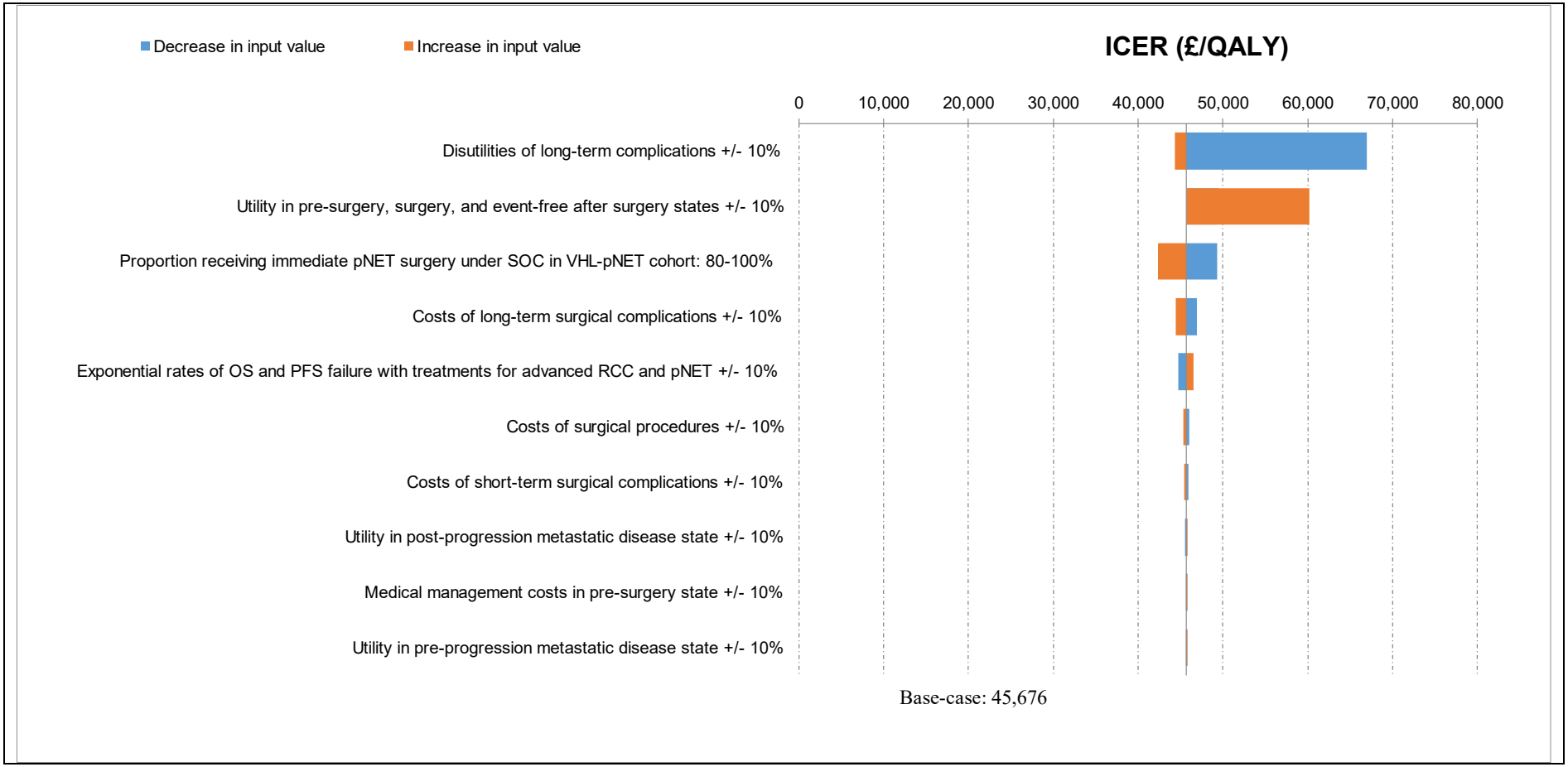
DSA: deterministic sensitivity analysis; RCC: renal cell carcinoma; SOC: standard of care; ToT: Time on Treatment; VHL: Von Hippel Lindau

Figure 32 DSA results – tornado diagram for belzutifan vs. SOC in the VHL-CNS Hb population (list price)



CNS: central nervous system; DSA: deterministic sensitivity analysis; Hb: haemangioblastoma; SOC: standard of care; ToT: Time on Treatment; VHL: Von Hippel Lindau

Figure 33 DSA results – tornado diagram for belzutifan vs. SOC in the VHL-pNET population (list price)



DSA: deterministic sensitivity analysis; pNET: pancreatic neuroendocrine tumour; SOC: standard of care; ToT: Time on Treatment; VHL: Von Hippel Lindau

Scenario analyses

To assess structural uncertainty of the cost-effectiveness results, scenario analyses were conducted by varying one model input or assumption at a time. The results of the 10 scenarios that have greatest influence on the ICER for each cohort sensitive parameters are presented from most to least sensitive in Table 102-Table 104 below. These results have the severity modifier applied in the QALY calculation, some of these results should be interpreted with caution as a lower weighted severity modifier is applied which is not appropriate for this appraisal.

Base-case results are generally robust to changes tested across the broad range of scenarios. The most impactful scenario across the three cohorts is the removal of treatment effect waning.

Table 102 Scenario analysis results of the most sensitive parameters for VHL-RCC cohort

Scenario	ICER (£/QALY), belzutifan vs SOC	% Change in ICER from base case
Base case	**42,997**	**-**
Assume no treatment effect waning: Model efficacy and ToT separately	15,831	-63.18%
Do not adjust surgery and metastases rates to account for real-world standard of care	63,178	46.94%
Time horizon: 30 years	59,623	38.67%
Time horizon: 20 years	57,096	32.79%
Distribution for belzutifan ToT: Weibull	53,685	24.86%
Include indirect costs (societal perspective)	35,597	-17.21%
Do not apply relative dose intensity	47,944	11.50%
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	39,535	-8.05%
Apply caregiver disutility	41,141	-4.32%
Distribution for pre-surgery→surgery in the belzutifan arm (VHL-RCC cohort): Gamma	44,781	4.15%

AE: adverse event; CNS: central nervous system; CR: complete response; DSA: deterministic sensitivity analysis; Hb: haemangioblastoma; ICER: incremental cost effectiveness ratio; IV: intravenous; PD: progressed disease; pNET: pancreatic neuroendocrine tumour; PR: partial response; QALY: quality adjusted life year; RCC: renal cell carcinoma; SD: stable disease; SOC: standard of care; TP: transition probability; VHL: Von Hippel Lindau

Table 103 Scenario analysis results of the most sensitive parameters for VHL-CNS Hb cohort

Scenario	ICER (£/QALY), belzutifan vs SOC	% Change in ICER from base case
Base case	**33,490**	**-**
Assume no treatment effect waning: Model efficacy and ToT separately	10,385	-68.99%
Time horizon: 30 years	46,867	39.93%
Time horizon: 20 years	46,025	37.43%
Distribution for belzutifan ToT: Weibull	42,057	25.58%
Do not adjust surgery and metastases rates to account for real-world standard of care	41,584	24.17%
Include indirect costs (societal perspective)	28,041	-16.27%
Do not apply relative dose intensity	37,272	11.29%
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	30,644	-8.50%
Annual discount rate: 0.0%	31,097	-7.15%
Apply caregiver disutility	31,468	-6.04%

AE: adverse event; CNS: central nervous system; CR: complete response; DSA: deterministic sensitivity analysis; Hb: haemangioblastoma; ICER: incremental cost effectiveness ratio; IV: intravenous; PD: progressed disease; pNET: pancreatic neuroendocrine tumour; PR: partial response; QALY: quality adjusted life year; RCC: renal cell carcinoma; SD: stable disease; SOC: standard of care; TP: transition probability; VHL: Von Hippel Lindau

Table 104 Scenario analysis results of the most sensitive parameters for VHL-pNET cohort

Scenario	ICER (£/QALY), belzutifan vs SOC	% Change in ICER from base case
Base case	**45,676**	**-**
Assume no treatment effect waning: Model efficacy and ToT separately	13,560	-70.31%
Time horizon: 20 years	65,994	44.48%
Time horizon: 30 years	64,661	41.57%
Distribution for belzutifan ToT: Weibull	56,748	24.24%
Include indirect costs (societal perspective)	35,367	-22.57%
Annual discount rate: 0.0%	38,457	-15.80%
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	39,930	-12.58%
Do not apply relative dose intensity	50,945	11.54%

Annual discount rate: 1.5%	41,523	-9.09%
Apply caregiver disutility	42,868	-6.15%

AE: adverse event; CNS: central nervous system; CR: complete response; DSA: deterministic sensitivity analysis; Hb: haemangioblastoma; ICER: incremental cost effectiveness ratio; IV: intravenous; PD: progressed disease; pNET: pancreatic neuroendocrine tumour; PR: partial response; QALY: quality adjusted life year; RCC: renal cell carcinoma; SD: stable disease; SOC: standard of care; TP: transition probability; VHL: Von Hippel Lindau

B.3.12 Subgroup analyses

No subgroup analyses were conducted as part of this economic evaluation.

B.3.13 Benefits not captured in the QALY calculation

VHL disease can result in severe and debilitating symptoms if tumours are not controlled. It is common for patients with VHL-associated RCC, CNS Hb or pNET tumours to undergo multiple surgeries throughout their lifetimes to control the disease, and these surgeries result in further deterioration in patient QoL. In such patients, the requirement for multiple surgeries results in not only reduced well-being of patients and direct costs to the NHS through increased resource use, but also a greater burden on social care services, family members, and other caregivers who are required to spend time and resources assisting patients as they lose independence due to VHL. Belzutifan will be the first systemic therapeutic intervention available in the UK for the treatment of VHL-associated RCC, CNS Hb and pNET tumours in patients for whom localised procedures are unsuitable or undesirable. By slowing disease progression and preventing the requirement for severely detrimental surgeries, belzutifan has the potential to provide further benefits to patients, carers, and wider society that are not captured in the cost per QALY calculations in Section B.3.10 Base-case results.

Value of belzutifan not currently captured, or not sufficiently captured in the QALY calculation includes:

1. The value of reducing symptoms across multiple systems: the structure of the model focuses on the ‘worst’ primary tumour. For the most adversely impacted patients, significant benefit will be gained by both preserving organ function AND preventing tumours’ advance at other tumour sites. This might include preserving

both renal and pancreatic function, maintaining sight, reducing neurological symptoms.

2. The VHL RW QoL Disease Burden Study EQ-5D study captured VHL-patient utility scores based on self-reported response status using the EQ-5D questionnaire. The study did not report utility scores for the types of interventions likely to be received by patients described in the MHRA indication. We believe this underestimates the negative impact on HRQoL for the target patient population.

3. Due to scarcity of data in the trial, we had to model OS in the non-metastatic health states (from pre-surgery and event-free after surgery) based on mortality data from the VHL Natural History Study and capped using life tables reflecting a minimum mortality risk of the general population. We cap OS at either background mortality or VHL Natural History Study mortality. This is likely to substantially overestimate life expectancy in the patient population described in the MHRA indication, therefore underestimating the potential value of belzutifan.

4. While the EQ-5D interviewed patients with VHL, we do not believe it is a sensitive enough tool to accurately reflect of the mental health impact of this disease. We note a suicide and a fentanyl overdose death in the study. This is extremely unusual in such a small patient population. Neither of these deaths were attributed to belzutifan, instead we postulate they were the result of living with VHL. Clinicians frequently refer to suicidal ideation in their VHL patients.

5. Reduced anxiety associated with frequent scans, fear of disability or death from surgery, and relief in knowing tumour manifestations as a whole are controlled in comparison to multiple different surgeries for different tumour sites.

6. We include some disutility for carers in the economic model. However, for this inherited disease, having no treatment option is significantly more negatively

impactful on family mental health than having a treatment option. This is not currently valued in the economic model.

The effects of VHL disease, particularly in severely affected patients for whom localised procedures are unsuitable or undesirable, have significant negative impacts on the psychological well-being of patients. The requirement of patients to undergo repeated surgical procedures, which in themselves result in deterioration in health, can often cause patients to lose hope for the future and reduce their ability to take part in their normal daily activities. As previously described, this may be reflected by the two patients who have died in the MK-6482-004 trial, one of whom died from suicide and the other from a fentanyl overdose, both following several surgical procedures. By delaying and slowing the need for surgeries, belzutifan is therefore expected to improve the mental well-being (including a reduction in patient anxiety associated with scans) and offer optimism and much-needed hope to patients, which may not be captured in the stringent framework of the EQ-5D questionnaire and QALY measures. Furthermore, this appraisal assesses the cost-effectiveness of belzutifan in the VHL-RCC, VHL-CNS Hb and VHL-pNET cohorts in isolation and does not account for its benefit in reducing tumour size for patients with multiple tumour manifestations since its mechanism of action is not restricted solely to the primary tumour. This provides further relief for patients in knowing that their tumour manifestations as a whole are being controlled rather than requiring multiple different surgeries for different tumour sites.

As a result of the reduced burden of surgical procedures expected with belzutifan treatment, patients may see increases in their work productivity through reduced sick leave from work and medical appointments. Patients may also be able to work later into life as a result of delaying the most severe and debilitating surgical complications, which can prevent patients from working and lead to early retirement or loss of job opportunities. The VHL RW QoL Disease Burden Study reported that 29.3% of hours of a standard work week were lost out due to VHL disease for patients without metastatic disease, whilst 40.6% of hours were lost by patient with metastatic disease. This

significant absenteeism is not only a burden to the workforce but to VHL patients as they pursue activities of daily living.

Caregivers of patients with VHL disease may also benefit from improved work productivity. Due to the severely debilitating effects of VHL tumours and the surgeries required to control them, the burden on family and/or social caregivers as a result of VHL disease can be significant. In particular, the serious complications of surgeries for CNS Hb tumours can include paralysis and severe nerve damage, resulting in a need for 24/7 care for some patients, a responsibility that most often falls to family members. For RCC surgeries, this may result in a need for support with transport to and from dialysis or assistance with home dialysis. A recent study exploring the quality of life in caregivers of patients receiving haemodialysis found that over one-third of carers spend 3 hours or more per day caring for patient receiving dialysis (125). The work productivity of caregivers of patients with VHL can therefore also be expected to improve due to belzutifan's impact on delaying and/or reducing surgeries over a patient's lifetime, and these benefits may spill over to other government bodies.

Some costs related to VHL management which are incurred by both patients and caregivers are not reimbursed by the NHS, such as the costs of travelling to medical appointments. These costs are like to be reduced in patients treated with belzutifan due to reduced need for surgical or other localised procedures, and cost savings would also accrue from reduced need for medical treatment for the complications of such procedures.

Belzutifan has obtained regulatory approval via the MHRA ILAP pathway, which is reserved for medicines which are innovative and aimed at the treatment of patient populations with a very high unmet need. With MHRA ILAP approval, the UK became the first country to confirm regulatory approval of belzutifan worldwide, demonstrating its commitment to bringing to market innovative therapies for rare diseases with a high unmet need. If given a positive recommendation by NICE, the UK will become one of the first countries globally to approve belzutifan for reimbursement, demonstrating the

success of the ILAP pathway and offering patients in the UK the opportunity for accelerated access to a step-change in the treatment of VHL. As belzutifan is one of the first drugs in the UK to be approved through the ILAP pathway, demonstrating the success of the ILAP may be an important milestone for future patient access in the UK.

B.3.14 Validation

Validation of cost-effectiveness analysis

Verification

To verify the results of the cost-effectiveness analysis, internal quality control procedures were undertaken by the model developer team to ensure that the mathematical calculations are performed correctly and are consistent with the model's specifications. An early version of the model was also independently reviewed by external health economists at the [redacted], who evaluated the model from an overall health economics perspective.

External validation

In addition to model verification and internal quality control, external validity of the model was also assessed. Due to the paucity of data available to validate the outcomes for the target population stipulated by the MHRA label, modelled efficacy outcomes are assessed against the original sources that informed the efficacy inputs without the assumption of immediate surgery for the SOC arm or adjustments to reflect real-world SOC to allow for interpretable comparisons. These are presented in *Validation of transition probabilities*.

No clinical trials or real-world evidence studies were identified which could be used to externally validate the modelled outcomes in the current appraisal, given the highly specific wording of the eligible population per the MHRA label (i.e. patients for whom localised procedures are unsuitable or undesirable). However, advice was sought in multiple engagements from clinical experts with experience treating the diverse manifestations of VHL disease. These UK clinical experts were consulted via an

advisory board and in individual consultation meetings to validate the assumptions used in the model regarding efficacy inputs and other key model assumptions from a clinical perspective (*Assumptions* for more details on assumptions validated by clinical experts).

The discussion at the advisory board focused on:

- General insights and unmet need on VHL
- Treatment pathways and insights for each of the 3 VHL tumour manifestations included in the MHRA label
- Other common VHL manifestations
- Primary goals of VHL treatment
- Interpretation of the MHRA label

A summary of the UK clinical expert validation process is provided below:

- Clinical experts were selected based on experience treating VHL-associated tumour manifestations, ensuring a broad geographic range in the UK
- The experts were engaged through individual consultation meetings. One expert was contracted for a series of consultations where an honorarium was paid. Some of the experts were later engaged through an advisory board, where an honorarium was paid.
- The experts were suitably qualified to provide input on the evidence submission, having the following experience:
    - A consultant endocrinologist who runs a VHL MDT that manages 50-60 VHL patients, inclusive of RCC, CNS Hb and pNET patients.

- Three additional consultant endocrinologists that attend VHL MDTs, who have expertise in recommending and referring VHL RCC, CNS Hb and pNET patients for treatment.
- Four consultant urological surgeons, the respective regional leads for VHL surgery, specialising in treatment for VHL RCC.
- A consultant neurologist, neurosurgeon and a neuro-oncologist who manage patients with VHL-CNS Hb.
- Three consultant clinical geneticists, respective leads of regional genetics services that manage patients with VHL.
- A professor of medical genetics and lead author on an evaluation of tumour surveillance protocols and outcomes in VHL disease in the UK.
- An interventional radiologist who provides non-surgical treatment for patients with VHL RCC.
- Two medical oncologists that have experience with treating VHL tumours.

- During individual consultation meetings and the advisory board, experts were asked questions about their experience managing VHL, typical patient profiles and patient management, how they would interpret the MHRA label population, expected consequences of surgery in the MHRA label population, interpretation of belzutifan clinical data, and validation of model assumptions.

Cross-validation

The current analysis represents the first cost-effectiveness evaluation of treatments for VHL-associated RCC, pNET, and CNS Hb. Consequently, there is limited potential for cross-validation of the current model results against other, independently developed economic evaluations in the same indication. Previous NICE appraisals provide precedent for some of the assumptions used in the economic analysis. This includes the

use of OS data from an ITC of clinical trials of treatments of metastatic disease to inform the risk of death in the metastatic disease health state, an approach previously accepted in the appraisal of pembrolizumab for adjuvant treatment of RCC (TA830). Other approaches previously validated include the consideration of Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan or Grade ≥3 TRAEs occurring in >0% of patients treated with belzutifan in the model, which is standard practice for NICE appraisals. Utility values previously accepted by NICE in TA830 were also used to inform some health state utility values (see *Health-related quality-of-life data used in the cost-effectiveness analysis* section for more information).

B.3.15 Interpretation and conclusions of economic evidence

VHL is a highly complex disease with multiple manifestations, which have diverse impacts on HRQL and survival. The current analysis not only reflects these multiple manifestations (including both primary and non-primary tumours), but also the manifold complications from surgery which vary significantly by primary tumour type. These sequelae have wide-ranging effects on health (both patient and caregiver) and costs (both direct medical costs and costs related social care/support).

The population assessed in the economic analysis reflected the multisystemic nature of VHL disease and considered different potential mechanisms of incremental health benefit in patients treated with belzutifan (relative to SOC). Belzutifan was estimated to improve QALY gains primarily through reducing the risk of both primary and non-primary VHL-related tumour surgeries, and in turn reducing surgical complications and perioperative deaths. The available evidence from MK-6482-004, which showed favourable rates of response and disease control (per RECIST v1.1 criteria) for belzutifan, was linked not only to improvements patients' HRQoL in non-metastatic health states, but to reductions in the risk of metastases from any VHL-related origin tumour and reduce mortality risk linked to progression of CNS Hb.

In the base case, belzutifan was estimated to generate [redacted], [redacted], [redacted] additional QALYs and [redacted], [redacted], [redacted] additional LYs compared to the current SOC (established clinical

management without belzutifan) in the VHL-RCC, VHL-CNS Hb and VHL-pNET cohorts respectively. Parameter and structural uncertainty were explored through PSA, univariate OWSA and scenario analysis. Overall, the sensitivity and scenario analyses explored indicate that, using the base-case assumptions as a basis, the cost effectiveness of belzutifan is most sensitive to the utility in the non-metastatic health states, the proportion to receive immediate surgery in the SOC arm and the removal of treatment effect waning. The current economic analysis specifically considered adult patients who require therapy for VHL-RCC, VHL-CNS Hb or VHL-pNET and for whom localised procedures are unsuitable or undesirable. This population is fully aligned with marketing authorisation per the MHRA license, and therefore results can be considered directly generalisable to patients eligible to receive belzutifan in UK clinical practice.

A key challenge in developing a robust economic analysis for this appraisal was the unanticipated expansion of the MHRA label to two additional tumour manifestations (CNS Hb and pNET) beyond RCC tumours, the population around which the MK-6482-004 trial and real-world evidence studies were designed. Despite the considerable data availability challenges, the three VHL cohorts specified in the MHRA have been each modelled separately, reflecting the diverse set of outcomes specific to each primary tumour site. The assessment per VHL cohort provides an indication of cost-effectiveness of belzutifan in each group but should be viewed in light of the unavoidable data gaps unique to each patient group. More importantly, given what is known about the progression of VHL, these VHL cohorts should not be strictly considered as discrete subgroups perfectly distinct from one another as clinical trial subgroups often are. Nor should any VHL cohort be considered perfectly representative of a patient group afflicted with a specific VHL-associated primary manifestation. For example, a patient with pNET as the primary VHL tumour manifestation may have a diverse mix of VHL-associated sequelae that may overlap with other VHL cohorts over the course of their lives.

Beyond the inclusion of the two additional VHL cohorts, the analysis also reflected a highly specific VHL population who require therapy and for whom localised procedures

are unsuitable or undesirable. This is a more restrictive population than the MK-6482-004 trial, which did not restrict recruitment in line with the MHRA licenced population, which adds some uncertainty into an evaluation in what is already a rare disease with significant data gaps. Where possible, adjustments to model parameters such as risks of surgical complications and perioperative mortality to reflect the population likely to be treated with belzutifan more in line with the MHRA label.

Despite the considerable challenges of developing a robust economic model for such a rare and complex disease, the current analysis made best use of data from the MK-6482-004 trial of belzutifan and two real world evidence studies: the VHL Natural History Study and Optum Clinformatics Data Mart study. The use of multiple data sources with varying study designs required calibration of the outcomes data to ensure comparability in how it fed into the model, not only across studies but alignment with the eligible population for belzutifan in the UK, taking into account differences between the MHRA label, the clinician trial, and RWE.

The Markov cohort structure utilized in the current analysis and multi-state parametric modelling approach for the estimation of transition probabilities are well-established modelling approaches that have been commonly used in published cost-effectiveness analyses and prior health technology appraisals in non-metastatic oncology indications such as non-small-cell lung cancer, melanoma, and RCC (53, 126-128). Given the heterogenous nature of VHL, it would be of interest to consider how a patient-level simulation model might have better reflect the diverse outcomes of VHL across its various manifestations that the current Markov model was not able to account for. There are extensive data requirements for a patient-level simulation model, however, which may present challenges in the context of a rare disease such as VHL.

Despite the limitations resulting from the paucity of available data in the specific VHL population indicated by the MHRA label, the model does provide a robust economic analysis of belzutifan for the treatment of VHL-RCC, VHL-CNS Hb and VHL-pNET in patients who require therapy for whom localised procedures are unsuitable or

undesirable. Clinical expert input has been utilized to the fullest extent possible to validate key clinical assumptions. Additionally, the adjustments accounting for the MHRA license misalignment and the elevated SOC in the VHL Natural History Study represent the care taken to ensure the economic analysis aligns as closely as possible to the target population of the appraisal.

Due to constraints of the model structure and/or data, the cost-effectiveness model developed for this appraisal does not capture the full benefit of belzutifan as described in *B.3.13 Benefits not captured in the QALY calculation*. Specifically, the societal benefits of increased work productivity for both patients and carers, reduced out-of-pocket costs for patients, reduced need for government assistance and the clinical benefits of belzutifan acting to reduce tumour burden on multiple sites simultaneously. Hence, the value of belzutifan in this patient population is likely underestimated in the cost-effectiveness analysis.

Due to the rarity of VHL, there is limited opportunity for future analyses to supplement the robustness of completeness of the results, other than future readouts from the MK-6482-004 trial. However, the phase 2, single-arm MK-6482-015 trial assessing belzutifan in a population partially overlapping the population relevant to the current appraisal may provide some confirmatory data when it reports results, estimated in 2026. The proposed patient registry collecting data on real world outcomes associated with belzutifan treatment is current under discussion with the MHRA and would provide data on patients in the UK receiving belzutifan in its licenced MHRA indication and would represent a dataset most aligned to the current appraisal.

The economic analysis developed for this NICE appraisal is the first and only cost-effectiveness analysis developed in patients with VHL-associated RCC, CNS Hb, or pNET, irrespective of suitability/desirability for localised procedures. This lack of previously published economic evaluations highlights that VHL patients have been an underserved population with no access to effective treatment options in VHL. The unmet need for effective therapies is extremely high in this vulnerable population, and

outcomes estimated in model for SOC and belzutifan highlight the transformative value of belzutifan as a step change in the pathway of care for VHL-associated RCC, pNET, and CNS Hb.

B.4 References

1. Chan VW-S, Lenton J, Smith J, Jagdev S, Ralph C, Vasudev N, et al. Multimodal image-guided ablation on management of renal cancer in Von-Hippel-Lindau syndrome patients from 2004 to 2021 at a specialist centre: A longitudinal observational study. European Journal of Surgical Oncology. 2022;48(3):672-9.
2. NICE. Highly Specialised Technologies (HST) criteria checklist - ID3932 Belzutifan for treating tumours associated with von Hippel-Lindau disease2022. Available from: https://www.nice.org.uk/guidance/gid-ta10817/documents/supporting-documentation.
3. MSD. HTA Advisory Board summary slides. [MSD Data On File] [Internet]. 2022.
4. Stueven AK, Kayser A, Wetz C, Amthauer H, Wree A, Tacke F, et al. Somatostatin Analogues in the Treatment of Neuroendocrine Tumors: Past, Present and Future. Int J Mol Sci. 2019;20(12).
5. MHRA. WELIREG 40 MG FILM-COATED TABLETS. Summary of Product Characteristics [Internet]. 2022. Available from: https://products.mhra.gov.uk/search/?search=welireg&page=1&doc=Spc&rerouteType=0.
6. EMA. EU/3/20/2324: Orphan designation for the treatment of von Hippel-Lindau disease2020. Available from: https://www.ema.europa.eu/en/medicines/human/orphan-designations/eu3202324.
7. Varshney N, Kebede AA, Owusu-Dapaah H, Lather J, Kaushik M, Bhullar JS. A Review of Von Hippel-Lindau Syndrome. J Kidney Cancer VHL. 2017;4(3):20-9.
8. FDA. Rare Diseases at FDA2021 17-MAR-2021. Available from: https://www.fda.gov/patients/rare-diseases-fda.
9. Ploussard G, Droupy S, Ferlicot S, Ples R, Rocher L, Richard S, et al. Local recurrence after nephron-sparing surgery in von Hippel-Lindau disease. Urology. 2007;70(3):435-9.
10. Kaelin WG, Jr. HIF2 Inhibitor Joins the Kidney Cancer Armamentarium. J Clin Oncol. 2018;36(9):908-10.
11. MSD. Mechanism of Action - Welireg: Merck Sharp & Dohme; 2022 [Available from: https://www.welireghcp.com/mechanism-of-action/.
12. Patel A. Benign vs Malignant Tumors. JAMA Oncology. 2020;6(9):1488-.
13. Maher ER, Adlard J, Barwell J, Brady AF, Brennan P, Cook J, et al. Evaluation of tumour surveillance protocols and outcomes in von Hippel-Lindau disease in a national health service. Br J Cancer. 2022.
14. Schunemann V, Huntoon K, Lonser RR. Personalized Medicine for Nervous System Manifestations of von Hippel-Lindau Disease. Front Surg. 2016;3:39.

15. Lonser RR, Glenn GM, Walther M, Chew EY, Libutti SK, Linehan WM, et al. von Hippel-Lindau disease. Lancet. 2003;361(9374):2059-67.

16. Wilding A, Ingham SL, Lalloo F, Clancy T, Huson SM, Moran A, et al. Life expectancy in hereditary cancer predisposing diseases: an observational study. J Med Genet. 2012;49(4):264-9.

17. VHL UK / Ireland. Stories: VHL UK / Ireland; 2022 [Available from: https://vhl-uk-ireland.org/stories/.

18. Casserly LF, Dember LM. Thrombosis in end-stage renal disease. Semin Dial. 2003;16(3):245-56.

19. Reinecke H, Brand E, Mesters R, Schäbitz WR, Fisher M, Pavenstädt H, et al. Dilemmas in the management of atrial fibrillation in chronic kidney disease. J Am Soc Nephrol. 2009;20(4):705-11.

20. Kuno T, Takagi H, Ando T, Sugiyama T, Miyashita S, Valentin N, et al. Oral Anticoagulation for Patients With Atrial Fibrillation on Long-Term Hemodialysis. J Am Coll Cardiol. 2020;75(3):273-85.

21. Tveit DP, Hypolite IO, Hshieh P, Cruess D, Agodoa LY, Welch PG, et al. Chronic dialysis patients have high risk for pulmonary embolism. Am J Kidney Dis. 2002;39(5):1011-7.

22. Cheema A, Singh T, Kanwar M, Chilukuri K, Maria V, Saleem F, et al. Chronic kidney disease and mortality in implantable cardioverter-defibrillator recipients. Cardiol Res Pract. 2010;2010.

23. Wachterman MW, O'Hare AM, Rahman OK, Lorenz KA, Marcantonio ER, Alicante GK, et al. One-Year Mortality After Dialysis Initiation Among Older Adults. JAMA Intern Med. 2019;179(7):987-90.

24. Woodcock L. Diabetes care after pancreatic surgery. Journal of Diabetes Nursing. 2019;23:69.

25. Cui Y, Andersen DK. Pancreatogenic diabetes: special considerations for management. Pancreatology. 2011;11(3):279-94.

26. Vonderau JS, Desai CS. Type 3c: Understanding pancreatogenic diabetes. JAAPA. 2022;35(11):20-4.

27. NHS. Whipples Operation (Pylorus preserving pancreatico-duodenectomy) - Information for patients2022. Available from: http://flipbooks.leedsth.nhs.uk/LN003925.pdf.

28. Feletti A, Anglani M, Scarpa B, Schiavi F, Boaretto F, Zovato S, et al. Von Hippel-Lindau disease: an evaluation of natural history and functional disability. Neuro Oncol. 2016;18(7):1011-20.

29. Lammens CR, Bleiker EM, Verhoef S, Hes FJ, Ausems MG, Majoor-Krakauer D, et al. Psychosocial impact of Von Hippel-Lindau disease: levels and sources of distress. Clin Genet. 2010;77(5):483-91.

30. VHL UK / Ireland. VHL UK/Ireland Patient/Carer Survey2022. Available from: teamnice@vhl-uk-ireland.org.

31. Evans DG, Howard E, Giblin C, Clancy T, Spencer H, Huson SM, et al. Birth incidence and prevalence of tumor-prone syndromes: estimates from a UK family genetic register service. Am J Med Genet A. 2010;152A(2):327-32.

32. NORD. Von Hippel-Lindau Disease: National Organization for Rare Disorders; 2021 [updated 06-OCT-2021. Available from: https://rarediseases.org/rare-diseases/von-hippel-lindau-disease/?filter=Affected+Populations.

33. Maher ER, Iselius L, Yates JR, Littler M, Benjamin C, Harris R, et al. Von Hippel-Lindau disease: a genetic study. J Med Genet. 1991;28(7):443-7.

34. Jonasch E, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Epidemiology and Economic Burden of von Hippel-Lindau Disease-Associated Renal Cell Carcinoma in the United States. Clinical Genitourinary Cancer. 2022.

35. Jonasch E, Donskov F, Iliopoulos O, Rathmell WK, Narayan VK, Maughan BL, et al. Belzutifan for Renal Cell Carcinoma in von Hippel-Lindau Disease. N Engl J Med. 2021;385(22):2036-46.

36. Srinivasan R, Iliopoulos O, Rathmell WK, Narayan V, Maughan BL, Oudard S, et al. LBA69 Belzutifan, a HIF-2α Inhibitor, for von Hippel-Lindau (VHL) disease-associated neoplasms: 36 months of follow-up of the phase II LITESPARK-004 study. Annals of Oncology. 2022;33:S1433-S4.

37. MSD. MK-6482-004 Clinical Study Report. [MSD Data On File] [Internet]. 2022.

38. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

39. Hernandez-Alava MW, A.; Pudney, S.;. Methods for mapping between the EQ 5D 5L and the 3L2017. Available from: https://drive.google.com/file/d/1ehLw3-Mo7WyHFiwyijI4Ur5kCAPTEAPe/view.

40. clinicaltrials.gov. Belzutifan/MK-6482 for the Treatment of Advanced Pheochromocytoma/Paraganglioma (PPGL), Pancreatic Neuroendocrine Tumor (pNET), or Von Hippel-Lindau (VHL) Disease-Associated Tumors (MK-6482-015): U.S. National Library of Medicine; 2022 [updated 22-DEC-2022. Available from: https://clinicaltrials.gov/ct2/show/NCT04924075.

41. MSD. Estimating prevalence, treatment patterns, healthcare resource utilization (HRU), and economic burden of von Hippel-Lindau (VHL)-associated renal cell carcinoma (RCC). [MSD Data On File] [Internet]. 2022.

42. NICE. Metreleptin for treating lipodystrophy - HST14: National Institute for Health and Care Excellence; 2021 [updated 24-FEB-2021. Available from: https://www.nice.org.uk/guidance/HST14.

43. NICE. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer (CDF review of TA447) - TA531 - Final appraisal determination2018. Available from: https://www.nice.org.uk/guidance/ta531/documents/final-appraisal-determination-document.

44. MSD. VHL-Related Disease Impact on Quality of Life and Work Productivity in Patients with Renal Cell Carcinoma, Central Nervous System Hemangioblastoma, or Pancreatic Neuroendocrine Tumor Manifestation. [MSD Data On File] [Internet]. 2022.

45. NICE. Pembrolizumab for adjuvant treatment of renal cell carcinoma [ID3810] - [TA830] - Committee Papers. Single Technology Appraisal [Internet]. 2022. Available from: https://www.nice.org.uk/guidance/ta830/documents/committee-papers.

46. MSD. Von Hippel-Lindau Natural History Study - Final Study Report v1.0. [MSD Data On File] [Internet]. 2021.

47. Duffey BG, Choyke PL, Glenn G, Grubb RL, Venzon D, Linehan WM, et al. The relationship between renal tumor size and metastases in patients with von Hippel-Lindau disease. J Urol. 2004;172(1):63-5.

48. Lonser RR, Butman JA, Huntoon K, Asthagiri AR, Wu T, Bakhtian KD, et al. Prospective natural history study of central nervous system hemangioblastomas in von Hippel-Lindau disease. J Neurosurg. 2014;120(5):1055-62.

49. Wang L, Bensimon AG, Sundaram M, Xu R, Lai Y, Liu Y, et al. Burden of surgeries and surgical complications in patients with Von Hippel Lindau (VHL) disease before and after treatment with belzutifan. Journal of Clinical Oncology. 2023;41(6_suppl):733-.

50. Johnson A, Sudarshan S, Liu J, Linehan WM, Pinto PA, Bratslavsky G. Feasibility and outcomes of repeat partial nephrectomy. J Urol. 2008;180(1):89-93; discussion

51. Lonser RR, Weil RJ, Wanebo JE, DeVroom HL, Oldfield EH. Surgical management of spinal cord hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg. 2003;98(1):106-16.

52. Krauss T, Ferrara AM, Links TP, Wellner U, Bancos I, Kvachenyuk A, et al. Preventive medicine of von Hippel-Lindau disease-associated pancreatic neuroendocrine tumors. Endocr Relat Cancer. 2018;25(9):783-93.

53. NICE. Pembrolizumab for adjuvant treatment of renal cell carcinoma - [TA830] - Final Appraisal document2022. Available from:

https://www.nice.org.uk/guidance/ta830/documents/final-appraisal-determination-document.

54. Pavel M, Öberg K, Falconi M, Krenning EP, Sundin A, Perren A, et al. Gastroenteropancreatic neuroendocrine neoplasms: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2020;31(7):844-60.

55. Rini BI, Plimack ER, Stus V, Waddell T, Gafanov R, Pouliot F, et al. Pembrolizumab (pembro) plus axitinib (axi) versus sunitinib as first-line therapy for advanced clear cell renal cell carcinoma (ccRCC): Results from 42-month follow-up of KEYNOTE-426. Journal of Clinical Oncology. 2021;39(15_suppl):4500-.

56. Bensimon AG, Zhong Y, Swami U, Briggs A, Young J, Feng Y, et al. Cost-effectiveness of pembrolizumab with axitinib as first-line treatment for advanced renal cell carcinoma. Curr Med Res Opin. 2020;36(9):1507-17.

57. Riaz IB, He H, Ryu AJ, Siddiqi R, Naqvi SAA, Yao Y, et al. A Living, Interactive Systematic Review and Network Meta-analysis of First-line Treatment of Metastatic Renal Cell Carcinoma. Eur Urol. 2021;80(6):712-23.

58. Motzer R, Alekseev B, Rha S-Y, Porta C, Eto M, Powles T, et al. Lenvatinib plus Pembrolizumab or Everolimus for Advanced Renal Cell Carcinoma. New England Journal of Medicine. 2021;384(14):1289-300.

59. Sun W, Lipsitz S, Catalano P, Mailliard JA, Haller DG, Eastern Cooperative Oncology G. Phase II/III study of doxorubicin with fluorouracil compared with streptozocin with fluorouracil or dacarbazine in the treatment of advanced carcinoid tumors: Eastern Cooperative Oncology Group Study E1281. J Clin Oncol. 2005;23(22):4897-904.

60. Faivre S, Niccoli P, Castellano D, Valle JW, Hammel P, Raoul JL, et al. Sunitinib in pancreatic neuroendocrine tumors: updated progression-free survival and final overall survival from a phase III randomized study. Ann Oncol. 2017;28(2):339-43.

61. Kaderli RM, Spanjol M, Kollar A, Butikofer L, Gloy V, Dumont RA, et al. Therapeutic Options for Neuroendocrine Tumors: A Systematic Review and Network Meta-analysis. JAMA Oncol. 2019;5(4):480-9.

62. Schuhmacher P, Kim E, Hahn F, Sekula P, Jilg CA, Leiber C, et al. Growth characteristics and therapeutic decision markers in von Hippel-Lindau disease patients with renal cell carcinoma. Orphanet Journal of Rare Diseases. 2019;14(1):235.

63. NICE. NICE health technology evaluation topic selection: the manual [PMG37]2022. Available from: https://www.nice.org.uk/process/pmg37/chapter/about-this-guide.

64. Siller S, Szelenyi A, Herlitz L, Tonn JC, Zausinger S. Spinal cord hemangioblastomas: significance of intraoperative neurophysiological monitoring for resection and long-term outcome. J Neurosurg Spine. 2017;26(4):483-93.

65. Rochette C, Baumstarck K, Canoni-Zattara H, Abdullah AE, Figarella-Branger D, Pertuit M, et al. Psychological impact of von Hippel-Lindau genetic screening in patients with a previous history of hemangioblastoma of the central nervous system. J Psychosoc Oncol. 2018;36(5):624-34.

66. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health. 2010;13(5):509-18.

67. Efthymiadou O, Mossman J, Kanavos P. Health related quality of life aspects not captured by EQ-5D-5L: Results from an international survey of patients. Health Policy. 2019;123(2):159-65.

68. Kiebert GM, Green C, Murphy C, Mitchell JD, O'Brien M, Burrell A, et al. Patients' health-related quality of life and utilities associated with different stages of amyotrophic lateral sclerosis. J Neurol Sci. 2001;191(1-2):87-93.

69. Gurusamy KS, Riviere D, van Laarhoven CJH, Besselink M, Abu-Hilal M, Davidson BR, et al. Cost-effectiveness of laparoscopic versus open distal pancreatectomy for pancreatic cancer. PLoS One. 2017;12(12):e0189631.

70. Lee AJ, Morgan CL, Conway P, Currie CJ. Characterisation and comparison of health-related quality of life for patients with renal failure. Curr Med Res Opin. 2005;21(11):1777-83.

71. Simianu VV, Gaertner WB, Kuntz K, Kwaan MR, Lowry AC, Madoff RD, et al. Cost-effectiveness Evaluation of Laparoscopic Versus Robotic Minimally Invasive Colectomy. Ann Surg. 2020;272(2):334-41.

72. Gandhi SK, Jensen MM, Fox KM, Smolen L, Olsson AG, Paulsson T. Cost-effectiveness of rosuvastatin in comparison with generic atorvastatin and simvastatin in a Swedish population at high risk of cardiovascular events. Clinicoecon Outcomes Res. 2012;4:1-11.

73. NICE. Neuroendocrine tumours (metastatic, unresectable, progressive) - everolimus and sunitinib [ID858]: National Institute for Health and Care Excellence; 2017 [Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10024.

74. Nisula S, Vaara ST, Kaukonen KM, Reinikainen M, Koivisto SP, Inkinen O, et al. Six-month survival and quality of life of intensive care patients with acute kidney injury. Crit Care. 2013;17(5):R250.

75. Zargar M, McFarlane T, Chan KKW, Wong WWL. Cost-Effectiveness of Nivolumab in Recurrent Metastatic Head and Neck Squamous Cell Carcinoma. Oncologist. 2018;23(2):225-33.

76. Poole CD, Chambers C, Allsopp R, Currie CJ. Quality of life and health-related utility analysis of adults with moderate and severe atopic dermatitis treated with tacrolimus ointment vs. topical corticosteroids. J Eur Acad Dermatol Venereol. 2010;24(6):674-8.

77. Stevenson SM, Danzig MR, Ghandour RA, Deibert CM, Decastro GJ, Benson MC, et al. Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer. Urol Oncol. 2014;32(8):1172-7.

78. Sankar A, Dixon PR, Sivanathan L, Memtsoudis SG, de Almeida JR, Singh M. Cost-effectiveness Analysis of Preoperative Screening Strategies for Obstructive Sleep Apnea among Patients Undergoing Elective Inpatient Surgery. Anesthesiology. 2020;133(4):787-800.

79. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.

80. Wang CY, Pham PN, Kim S, Lingineni K, Schmidt S, Diaby V, et al. Predicting Cost-Effectiveness of Generic vs. Brand Dabigatran Using Pharmacometric Estimates Among Patients with Atrial Fibrillation in the United States. Clin Transl Sci. 2020;13(2):352-61.

81. Memeh K, Ruhle B, Vaghaiwalla T, Kaplan E, Keutgen X, Angelos P. Thyroidectomy for euthyroid patients with Hashimoto thyroiditis and persisting symptoms: A cost-effectiveness analysis. Surgery. 2021;169(1):7-13.

82. Verkleij ML, Heijnsdijk EAM, Bussé AML, Carr G, Goedegebure A, Mackey AR, et al. Cost-Effectiveness of Neonatal Hearing Screening Programs: A Micro-Simulation Modeling Analysis. Ear Hear. 2021;42(4):909-16.

83. NICE. Ocrelizumab for treating relapsing–remitting multiple sclerosis - Technology appraisal guidance [TA533]2018. Available from: https://www.nice.org.uk/guidance/ta533.

84. Turner D, Adams E, Boulton M, Harrison S, Khan N, Rose P, et al. Partners and close family members of long-term cancer survivors: health status, psychosocial well-being and unmet supportive care needs. Psychooncology. 2013;22(1):12-9.

85. Song JI, Shin DW, Choi JY, Kang J, Baek YJ, Mo HN, et al. Quality of life and mental health in the bereaved family members of patients with terminal cancer. Psychooncology. 2012;21(11):1158-66.

86. Jonasch E, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Healthcare resource utilization (HRU) and costs among patients with Von Hippel-Lindau disease (VHL)-associated renal cell carcinoma (RCC): A retrospective administrative claims analysis. Journal of Clinical Oncology. 2022;40(6_suppl):305-.

87. Sundaram M, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Real-world treatment patterns in von Hippel-Lindau (VHL) disease-associated renal cell carcinoma (RCC): Costs of tumor reduction procedures and their complications. Journal of Clinical Oncology. 2022;40(16_suppl):4539-.

88. NICE. Pembrolizumab for adjuvant treatment of completely resected stage 3 melanoma [TA766]. Technology appraisal guidance [Internet]. 2022. Available from: https://www.nice.org.uk/guidance/ta766.

89. NICE. Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma [TA169]. Technology appraisal guidance [Internet]. 2009. Available from: https://www.nice.org.uk/guidance/ta169.

90. NICE. Pazopanib for the first-line treatment of advanced renal cell carcinoma - Technology appraisal guidance [TA215]2013. Available from: https://www.nice.org.uk/guidance/ta215.

91. BNF. British National Formulary2022. Available from: https://bnf.nice.org.uk/.

92. NHS. Drugs and pharmaceutical electronic market information tool (eMIT)2022. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.

93. MSD. MSD KEYNOTE-426 Clinical Study Report. [MSD Data On File] [Internet]. 2018.

94. NICE. Tivozanib for treating advanced renal cell carcinoma - Technology appraisal guidance [TA512]2018. Available from: https://www.nice.org.uk/guidance/ta512.

95. NICE. Cabozantinib for untreated advanced renal cell carcinoma - Technology appraisal guidance [TA542]2018. Available from: https://www.nice.org.uk/guidance/TA542.

96. Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, et al. Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. New England Journal of Medicine. 2019;380(12):1103-15.

97. NICE. Nivolumab for previously treated advanced renal cell carcinoma - Technology appraisal guidance [TA417]2016. Available from: https://www.nice.org.uk/guidance/ta417.

98. NICE. Axitinib for treating advanced renal cell carcinoma after failure of prior systemic treatment [TA333]2015. Available from: https://www.nice.org.uk/guidance/ta333.

99. NICE. Cabozantinib for previously treated advanced renal cell carcinoma - Technology appraisal guidance [TA463]2017. Available from: https://www.nice.org.uk/guidance/ta463.

100. Motzer RJ, Hutson TE, Glen H, Michaelson MD, Molina A, Eisen T, et al. Lenvatinib, everolimus, and the combination in patients with metastatic renal cell carcinoma: a randomised, phase 2, open-label, multicentre trial. Lancet Oncol. 2015;16(15):1473-82.

101. NICE. Everolimus for the second-line treatment of advanced renal cell carcinoma - Technology appraisal guidance [TA219]2011. Available from: https://www.nice.org.uk/guidance/ta219.

102. NICE. Everolimus for advanced renal cell carcinoma after previous treatment - Technology appraisal guidance [TA432]2017. Available from: https://www.nice.org.uk/guidance/ta432.

103. Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med. 2007;356(22):2271-81.

104. Strosberg JR, Fine RL, Choi J, Nasir A, Coppola D, Chen DT, et al. First-line chemotherapy with capecitabine and temozolomide in patients with metastatic pancreatic endocrine carcinomas. Cancer. 2011;117(2):268-75.

105. Faiss S, Pape UF, Böhmig M, Dörffel Y, Mansmann U, Golder W, et al. Prospective, randomized, multicenter trial on the antiproliferative effect of lanreotide, interferon alfa, and their combination for therapy of metastatic neuroendocrine gastroenteropancreatic tumors--the International Lanreotide and Interferon Alfa Study Group. J Clin Oncol. 2003;21(14):2689-96.

106. Iwasa S, Morizane C, Okusaka T, Ueno H, Ikeda M, Kondo S, et al. Cisplatin and etoposide as first-line chemotherapy for poorly differentiated neuroendocrine carcinoma of the hepatobiliary tract and pancreas. Jpn J Clin Oncol. 2010;40(4):313-8.

107. Hentic O, Hammel P, Couvelard A, Rebours V, Zappa M, Palazzo M, et al. FOLFIRI regimen: an effective second-line chemotherapy after failure of etoposide-platinum combination in patients with neuroendocrine carcinomas grade 3. Endocr Relat Cancer. 2012;19(6):751-7.

108. Faure M, Niccoli P, Autret A, Cavaglione G, Mineur L, Raoul JL. Systemic chemotherapy with FOLFOX in metastatic grade 1/2 neuroendocrine cancer. Mol Clin Oncol. 2017;6(1):44-8.

109. NHS England. 2019/20 National Cost Collection Data Publication2021. Available from: https://www.england.nhs.uk/publication/2019-20-national-cost-collection-data-publication/.

110. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med. 2015;373(19):1803-13.

111. Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier S, et al. Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. Lancet Oncol. 2013;14(6):552-62.

112. Motzer RJ, Escudier B, Powles T, Scheffold C, Choueiri TK. Long-term follow-up of overall survival for cabozantinib versus everolimus in advanced renal cell carcinoma. Br J Cancer. 2018;118(9):1176-8.

113. Sternberg CN, Hawkins RE, Wagstaff J, Salman P, Mardiak J, Barrios CH, et al. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. Eur J Cancer. 2013;49(6):1287-96.

114. Hutson TE, Escudier B, Esteban E, Bjarnason GA, Lim HY, Pittman KB, et al. Randomized phase III trial of temsirolimus versus sorafenib as second-line therapy after sunitinib in patients with metastatic renal cell carcinoma. J Clin Oncol. 2014;32(8):760-7.

115. Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, Ou S-S, et al. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(33):5422-8.

116. Kanno H, Kobayashi N, Nakanowatari S. Pathological and Clinical Features and Management of Central Nervous System Hemangioblastomas in von Hippel-Lindau Disease. J Kidney Cancer VHL. 2014;1(4):46-55.

117. Maher ER, Neumann HP, Richard S. von Hippel-Lindau disease: a clinical and scientific review. Eur J Hum Genet. 2011;19(6):617-23.

118. NICE. Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer - Technology appraisal guidance [TA476]2017. Available from: https://www.nice.org.uk/guidance/ta476.

119. Georghiou TB, M;. Exploring the cost of care at the end of life2014. Available from: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life.

120. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrol Dial Transplant. 2012;27 Suppl 3(Suppl 3):iii73-80.

121. Coronini-Cronberg S, Appleby J, Thompson J. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. J R Soc Med. 2013;106(7):278-87.

122. Patel A, Berdunov V, Quayyum Z, King D, Knapp M, Wittenberg R. Estimated societal costs of stroke in the UK based on a discrete event simulation. Age Ageing. 2020;49(2):270-6.

123. Fineberg NA, Haddad PM, Carpenter L, Gannon B, Sharpe R, Young AH, et al. The size, burden and cost of disorders of the brain in the UK. J Psychopharmacol. 2013;27(9):761-70.

124. Moore A, Young CA, Hughes DA. Health Utilities and Costs for Motor Neurone Disease. Value in Health. 2019;22(11):1257-65.

125. Nataatmadja M, Krishnasamy R, Zuo L, Hong D, Smyth B, Jun M, et al. Quality of Life in Caregivers of Patients Randomized to Standard- Versus Extended-Hours Hemodialysis. Kidney Int Rep. 2021;6(4):1058-65.

126. NICE. Pembrolizumab for adjuvant treatment of resected stage 2B or 2C melanoma - Technology appraisal guidance [TA837]2022. Available from: https://www.nice.org.uk/guidance/ta837.

127. NICE. Atezolizumab for adjuvant treatment of resected non-small-cell lung cancer - Technology appraisal guidance [TA823]2022. Available from: https://www.nice.org.uk/guidance/ta823.

128. NICE. Pembrolizumab for neoadjuvant and adjuvant treatment of triple-negative early or locally advanced breast cancer - Technology appraisal guidance [TA851]2022. Available from: https://www.nice.org.uk/guidance/ta851.

129. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.

130. Catapano D, Muscarella LA, Guarnieri V, Zelante L, D'Angelo VA, D'Agruma L. Hemangioblastomas of central nervous system: molecular genetic analysis and clinical management. Neurosurgery. 2005;56(6):1215-21.

B.5 Appendices

Appendix C: Summary of product characteristics (SmPC) and UK public assessment report

C1.1 SmPC

Welireg SPC.pdf

C1.2 UK public assessment report

Welireg PAR.pdf

Appendix D: Identification, selection and synthesis of clinical evidence

D1.1 Identification and selection of relevant studies

To identify and select relevant studies, a systematic literature review (SLR) was carried out in accordance with NICE guidance, according to a protocol developed a priori, to identify relevant studies that investigated belzutifan and any relevant comparator treatments for the indication of interest for this appraisal as described in Table 1 of Document B section B1.1.

Specifically, the SLR was to gather clinical trial and observational evidence regarding the efficacy/effectiveness and safety of interventions in patients with Von Hippel-Lindau (VHL)-associated renal cell carcinoma (RCC) or central nervous system (CNS) hemangioblastoma, or pancreatic neuro-endocrine tumours (pNET) that have not metastasised.

Search strategy

Information sources

The following databases were searched:

- EMBASE (through the OVID portal)
- MEDLINE (through the OVID portal)
- Cochrane Registry of Controlled Trials (through the OVID portal)

This search used terms related to each intervention of interest, disease area, and study design of interest. The study design filters recommended by the Scottish Intercollegiate Guidelines Network (SIGN) for EMBASE and MEDLINE was used to identify studies (http://www.sign.ac.uk/methodology/filters.html). Where possible, databases were

searched with filters to allow identification of citations released online ahead of their print date.

Manual searches were conducted in clinicaltrials.gov to identify clinical trials that have not been published but are potentially eligible for inclusion. Furthermore, conference proceedings were searched to identify relevant evidence presented at the following meetings in the past 2 years:

- American Society of Clinical Oncology (ASCO)
- European Society for Medical Oncology (ESMO)

Search strategies

Embase

Table 105 Embase Search Strategy, Database(s): Embase 1974 to 2022 June 14, Searched on June 15, 2022

#	Searches	Results
1	exp renal cell carcinoma/	22608
2	((renal or kidney) adj2 cell adj2 (carcinoma or cancer* or cancer* or malignan* or tumor* or tumour* or neoplas* or adenocarcinoma*)).ti,ab.	67820
3	(renal cell cancer or RCC or renal cell carcinoma or kidney cancer or kidney carcinoma).ti,ab.	74534
4	exp hemangioblastoma/	3804
5	hemangioblastoma.mp.	4232
6	exp pancreatic neuroendocrine tumor/	20870
7	pancreatic neuroendocrine tumor.mp.	4301
8	or/1-7	110505
9	(von Hippel Lindau Disease or Familial Cerebello-Retinal Angiomatosis or Angiomatoses, Familial Cerebello-Retinal or Angiomatosis, Familial Cerebello-Retinal or Cerebello-Retinal Angiomatoses, Familial or Cerebello-Retinal Angiomatosis, Familial or Familial Cerebello Retinal Angiomatosis or Familial Cerebello-Retinal Angiomatoses or Hippel-Lindau Disease or Hippel Lindau Disease or VHL Syndrome or VHL Syndromes or Lindau's Disease or Lindau's Diseases or Lindaus Disease or von Hippel-Lindau Syndrome or von Hippel Lindau Syndrome or Angiomatosis Retinae or Cerebelloretinal Angiomatosis, Familial or Angiomatoses, Familial Cerebelloretinal or Angiomatosis, Familial Cerebelloretinal or Cerebelloretinal Angiomatoses, Familial or Familial Cerebelloretinal Angiomatoses or Familial Cerebelloretinal Angiomatosis or Lindau Disease).mp.	6172
10	8 and 9	3129

#	Searches	Results
11	Clinical Trial/	1036355
12	Randomized Controlled Trial/	712798
13	controlled clinical trial/	465795
14	multicenter study/	325984
15	Phase 3 clinical trial/	60943
16	Phase 4 clinical trial/	4789
17	exp RANDOMIZATION/	94272
18	Single Blind Procedure/	46421
19	Double Blind Procedure/	195772
20	Crossover Procedure/	70598
21	PLACEBO/	381471
22	randomi?ed controlled trial$.tw.	287356
23	rct.tw.	47112
24	(random$ adj2 allocat$).tw.	50221
25	single blind$.tw.	28996
26	double blind$.tw.	230877
27	((treble or triple) adj blind$).tw.	1578
28	placebo$.tw.	344115
29	Prospective Study/	771811
30	or/11-29	2699134
31	Case Study/	85976
32	case report.tw.	489145
33	abstract report/ or letter/	1243104
34	Editorial.pt.	729182
35	Note.pt.	897173
36	or/31-35	3425742
37	30 not 36	2566213
38	Clinical study/	158300
39	case control study/	188885
40	Family study/	25442
41	Longitudinal study/	173500
42	Retrospective study/	1257948
43	Prospective study/	771811
44	Cohort analysis/	851929
45	(Cohort adj (study or studies)).mp.	404942
46	(Case control adj (study or studies)).tw.	154932
47	(follow up adj (study or studies)).tw.	69465
48	(observational adj (study or studies)).tw.	218506
49	(epidemiologic$ adj (study or studies)).tw.	116194
50	(cross sectional adj (study or studies)).tw.	291507
51	or/38-50	3450023
52	37 or 51	5050019
53	10 and 52	461

MEDLINE

Table 106 MEDLINE Search Strategy, Database(s): Ovid MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions) 1946 June 14, 2022, Searched on June 15, 2022

#	Searches	Results
1	exp renal cell carcinoma/	37942
2	((renal or kidney) adj2 cell adj2 (carcinoma or cancer* or cancer* or malignan* or tumor* or tumour* or neoplas* or adenocarcinoma*)).ti,ab.	45444
3	(renal cell cancer or RCC or renal cell carcinoma or kidney cancer or kidney carcinoma).ti,ab.	48744
4	exp hemangioblastoma/	1460
5	hemangioblastoma.mp.	2379
6	pancreatic neuroendocrine tumor.mp.	1139
7	or/1-6	61105
8	(von Hippel Lindau Disease or Familial Cerebello-Retinal Angiomatosis or Angiomatoses, Familial Cerebello-Retinal or Angiomatosis, Familial Cerebello-Retinal or Cerebello-Retinal Angiomatoses, Familial or Cerebello-Retinal Angiomatosis, Familial or Familial Cerebello Retinal Angiomatosis or Familial Cerebello-Retinal Angiomatoses or Hippel-Lindau Disease or Hippel Lindau Disease or VHL Syndrome or VHL Syndromes or Lindau's Disease or Lindau's Diseases or Lindaus Disease or von Hippel-Lindau Syndrome or von Hippel Lindau Syndrome or Angiomatosis Retinae or Cerebelloretinal Angiomatosis, Familial or Angiomatoses, Familial Cerebelloretinal or Angiomatosis, Familial Cerebelloretinal or Cerebelloretinal Angiomatoses, Familial or Familial Cerebelloretinal Angiomatoses or Familial Cerebelloretinal Angiomatosis or Lindau Disease).mp.	3911
9	7 and 8	1819
10	Randomized Controlled Trials as Topic/	155975
11	randomized controlled trial/	570787
12	Random Allocation/	106855
13	Double Blind Method/	172108
14	Single Blind Method/	31996
15	clinical trial/	535368
16	clinical trial, phase i.pt.	23953
17	clinical trial, phase ii.pt.	38190
18	clinical trial, phase iii.pt.	20674
19	clinical trial, phase iv.pt.	2339
20	controlled clinical trial.pt.	94905
21	randomized controlled trial.pt.	570787
22	multicenter study.pt.	322481
23	clinical trial.pt.	535368
24	exp Clinical Trials as topic/	374859
25	or/10-24	1519265
26	(clinical adj trial$).tw.	439105
27	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.	189143
28	placebos/	35916

#	Searches	Results
29	placebo$.tw.	236501
30	randomly allocated.tw.	33737
31	(allocated adj2 random$).tw.	37342
32	or/26-31	731510
33	25 or 32	1832577
34	case report.tw.	364583
35	letter/	1183538
36	historical article/	368451
37	or/34-36	1898698
38	33 not 37	1791276
39	Epidemiologic studies/	9112
40	exp case control studies/	1329464
41	exp cohort studies/	2360728
42	Case control.tw.	143950
43	(cohort adj (study or studies)).tw.	275192
44	Cohort analy$.tw.	10372
45	(Follow up adj (study or studies)).tw.	53755
46	(observational adj (study or studies)).tw.	141179
47	Longitudinal.tw.	293938
48	Retrospective.tw.	666311
49	Cross sectional.tw.	453589
50	Cross-sectional studies/	429988
51	or/39-50	3541669
52	38 or 51	4854629
53	9 and 52	345

Cochrane

Table 107 Cochrane Central Register of Controlled Trials Search Strategy, Database(s): EBM Reviews - Cochrane Central Register of Controlled Trials May 2022, Searched on June 15, 2022

#	Searches	Results
1	exp Carcinoma, Renal Cell/	1056
2	((renal or kidney) adj2 cell adj2 (carcinoma or cancer* or cancer* or malignan* or tumor* or tumour* or neoplas* or adenocarcinoma*)).ti,ab.	3052
3	(renal cell cancer or RCC or renal cell carcinoma or kidney cancer or kidney carcinoma).ti,ab.	3433
4	exp hemangioblastoma/	1
5	hemangioblastoma.mp.	20
6	pancreatic neuroendocrine tumor.mp.	105
7	or/1-6	3725

#	Searches	Results
8	(von Hippel Lindau Disease or Familial Cerebello-Retinal Angiomatosis or Angiomatoses, Familial Cerebello-Retinal or Angiomatosis, Familial Cerebello-Retinal or Cerebello-Retinal Angiomatoses, Familial or Cerebello-Retinal Angiomatosis, Familial or Familial Cerebello Retinal Angiomatosis or Familial Cerebello-Retinal Angiomatoses or Hippel-Lindau Disease or Hippel Lindau Disease or VHL Syndrome or VHL Syndromes or Lindau's Disease or Lindau's Diseases or Lindaus Disease or von Hippel-Lindau Syndrome or von Hippel Lindau Syndrome or Angiomatosis Retinae or Cerebelloretinal Angiomatosis, Familial or Angiomatoses, Familial Cerebelloretinal or Angiomatosis, Familial Cerebelloretinal or Cerebelloretinal Angiomatoses, Familial or Familial Cerebelloretinal Angiomatoses or Familial Cerebelloretinal Angiomatosis or Lindau Disease).mp.	27
9	7 and 8	15

Study selection

Two reviewers, working independently, reviewed all abstracts and proceedings identified by the searches according to the PICOTS criteria (summarised in Table 108 in the next section), with the exception of outcomes criteria, which was only applied to the full-text selection. Studies identified as eligible during abstract screening and were then screened again by the same two reviewers by viewing the full-text versions of the study. Studies remaining eligible for inclusion after reviewing the full-text articles were then moved to data extraction. In each selection phase, the independent reviewers reconciled differences between them. A third reviewer was included to reach consensus on any discrepancies that were insolvable between the two reviewers. The process of study identification and selection was summarised with a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

While screening the studies the following types of studies were excluded:

- Studies reporting only surgical-related complications as an outcome, as surgical complications are not listed as outcomes of interest.
- Studies reporting only surgical-related death as an outcome

- Studies including a mixed population of VHL and non-VHL patients that did not report outcomes separately for the VHL subgroup

- Case studies and case series, studies in which outcomes are reported for each patient

Eligibility criteria

To support the research objectives, the SLR focused on identifying RCTs, controlled clinical trials, and other relevant scientific studies. The PICOTS (Population, Interventions, Comparison, Outcomes, Time, and Study Design) statement presented in Table 108 summarises the eligibility criteria for the SLR.

Table 108 PICOTS criteria for study inclusion

	Inclusion Criteria

Population	Inclusion criteria:
	• Patients with von Hippel-Lindau (VHL) disease who require therapy for associated
	  ○ Renal cell carcinoma (RCC)
	  ○ Central nervous system (CNS) hemangioblastomas
	  ○ Pancreatic neuroendocrine tumors (pNET)
	
	Exclusion criteria:
	• Patients with metastatic disease
	• Stage IV (M1)

Interventions	Any intervention studied in the population of interest including, but not limited to: • MK-6482 • Surgery
Comparisons	With respect to comparative evidence (e.g. randomized controlled trials [RCTs]), the following comparators are of interest: • Placebo or best supportive care • Any intervention of interest With respect to non-comparative evidence (e.g. 'single-arm' trials, observational cohorts), no comparator treatment is required.

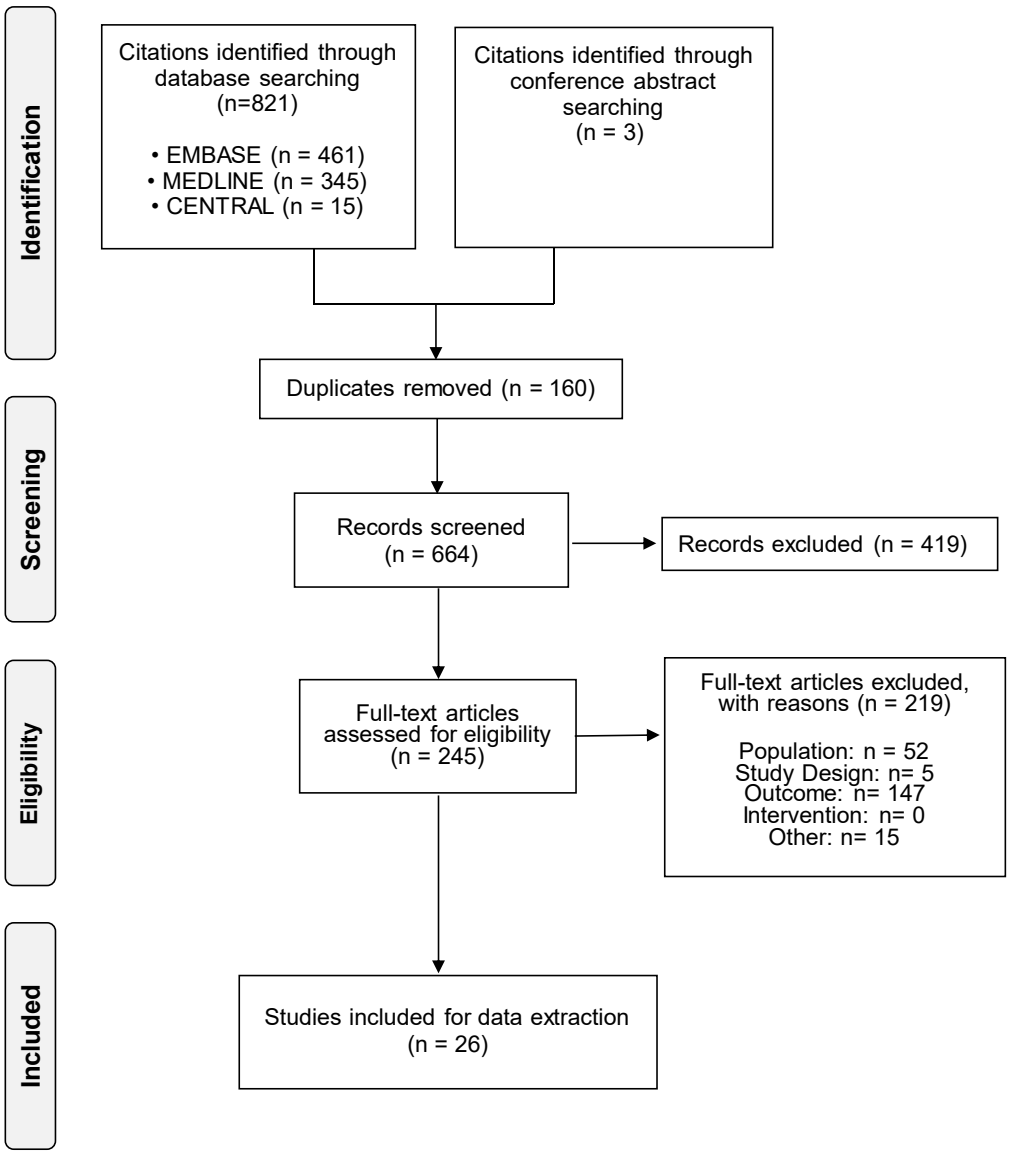
Outcomes	• Objective response rate (ORR) • Duration of response (DOR) • Time to response (TTR) • Time to surgery (TTS) • Progression-free survival (PFS) • Overall survival (OS) • Drug-related adverse events (AEs) • Grade 3-5 AEs (all, drug-related) • Discontinuation due to AE (DAEs) • Serious AEs (SAEs)
Time	Not Applicable
Study design	Inclusion criteria: • Randomised controlled trials (RCTs) • Controlled clinical trials • Non-randomized clinical trials, including single-arm prospective interventional trials • Prospective and retrospective cohort studies Exclusion criteria: • Case-control studies • Case series / Case reports • Cross-sectional studies • Systematic reviews and meta-analyses
Other	Only studies published in English language will be included.

Systematic literature review findings

The selection process to identify studies of interest is summarised with a PRISMA flow diagram (Figure 34). The literature search was performed on 15-JUN-2022. A total of 824 citations were identified by searching the bibliographic databases and conferences. After excluding 160 duplicates, a total of 664 citations were screened. This resulted in identification of 245 citations eligible for full-text review. Of the 245 full-text articles screened, 219 were excluded: 52 due to a population that was not of interest; five due to study design that was not of interest; 147 for no outcomes of interest and 15 for other reasons. Thus, a total of 26 citations representing 26 unique studies were initially included in this review.

However, only one of the 26 studies identified investigated the efficacy of belzutifan, specifically the MK-6482-004 study as reported in the Jonasch et al. 2021 publication (35), and so only that one study has been included for the purposes of this submission.

Figure 34 PRISMA flow diagram of study selection



Complete reference lists for included studies and excluded studies

List of citations initially included after full-text screening

The 26 citations initially included after full-text screening are listed in Table 109. Please note that only one of these 26 citations were for a trial that investigated the efficacy of belzutifan in the indication of relevance to this appraisal, specifically the MK-6482-004 trial reported in the Jonasch et al. 2021 citation (35), and so only that one trial has been included for the purposes of this submission.

Table 109 List of citations initially included at full-text screening phase

Citation
Asthagiri AR, Mehta GU, Zach L, et al. Prospective evaluation of radiosurgery for hemangioblastomas in von Hippel-Lindau disease. Neuro-Oncology. 2010;12(1):80-86.
Capitanio U, Rosiello G, Erdem S, et al. Clinical, surgical, pathological and follow-up features of kidney cancer patients with Von Hippel-Lindau syndrome: novel insights from a large consortium. World J Urol. 2021;39(8):2969-2975.
Chan VW, Lenton J, Smith J, et al. Multimodal image-guided ablation on management of renal cancer in Von-Hippel-Lindau syndrome patients from 2004 to 2021 at a specialist centre: A longitudinal observational study. Eur J Surg Oncol. 2022;48(3):672-679.
Chang SD, Meisel JA, Hancock SL, et al. Treatment of hemangioblastomas in von Hippel-Lindau disease with linear accelerator-based radiosurgery. Neurosurgery. 1998;43(1):28-34; discussion 34-25.
Cvek J, Knybel L, Reguli S, et al. Stereotactic radiotherapy for spinal hemangioblastoma - disease control and volume analysis in long-term follow up. Rep Pract Oncol Radiother. 2022;27(1):134-141.
Eggener SE, Rubenstein JN, Smith ND, et al. Renal tumors in young adults. J Urol. 2004;171(1):106-110.
Frydenberg M, Malek RS, Zincke H. Conservative renal surgery for renal cell carcinoma in von Hippel-Lindau's disease. J Urol. 1993;149(3):461-464.
Goldfarb DA, Neumann HP, Penn I, et al. Results of renal transplantation in patients with renal cell carcinoma and von Hippel-Lindau disease. Transplantation. 1997;64(12):1726-1729.
Jilg CA, Neumann HP, Glasker S, et al. Nephron sparing surgery in von Hippel-Lindau associated renal cell carcinoma; clinicopathological long-term follow-up. Fam Cancer. 2012;11(3):387-394.
Jonasch E, Donskov F, Iliopoulos O, et al. Belzutifan for Renal Cell Carcinoma in von Hippel–Lindau Disease. New England Journal of Medicine. 2021;385(22):2036-2046.
Jonasch E, McCutcheon IE, Waguespack SG, et al. Pilot trial of sunitinib therapy in patients with von Hippel-Lindau disease. Ann Oncol. 2011;22(12):2661-2666.
Jonasch E, McCutcheon, Gombos DS, et al. Pazopanib in patients with von Hippel-Lindau disease: a single-arm, single-centre, phase 2 trial. The Lancet Oncology. 2018;19(10):1351-1359.
Kano H, Niranjan A, Mongia S, et al. The role of stereotactic radiosurgery for intracranial hemangioblastomas. Neurosurgery. 2008;63(3):443-450; discussion 450-441.

Citation
Kano H, Shuto T, Iwai Y, et al. Stereotactic radiosurgery for intracranial hemangioblastomas: a retrospective international outcome study. J Neurosurg. 2015;122(6):1469-1478.
Koh ES, Nichol A, Millar BA, et al. Role of fractionated external beam radiotherapy in hemangioblastoma of the central nervous system. Int J Radiat Oncol Biol Phys. 2007;69(5):1521-1526.
Ma K, Hong B, Zhou J, et al. The Efficacy and Safety of Tyrosine Kinase Inhibitors for Von Hippel-Lindau Disease: A Retrospective Study of 32 Patients. Front Oncol. 2019;9:1122.
Morgan WR, Zincke H. Progression and survival after renal-conserving surgery for renal cell carcinoma: experience in 104 patients and extended followup. J Urol. 1990;144(4):852-857; discussion 857-858.
Persad RA, Probert JL, Sharma SD, et al. Surgical management of the renal manifestations of von Hippel-Lindau disease: a review of a United Kingdom case series. Br J Urol. 1997;80(3):392-396.
Ploussard G, Droupy S, Ferlicot S, et al. Local recurrence after nephron-sparing surgery in von Hippel-Lindau disease. Urology. 2007;70(3):435-439.
Roma A, Maruzzo M, Basso U, et al. First-Line sunitinib in patients with renal cell carcinoma (RCC) in von Hippel-Lindau (VHL) disease: clinical outcome and patterns of radiological response. Fam Cancer. 2015;14(2):309-316.
Roupret M, Hopirtean V, Mejean A, et al. Nephron sparing surgery for renal cell carcinoma and von Hippel-Lindau's disease: a single center experience. J Urol. 2003;170(5):1752-1755.
Simone CB, 2nd, Lonser RR, Ondos J, et al. Infratentorial craniospinal irradiation for von Hippel-Lindau: a retrospective study supporting a new treatment for patients with CNS hemangioblastomas. Neuro Oncol. 2011;13(9):1030-1036.
Steinbach F, Novick AC, Zinke H, et al. Treatment of renal cell carcinoma in von Hippel-Lindau disease: A multicenter study. J Urol. 1995;153(6):1812-1816.
Wessendorf J, König A, Heers H, et al. Repeat Percutaneous Radiofrequency Ablation of T1 Renal Cell Carcinomas is Safe in Patients with Von Hippel-Lindau Disease. Cardiovasc Intervent Radiol. 2021;44(12):2022-2025.
Yao M, Yoshida M, Kishida T, et al. VHL tumor suppressor gene alterations associated with good prognosis in sporadic clear-cell renal carcinoma. Journal of the National Cancer Institute. 2002;94(20):1569-1575.
Yousef A, Rutkowski MJ, Yalcin CE, et al. Sporadic and Von-Hippel Lindau disease-associated spinal hemangioblastomas: institutional experience on their similarities and differences. J Neurooncol. 2019;143(3):547-552.

List of studies initially excluded after full-text screening

Table 110 List of citations initially excluded at full-text screening phase

Study	Reason
A single-arm, phase II study of SU11248 (sunitinib) in patients with von Hippel-Lindau (VHL) disease. - VHLSUT. 2009.	Other
Albinana V, Escribano RMJ, Soler I, et al. Repurposing propranolol as a drug for the treatment of retinal haemangioblastomas in von Hippel-Lindau disease. Orphanet J Rare Dis. 2017;12(1):122.	Outcomes

Study	Reason
Allaf ME, Bhayani SB, Rogers C, et al. Laparoscopic partial nephrectomy: evaluation of long-term oncological outcome. J Urol. 2004;172(3):871-873.	Population
Allasia M, Soria F, Battaglia A, et al. Radiofrequency Ablation for Renal Cancer in Von Hippel-Lindau Syndrome Patients: A Prospective Cohort Analysis. Clinical Genitourinary Cancer. 2018;16(1):28-34.	Outcomes
Allasia M, Soria F, Battaglia A, et al. Radiofrequency Ablation for Renal Cancer in Von Hippel-Lindau Syndrome Patients: A Prospective Cohort Analysis. Clinical Genitourinary Cancer. 2018;16(1):28-34.	Outcomes
Alvarez R, Mastorakos P, Hogan E, et al. Retrobulbar Hemangioblastomas in von Hippel-Lindau Disease: Clinical Course and Management. Neurosurgery. 2021;88(5):1012-1020.	Population
Ammerman JM, Lonser RR, Dambrosia J, Butman JA, Oldfield EH. Long-term natural history of hemangioblastomas in patients with von Hippel-Lindau disease: implications for treatment. J Neurosurg. 2006;105(2):248-255.	Outcomes
Arnon L, Halperin R, Tirosh A. Impact of Pancreatic Neuroendocrine Tumor on Mortality in Patients With von Hippel-Lindau Disease. Endocrine Practice. 2021;27(10):1040-1045.	Outcomes
Asthagiri AR, Mehta GU, Butman JA, Baggenstos M, Oldfield EH, Lonser RR. Long-term stability after multilevel cervical laminectomy for spinal cord tumor resection in von Hippel-Lindau disease. J Neurosurg Spine. 2011;14(4):444-452.	Outcomes
Avci R, Yilmaz S, Inan UU, Kaderli B, Cevik SG. Vitreoretinal surgery for patients with severe exudative and proliferative manifestations of retinal capillary hemangioblastoma because of von hippel-lindau disease. Retina. 2017;37(4):782-788.	Outcomes
Baiocco JA, Ball MW, Pappajohn AK, et al. A comparison of outcomes for standard and multiplex partial nephrectomy in a solitary kidney: The National Cancer Institute experience. Urologic Oncology: Seminars and Original Investigations. 2019;37(6):356.e351-356.e357.	Population
Binderup ML, Budtz-Jorgensen E, Bisgaard ML. New von Hippel-Lindau manifestations develop at the same or decreased rates in pregnancy. Neurology. 2015;85(17):1500-1503.	Population
Binderup ML, Jensen AM, Budtz-Jorgensen E, Bisgaard ML. Survival and causes of death in patients with von Hippel-Lindau disease. J Med Genet. 2017;54(1):11-18.	Population
Bostrom A, Hans FJ, Reinacher PC, et al. Intramedullary hemangioblastomas: timing of surgery, microsurgical technique and follow-up in 23 patients. Eur Spine J. 2008;17(6):882-886.	Population
Bowen RC, Boldt HC, Mullins RF, et al. Intrafamilial Variability of Ocular Manifestations of von Hippel-Lindau Disease. Ophthalmology Retina. 2022;6(1):89-91.	Population
Brundl E, Schodel P, Ullrich OW, Brawanski A, Schebesch KM. Surgical resection of sporadic and hereditary hemangioblastoma: Our 10-year experience and a literature review. Surgical Neurology International. 2014;5(Supplement Supplement) (no pagination).	Population

Study	Reason
Byun J, Yoo HJ, Kim JH, et al. Growth rate and fate of untreated hemangioblastomas: clinical assessment of the experience of a single institution. Journal of Neuro-Oncology. 2019;144(1):147-154.	Outcomes
Carrion DM, Linares-Espinos E, Rios Gonzalez E, Bazan AA, Alvarez-Maestro M, Martinez-Pineiro L. Invasive management of renal cell carcinoma in von Hippel-Lindau disease. Cent. 2020;73(2):167-172.	Outcomes
Cervio A, Villalonga J, Mormandi R, Alcorta S, Sevlever G, Salvat J. Surgical treatment of cerebellar hemangioblastomas. Surgical Neurology International. 2017;8(1) (no pagination).	Population
Chauveau D, Duvic C, Chretien Y, et al. Renal involvement in von Hippel-Lindau disease. Kidney Int. 1996;50(3):944-951.	Study design
Chen X, Sanfilippo CJ, Nagiel A, et al. EARLY DETECTION of RETINAL HEMANGIOBLASTOMAS in von HIPPEL-LINDAU DISEASE USING ULTRA-WIDEFIELD FLUORESCEIN ANGIOGRAPHY. Retina. 2018;38(4):748-754.	Outcomes
Cheng J, Liu W, Hui X, Zhang S, Ju Y. Pediatric central nervous system hemangioblastomas: different from adult forms? A retrospective series of 25 cases. Acta Neurochir (Wien). 2017;159(9):1603-1611.	Outcomes
Cheng J, Liu W, Zhang S, Lei D, Hui X. Clinical Features and Surgical Outcomes in Patients with Cerebellopontine Angle Hemangioblastomas: Retrospective Series of 23 Cases. World Neurosurgery. 2017;103:248-256.	Outcomes
Chew EY. Ocular manifestations of von Hippel-Lindau disease: clinical and genetic investigations. Trans Am Ophthalmol Soc. 2005;103:495-511.	Study design
Chouliaras K, Newman NA, Shukla M, et al. Analysis of recurrence after the resection of pancreatic neuroendocrine tumors. Journal of Surgical Oncology. 2018;118(3):416-421.	Outcomes
Chretien Y, Chauveau D, Richard S, et al. [Treatment of von Hippel-Lindau disease with renal involvement]. Prog Urol. 1997;7(6):939-947.	Other
Clark AJ, Lu DC, Richardson RM, et al. Surgical technique of temporary arterial occlusion in the operative management of spinal hemangioblastomas. World Neurosurgery. 2010;74(1):200-205.	Outcomes
Coco D, Leanza S. Von Hippel-Lindau Syndrome: Medical Syndrome or Surgical Syndrome? A Surgical Perspective. J Kidney Cancer VHL. 2022;9(1):27-32.	Other
Coronel E, Gonzalez-Haba Ruiz M, Nielsen SM, et al. Pancreatic manifestations of von Hippel Lindau: A case series. American Journal of Gastroenterology. 2015;110(Supplement 1):S32.	Outcomes
Cristante L, Herrmann HD. Surgical management of intramedullary hemangioblastoma of the spinal cord. Acta Neurochir (Wien). 1999;141(4):333-339; discussion 339-340.	Outcomes
Cui H, Zou J, Bao YH, Wang MS, Wang Y. Surgical treatment of solid hemangioblastomas of the posterior fossa: A report of 28 cases. Oncology Letters. 2017;13(3):1125-1130.	Outcomes
Dal Bianco M, Artibani W, Bassi PF, Pescatori E, Pagano F. Prognostic factors in renal cell carcinoma. Eur Urol. 1988;15(1-2):73-76.	Population
Dalbah S, Bechrakis NE, Thomasen H, et al. Brachytherapy for Peripheral Retinal Capillary Haemangioblastoma in von Hippel-Lindau Disease. Klinische Monatsblatter fur Augenheilkunde. 2021;238(7):781-786.	Outcomes

Study	Reason
Dalvin LA, Yu MD, Ancona-Lezama DA, Pulido JS, Olsen TW, Shields CL. Retinal haemangioblastoma associated with peripheral non-perfusion: Widefield fluorescein angiography analysis of 41 cases. British Journal of Ophthalmology. 2020;104(2):167-172.	Population
Das JM, Kesavapisharady K, Sadasivam S, Nair SN. Microsurgical Treatment of Sporadic and von Hippel-Lindau Disease Associated Spinal Hemangioblastomas: A Single-Institution Experience. Asian spine j. 2017;11(4):548-555.	Outcomes
David P, Messerer M, Aghakhani N, et al. Intramedullary hemangioblastomas. Neurochirurgie. 2017;63(5):366-371.	Other
Deng X, Wang K, Wu L, et al. Intraspinal hemangioblastomas: Analysis of 92 cases in a single institution: Clinical article. Journal of Neurosurgery: Spine. 2014;21(2):260-269.	Outcomes
Dollfus H, Massin P, Taupin P, et al. Retinal hemangioblastoma in von Hippel-Lindau disease: A clinical and molecular study. Investigative Ophthalmology and Visual Science. 2002;43(9):3067-3074.	Outcomes
Dwarakanath S, Suri A, Sharma BS, Mehta VS. Intracranial hemangioblastomas: an institutional experience. Neurol India. 2006;54(3):276-278.	Outcomes
Egorov V, Petrov R, Beltsevich D. Organ-preserving and extensive pancreatic surgery for von HIPPELLINDAU disease. Six cases of 42 patients under surveillance. Pancreatology. 2018;18(4 Supplement 1):S135-S136.	Outcomes
Elborgy E, Pulido JS. Posterior Pole and Peripheral Retinal Fibrovascular Proliferation in von Hippel Lindau Disease. Asia-Pacific journal of ophthalmology (Philadelphia, Pa). 2017;6(3):256-260.	Outcomes
Eras M, Yenigun M, Acar C, Kumbasar B, Sar F, Bilge T. Pancreatic involvement in Von Hippel-Lindau disease. Indian J Cancer. 2004;41(4):159-161.	Other
Farrell MA, Charboneau WJ, DiMarco DS, et al. Imaging-guided radiofrequency ablation of solid renal tumors. American Journal of Roentgenology. 2003;180(6):1509-1513.	Population
Feletti A, Anglani M, Scarpa B, et al. Von Hippel-Lindau disease: An evaluation of natural history and functional disability. Neuro-Oncology. 2016;18(7):1011-1020.	Outcomes
Feletti A, Boaro A, Giampiccolo D, et al. Spinal hemangioblastomas: analysis of surgical outcome and prognostic factors. Neurosurgical Review. 2022;45(2):1645-1661.	Outcomes
Foubert F, Salimon M, Dumars C, et al. Survival and prognostic factors analysis of 151 intestinal and pancreatic neuroendocrine tumors: A single center experience. Journal of Gastrointestinal Oncology. 2019;10(1):103-111.	Population
Frantzen C, Kruizinga RC, van Asselt SJ, et al. Pregnancy-related hemangioblastoma progression and complications in von Hippel-Lindau disease. Neurology. 2012;79(8):793-796.	Outcomes
Frantzen C, Kruizinga RC, Van Asselt SJ, et al. Pregnancy stimulates cerebellar hemangioblastoma growth in von Hippel-Lindau disease. Pregnancy Hypertension. 2011;1(3-4):297.	Outcomes

Study	Reason
Ganesh HK, Gopal RA, George J, Bandgar T, Menon P, Shah N. Von Hippel-Lindau disease: A case series of unusual familial cancer syndrome. Endocrinologist. 2010;20(3):134-136.	Outcomes
Giammattei L, Messerer M, Aghakhani N, et al. Surgical resection of medulla oblongata hemangioblastomas: outcome and complications. Acta Neurochir (Wien). 2016;158(7):1333-1341.	Outcomes
Glasker S, Kruger MT, Klingler JH, et al. Hemangioblastomas and neurogenic polyglobulia. Neurosurgery. 2013;72(6):930-935; discussion 935.	Outcomes
Glasker S, Shah MJ, Hippchen B, Neumann HP, van Velthoven V. Doppler-sonographically guided resection of central nervous system hemangioblastomas. Neurosurgery. 2011;68(2 Suppl Operative):267-275; discussion 274-265.	Outcomes
Gonzalez-Rodriguez B, Villar Gomez de Las Heras K, Aguirre DT, et al. Evaluation of the safety and effectiveness of oral propranolol in patients with von Hippel-Lindau disease and retinal hemangioblastomas: phase III clinical trial. BMJ Open Ophthalmol. 2019;4(1):e000203.	Outcomes
Goyal N, Agrawal D, Singla R, Kale SS, Singh M, Sharma BS. Stereotactic radiosurgery in hemangioblastoma: Experience over 14 years. Journal of Neurosciences in Rural Practice. 2016;7(1):23-27.	Outcomes
Hajjaj A, Van Overdam K, Oldenburg R, De Klein A, Kilic E. Retinal hemangioblastomas in von Hippel-Lindau germline mutation carriers: Progression, complications and treatment outcome. Investigative Ophthalmology and Visual Science Conference. 2019;60(9).	Outcomes
Hajjaj A, van Overdam KA, Oldenburg RA, et al. Retinal haemangioblastomas in von Hippel-Lindau germline mutation carriers: progression, complications and treatment outcome. Acta Ophthalmologica. 2020;98(5):464-471.	Outcomes
Hanakita S, Koga T, Shin M, et al. The long-term outcomes of radiosurgery for intracranial hemangioblastomas. Neuro-Oncology. 2014;16(3):429-433.	Outcomes
Hasanov E, Jonasch E. MK-6482 as a potential treatment for von Hippel-Lindau disease-associated clear cell renal cell carcinoma. Expert Opinion on Investigational Drugs. 2021;30(5):495-504.	Population
Hasselblatt M, Jeibmann A, Gerss J, et al. Cellular and reticular variants of haemangioblastoma revisited: a clinicopathologic study of 88 cases. Neuropathol Appl Neurobiol. 2005;31(6):618-622.	Population
Hebbadj S, Cazzato RL, Garnon J, et al. Safety Considerations and Local Tumor Control Following Percutaneous Image-Guided Cryoablation of T1b Renal Tumors. CardioVascular and Interventional Radiology. 2018;41(3):449-458.	Population
Hes FJ, Slootweg PJ, van Vroonhoven TJ, et al. Management of renal cell carcinoma in von Hippel-Lindau disease. Eur J Clin Invest. 1999;29(1):68-75.	Outcomes
Hoang MP, Hruban RH, Albores-Saavedra J. Clear cell endocrine pancreatic tumor mimicking renal cell carcinoma: a distinctive neoplasm of von Hippel-Lindau disease. Am J Surg Pathol. 2001;25(5):602-609.	Population
Huntoon K, Lonser RR. Findings from the natural history of central nervous system hemangioblastomas in von Hippel-Lindau Disease. Clinical Neurosurgery. 2014;61(Supplement 1):159-162.	Outcomes

Study	Reason
Huntoon K, Wu T, Elder JB, et al. Biological and clinical impact of hemangioblastoma-associated peritumoral cysts in von Hippel-Lindau disease. J Neurosurg. 2016;124(4):971-976.	Outcomes
Hwang CK, Chew EY, Cukras CA, et al. Intravitreous treatment of severe ocular von Hippel-Lindau disease using a combination of the VEGF inhibitor, ranibizumab and PDGF inhibitor, E10030: Results from a phase 1/2 clinical trial. Clinical and Experimental Ophthalmology. 2021;49(9):1048-1059.	Population
Hwang JJU, E. M.;Pavlovich, C. P.;Pautler, S. E.;Libutti, S. K.;Linehan, W. M.;Walther, M. M. Surgical management of multi-organ visceral tumors in patients with von Hippel-Lindau disease: a single stage approach. J Urol. 2003;169(3):895-898.	Outcomes
Igarashi H, Ito T, Nishimori I, et al. Pancreatic involvement in Japanese patients with von Hippel-Lindau disease: Results of a nationwide survey. Journal of Gastroenterology. 2014;49(3):511-516.	Outcomes
Imagama S, Ito Z, Ando K, et al. Rapid Worsening of Symptoms and High Cell Proliferative Activity in Intra- and Extramedullary Spinal Hemangioblastoma. Global Spine Journal. 2017;7(1):6-13.	Outcomes
Iwamoto YK, H.;Yamakado, K.;Soga, N.;Arima, K.;Takeda, K.;Sugimura, Y. Management of renal tumors in Von Hippel-Lindau disease by percutaneous CT fluoroscopic guided radiofrequency ablation: preliminary results. Fam Cancer. 2011;10(3):529-534.	Outcomes
Jagannathan J, Lonser RR, Smith R, DeVroom HL, Oldfield EH. Surgical management of cerebellar hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg. 2008;108(2):210-222.	Outcomes
Jayanth ST, Mukherjee P, George AJP, et al. Outcomes of nephron sparing surgery and cortical sparing adrenalectomy in the management of Von Hippel-Lindau syndrome. African Journal of Urology. 2021;27(1) (no pagination).	Outcomes
Joaquim AF, Ghizoni E, dos Santos MJ, Valadares MG, da Silva FS, Tedeschi H. Intramedullary hemangioblastomas: surgical results in 16 patients. Neurosurg. 2015;39(2):E18.	Outcomes
Joly D, Mejean A, Correas JM, et al. Progress in nephron sparing therapy for renal cell carcinoma and von Hippel-Lindau disease. J Urol. 2011;185(6):2056-2060.	Population
Jonasch E, Donskov F, Iliopoulos O, et al. Phase II study of the oral HIF-2alpha inhibitor MK-6482 for Von Hippel-Lindau disease-associated renal cell carcinoma. Journal of Clinical Oncology Conference. 2020;38(15).	Other
Kanno H, Kuratsu J, Nishikawa R, et al. Clinical features of patients bearing central nervous system hemangioblastoma in von Hippel-Lindau disease. Acta Neurochir (Wien). 2013;155(1):1-7.	Outcomes
Kanno H, Yamamoto I, Nishikawa R, et al. Spinal cord hemangioblastomas in von Hippel-Lindau disease. Spinal Cord. 2009;47(6):447-452.	Outcomes
Kanno H, Kuratsu J, Nishikawa R, et al. Clinical features of patients bearing central nervous system hemangioblastoma in von Hippel-Lindau disease. Acta Neurochir (Wien). 2013;155(1):1-7.	Outcomes
Karabagli H, Genc A, Karabagli P, Abacioglu U, Seker A, Kilic T. Outcomes of gamma knife treatment for solid intracranial hemangioblastomas. J Clin Neurosci. 2010;17(6):706-710.	Outcomes

Study	Reason
Kataoka K, Ishikawa T, Ohno E, et al. Differentiation between pancreatic metastases from renal cell carcinoma and pancreatic neuroendocrine neoplasm using endoscopic ultrasound. Pancreatology. 2021;21(7):1364-1370.	Outcomes
Kelley RK, Van Cutsem E, Lee MS, et al. Phase 2 open-label study of pembrolizumab plus lenvatinib and belzutifan in patients with advanced solid tumors. Journal of Clinical Oncology Conference. 2022;40(4 SUPPL).	Population
Kim EH, Moon JH, Kang SG, Lee KS, Chang JH. Diagnostic challenges of posterior fossa hemangioblastomas: Refining current radiological classification scheme. Scientific reports. 2020;10(1):6267.	Outcomes
Kishi Y, Shimada K, Nara S, Esaki M, Hiraoka N, Kosuge T. Basing treatment strategy for non-functional pancreatic neuroendocrine tumors on tumor size. Annals of Surgical Oncology. 2014;21(9):2882-2888.	Population
Klingler JH, Steiert C, Glasker S, Kruger MT. Minimally invasive removal of spinal hemangioblastomas. Eur Spine J. 2019;28:2724.	Outcomes
Knickelbein JE, Jacobs-El N, Wong WT, et al. Systemic Sunitinib Malate Treatment for Advanced Juxtapapillary Retinal Hemangioblastomas Associated with von Hippel-Lindau Disease. Ophthalmology Retina. 2017;1(3):181-187.	Outcomes
Kobayashi N, Sato T, Kato S, et al. Imaging findings of pancreatic cystic lesions in von Hippel-Lindau disease. Internal Medicine. 2012;51(11):1301-1307.	Outcomes
Krivosic V, Kamami-Levy C, Jacob J, Richard S, Tadayoni R, Gaudric A. Laser Photocoagulation for Peripheral Retinal Capillary Hemangioblastoma in von Hippel-Lindau Disease. Ophthalmology Retina. 2017;1(1):59-67.	Outcomes
Kruger MT, Steiert C, Glasker S, Klingler JH. Minimally invasive resection of spinal hemangioblastoma: Feasibility and clinical results in a series of 18 patients. Journal of Neurosurgery: Spine. 2019;31(6):880-889.	Outcomes
Krzystolik K, Stopa M, Kuprjanowicz L, et al. Pars plana vitrectomy in advanced cases of von Hippel-Lindau eye disease. Retina. 2016;36(2):325-334.	Outcomes
Kumar P, Ravani R, Agarwal S, Dhanda S, Kumar V. Insights into retinal hemangioblastoma using ultra widefield imaging. Indian journal of ophthalmology. 2019;67(12):2029-2034.	Population
Kwon T, Jeong IG, Pak S, et al. Renal tumor size is an independent prognostic factor for overall survival in von Hippel-Lindau disease. Journal of Cancer Research and Clinical Oncology. 2014;140(7):1171-1177.	Outcomes
Le Reste PJ, Henaux PL, Morandi X, Carsin-Nicol B, Brassier G, Riffaud L. Sporadic intracranial haemangioblastomas: Surgical outcome in a single institution series. Acta Neurochir (Wien). 2013;155(6):1003-1009.	Population
Lee GJ, Jung TY, Kim IY, et al. The clinical experience of recurrent central nervous system hemangioblastomas. Clinical Neurology and Neurosurgery. 2014;123:90-95.	Outcomes
Lefevre A, Mathis T, Denis P, Kodjikian L. Retinal hemangioblastoma: Treatment strategy and long-term follow-up in a retrospective cohort. [French]. Journal francais d'ophtalmologie. 2018;41(2):164-169.	Outcomes
Levine E, Weigel JW, Collins DL. Diagnosis and management of asymptomatic renal cell carcinomas in von Hippel-Lindau syndrome. Urology. 1983;21(2):146-150.	Study design
Li X, Wang J, Niu J, Hong J, Feng Y. Diagnosis and microsurgical treatment of spinal hemangioblastoma. Neurological Sciences. 2016;37(6):899-906.	Outcomes

Study	Reason
Li Z, Zhang J, Zhang L, et al. Natural history and growth kinetics of clear cell renal cell carcinoma in sporadic and von Hippel-Lindau disease. Translational Andrology and Urology. 2021;10(3):1064-1070.	Outcomes
Liebenow B, Tatter A, Dezarn WA, Isom S, Chan MD, Tatter SB. Gamma Knife Stereotactic Radiosurgery favorably changes the clinical course of hemangioblastoma growth in von Hippel-Lindau and sporadic patients. Journal of Neuro-Oncology. 2019;142(3):471-478.	Outcomes
Lonser RR, Butman JA, Huntoon K, et al. Prospective natural history study of central nervous system hemangioblastomas in von Hippel-Lindau disease. J Neurosurg. 2014;120(5):1055-1062.	Outcomes
Lonser RR, Butman JA, Kiringoda R, Song D, Oldfield EH. Pituitary stalk hemangioblastomas in von Hippel-Lindau disease: Clinical article. J Neurosurg. 2009;110(2):350-353.	Outcomes
Lonser RR, Huntoon K, Butman JA, et al. Natural history of central nervous system hemangioblastomas in von hippel-lindau disease. Clinical Neurosurgery. 2013;1):168.	Outcomes
Lonser RR, Wait SD, Butman JA, et al. Surgical management of lumbosacral nerve root hemangioblastomas in von Hippel-Lindau syndrome. J Neurosurg. 2003;99(1 Suppl):64-69.	Outcomes
Lonser RR, Weil RJ, Wanebo JE, DeVroom HL, Oldfield EH. Surgical management of spinal cord hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg. 2003;98(1):106-116.	Outcomes
Lund GOF, B.;Curtis, M. A.;Williams, R. D. Conservative surgical therapy of localized renal cell carcinoma in von Hippel-Lindau disease. Cancer. 1994;74(9):2541-2545.	Outcomes
Ma D, Wang Y, Du G, Zhou L. Neurosurgical Management of Brainstem Hemangioblastomas: A Single-Institution Experience with 116 Patients. World Neurosurgery. 2015;84(4):1030-1038.	Outcomes
Ma K, Cai L, Ging K. Hazard of Death in Von Hippel-Lindau Disease Patients Depends on Genotype and Onset Age as Well as First Affected Lesion. J Urol. 2022;207(SUPPL 5):e343-e344.	Outcomes
Ma K, Gong K, Hong B, Xie H. The efficacy and safety of tyrosine kinase inhibitors in patients with von Hippel-Lindau disease. J Urol. 2019;201(4 Supplement 1):e345.	Other
Ma K, Li L, Cai L, Gong K. Risk factors for survival in patients with von hippel-lindau disease: A large single-center retrospective study of 734 vhl patients. J Urol. 2020;203(Supplement 4):e1230.	Population
Ma K, Zhang K, Cai L, Gong K. The expression level of PD-L1 is correlated with clinicopathological characteristics of patients with VHL-related renal cell carcinoma. J Urol. 2021;206(SUPPL 3):e772-e773.	Outcomes
Ma K, Zhang K, Zhou J, Cai L, Gong K. Analysis of risk factors for survival and prognosis of patients with von hippel-lindau syndrome: A large retrospective study. J Urol. 2021;206(SUPPL 3):e913-e914.	Outcomes
Ma KF, Li L, Kenan Z, Gong K, Cai L. PD-L1 expression was associated with aggressive clinicopathological features in patients with VHL-related renal cell carcinoma. European Urology Open Science. 2020;19(Supplement 2):e551-e552.	Outcomes

Study	Reason
Ma X, Tan Z, Zhang Q, et al. VHL Ser65 mutations enhance HIF2alpha signaling and promote epithelial-mesenchymal transition of renal cancer cells. Cell and Bioscience. 2022;12(1) (no pagination).	Outcomes
Ma XJ, Zhang GB, Guo TX, et al. Management and outcomes of pregnant patients with central nervous system hemangioblastoma. J Clin Neurosci. 2018;57:126-130.	Outcomes
Ma Z, Xie J, Zhao Y. [Hemangioblastomas of the central nervous system with von Hippel-Lindau's disease]. Chung Hua I Hsueh Tsa Chih. 1998;78(6):460-461.	Outcomes
Malek RS, Omess PJ, Benson RC, Jr., Zincke H. Renal cell carcinoma in von Hippel-Lindau syndrome. Am J Med. 1987;82(2):236-238.	Population
Margue G, Michiels C, Allenet C, et al. Feasibility of salvage robotic partial nephrectomy after ablative treatment failure (UroCCr-62 study). Minerva Urology and Nephrology. 2022;74(22):209-215.	Outcomes
Maroun FB, Green JS. Re: a case report of a family with 7 patients of the Von Hippel-Lindau disease (Violaris et al. Surg Neurol 2007;68:650-654). Surgical Neurology. 2009;71(2):261.	Outcomes
Matsunaga S, Shuto T, Inomori S, Fujino H, Yamamoto I. Gamma knife radiosurgery for intracranial haemangioblastomas. Acta Neurochir (Wien). 2007;149(10):1007-1013.	Outcomes
McDonald JD, Gupta S, Shindorf ML, et al. Pancreatic insufficiency following pancreatectomy: Does underlying tumor syndrome confer a greater risk? American Journal of Surgery. 2021;221(2):465-471.	Outcomes
Mehta GU, Asthagiri AR, Bakhtian KD, Auh S, Oldfield EH, Lonser RR. Functional outcome after resection of spinal cord hemangioblastomas associated with von Hippel-Lindau disease. J Neurosurg Spine. 2010;12(3):233-242.	Outcomes
Mehta GU, Montgomery BK, Maggio DM, Chittiboina P, Oldfield EH, Lonser RR. Functional outcome after resection of von Hippel-Lindau disease-associated Cauda equina hemangioblastomas: An observational cohort study. Operative Neurosurgery. 2017;13(4):435-440.	Outcomes
Meyerle CB, Dahr SS, Wetjen NM, et al. Clinical Course of Retrobulbar Hemangioblastomas in von Hippel-Lindau Disease. Ophthalmology. 2008;115(8):1382-1389.	Outcomes
Miller DL, Choyke PL, Walther MM, et al. von Hippel-Lindau disease: inadequacy of angiography for identification of renal cancers. Radiology. 1991;179(3):833-836.	Population
Minowada S, Homma Y, Takeuchi T, et al. [Surgical outcomes of nephron-sparing surgery for renal tumors]. Nippon Hinyokika Gakkai Zasshi. 2002;93(4):555-561.	Other
Moscovici S, Candanedo C, Spektor S, Cohen JE, Kaye AH. Solid vs. cystic predominance in posterior fossa hemangioblastomas: implications for cerebrovascular risks and patient outcome. Acta Neurochir (Wien). 2022;164(5):1357-1364.	Population
Moss JM, Choi CY, Adler JR, Jr., Soltys SG, Gibbs IC, Chang SD. Stereotactic radiosurgical treatment of cranial and spinal hemangioblastomas. Neurosurgery. 2009;65(1):79-85; discussion 85.	Outcomes

Study	Reason
Murro V, Lippera M, Mucciolo DP, et al. Outcome and genetic analysis of patients affected by retinal capillary hemangioblastoma in von hippel lindau syndrome. Molecular Vision. 2021;27:542-554.	Outcomes
Nakamura M, Ishii K, Watanabe K, et al. Surgical treatment of intramedullary spinal cord tumors: Prognosis and complications. Spinal Cord. 2008;46(4):282-286.	Population
Neumann HP, Eggert HR, Scheremet R, et al. Central nervous system lesions in von Hippel-Lindau syndrome. J Neurol Neurosurg Psychiatry. 1992;55(10):898-901.	Outcomes
Niemela M, Lemeta S, Sainio M, et al. Hemangioblastomas of the retina: impact of von Hippel-Lindau disease. Invest Ophthalmol Vis Sci. 2000;41(7):1909-1915.	Outcomes
Niemela M, Lemeta S, Summanen P, et al. Long-term prognosis of haemangioblastoma of the CNS: impact of von Hippel-Lindau disease. Acta Neurochir (Wien). 1999;141(11):1147-1156.	Population
Novick AC, Zincke H, Neves RJ, Topley HM. Surgical enucleation for renal cell carcinoma. J Urol. 1986;135(2):235-238.	Outcomes
Novick ACS, S. B. Long-term followup after nephron sparing surgery for renal cell carcinoma in von Hippel-Lindau disease. J Urol. 1992;147(6):1488-1490.	Population
O'Brien FJ, Danapal M, Jairam S, et al. Manifestations of von hippel lindau syndrome: A retrospective national review. Qjm. 2014;107(4):291-296.	Outcomes
Osaka K, Noguchi G, Kishida T, Yao M. Surgical management and outcomes of von Hippel-Lindau disease associated renal cell carcinoma: A single institution experience according to the 2 cm rule. European Urology, Supplements. 2019;18(11):e3533-e3534.	Outcomes
Oudard S, Elaidi R, Brizard M, et al. Sunitinib for the treatment of benign and malignant neoplasms from von Hippel-Lindau disease: A single-arm, prospective phase II clinical study from the PREDIR group. Oncotarget. 2016;7(51):85306-85317.	Population
Ozgen Saydam B, Adiyaman SC, Bozkurt O, et al. Pheochromocytoma: 16 years of experience in a single center. Turkish Journal of Endocrinology and Metabolism. 2021;25(1):54-67.	Other
Park BK, Kim CK. Percutaneous radio frequency ablation of renal tumors in patients with von Hippel-Lindau disease: preliminary results. J Urol. 2010;183(5):1703-1707.	Population
Park CH, Lee CH, Hyun SJ, Jahng TA, Kim HJ, Kim KJ. Surgical Outcome of Spinal Cord Hemangioblastomas. Journal of Korean Neurosurgical Society. 2012;52(3):221-227.	Population
Park SY, Park BK, Kim CK, et al. Percutaneous radiofrequency ablation of renal cell carcinomas in patients with von Hippel Lindau disease previously undergoing a radical nephrectomy or repeated nephron-sparing surgery. Acta radiologica (Stockholm, Sweden : 1987). 2011;52(6):680-685.	Population
Park TY, Lee SK, Park JS, et al. Clinical features of pancreatic involvement in von Hippel-Lindau disease: A retrospective study of 55 cases in a single center. Scandinavian Journal of Gastroenterology. 2014;50(3):360-367.	Outcomes
Park YS, Chang JH, Chang JW, Chung SS, Park YG. Gamma knife surgery for multiple hemangioblastomas. J Neurosurg. 2005;102 Suppl:97-101.	Population

Study	Reason
Parker F, Aghakhani N, Ducati LG, et al. Results of microsurgical treatment of medulla oblongata and spinal cord hemangioblastomas: a comparison of two distinct clinical patient groups. Journal of Neuro-Oncology. 2009;93(1):133-137.	Outcomes
Patrice SJ, Sneed PK, Flickinger JC, et al. Radiosurgery for hemangioblastoma: results of a multiinstitutional experience. Int J Radiat Oncol Biol Phys. 1996;35(3):493-499.	Outcomes
Pavesi G, Feletti A, Berlucchi S, et al. Neurosurgical treatment of von Hippel-Lindau-associated hemangioblastomas: benefits, risks and outcome. J Neurosurg Sci. 2008;52(2):29-36.	Outcomes
Peng XC, J.;Wang, J.;Peng, S.;Liu, S.;Ma, K.;Zhou, J.;Hong, B.;Zhou, B.;Zhang, J.;Cai, L.;Gong, K. Natural history of renal tumours in von Hippel-Lindau disease: A large retrospective study of Chinese patients. Journal of Medical Genetics. 2019.	Outcomes
Penitenti F, Landoni L, Scardoni M, et al. Clinical presentation, genotype-phenotype correlations, and outcome of pancreatic neuroendocrine tumors in Von Hippel-Lindau syndrome. Endocrine. 2021;74(1):180-187.	Outcomes
Peyre M, David P, Van Effenterre R, et al. Natural history of supratentorial hemangioblastomas in von Hippel-Lindau disease. Neurosurgery. 2010;67(3):577-587; discussion 587.	Outcomes
Pietila TA, Stendel R, Schilling A, Krznaric I, Brock M. Surgical treatment of spinal hemangioblastomas. Acta Neurochir (Wien). 2000;142(8):879-886.	Outcomes
Pilotto E, Midena G, Torresin T, et al. Retinal glial cells in von hippel-lindau disease: A novel approach in the pathophysiology of retinal hemangioblastoma. Cancers. 2022;14(1) (no pagination).	Outcomes
Pilotto E, Nacci EB, De Moja G, et al. Structural and microvascular changes of the peripapillary retinal nerve fiber layer in Von Hippel-Lindau disease: an OCT and OCT angiography study. Scientific reports. 2021;11(1):25.	Outcomes
Pilotto E, Nacci EB, Ferrara AM, et al. Macular flow in von Hippel-Lindau disease. Investigative Ophthalmology and Visual Science Conference. 2020;61(7).	Outcomes
Pinzi V, Viola A, De Martin E, et al. Radiosurgery for cranial and spinal haemangioblastomas: monoinstitutional analysis. Radiotherapy and Oncology. 2019;133(Supplement 1):***-***.	Population
Pitsika M, Pexas G, Joshi A, Mitchell P. Solid Component Volume as a Proxy to Identify Distinct Hemangioblastoma Populations. World Neurosurgery. 2021;146:e664-e669.	Outcomes
Pluta RM, Iuliano B, DeVroom HL, Nguyen T, Oldfield EH. Comparison of anterior and posterior surgical approaches in the treatment of ventral spinal hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg. 2003;98(1 SUPPL.):117-124.	Outcomes
Prasad R, Johnston LB, Savage MO, Martin L, Perry LA, Storr HL. Pediatric endocrine screening for von Hippel-Lindau disease: Benefits and the challenge of compliance. Journal of Endocrinological Investigation. 2011;34(4):296-299.	Outcomes
Prokopienko M, Kunert P, Podgorska A, Marchel A. Surgical treatment of sporadic and von Hippel-Lindau syndrome-associated intramedullary hemangioblastomas. Neurologia i Neurochirurgia Polska. 2016;50(5):349-355.	Outcomes

Study	Reason
Qiu J, Zhang K, Ma K, et al. The Genotype-Phenotype Association of Von Hipple Lindau Disease Based on Mutation Locations: A Retrospective Study of 577 Cases in a Chinese Population. Frontiers in Genetics. 2020;11 (no pagination).	Outcomes
Rajaraman C, Rowe JG, Walton L, Malik I, Radatz M, Kemeny AA. Treatment options for von Hippel-Lindau's haemangioblastomatosis: the role of gamma knife stereotactic radiosurgery. Br J Neurosurg. 2004;18(4):338-342.	Outcomes
Roy-Choudhury SH, Cast JE, Cooksey G, Puri S, Breen DJ. Early experience with percutaneous radiofrequency ablation of small solid renal masses. AJR Am J Roentgenol. 2003;180(4):1055-1061.	Population
Sadashivam S, Abraham M, Kesavapisharady K, Nair SN. Long-term outcome and prognostic factors of intramedullary spinal hemangioblastomas. Neurosurgical Review. 2020;43(1):169-175.	Outcomes
Saliou G, Giammattei L, Ozanne A, Messerer M. Role of preoperative embolization of intramedullary hemangioblastoma. Neurochirurgie. 2017;63(5):372-375.	Outcomes
Salome F, Colombeau P, Fermeaux V, et al. Renal lesions in Von Hippel-Lindau disease: The benign, the malignant, the unknown. Eur Urol. 1998;34(5):383-392.	Population
Sankaredja J, Brac B, Thines L, et al. Epidemiology, treatment and follow-up of central nervous system hemangioblastomas in von Hippel-Lindau disease. [French]. Revue Neurologique. 2014;170(4):288-296.	Other
Sayer FT, Nguyen J, Starke RM, Yen CP, Sheehan JP. Gamma knife radiosurgery for intracranial hemangioblastomas-outcome at 3 years. World Neurosurgery. 2011;75(1):99-105.	Outcomes
Schuhmacher P, Kim E, Hahn F, et al. Growth characteristics and therapeutic decision markers in von Hippel-Lindau disease patients with renal cell carcinoma. Orphanet J Rare Dis. 2019;14(1) (no pagination).	Outcomes
Selch MT, Tenn S, Agazaryan N, Lee SP, Gorgulho A, De Salles AAF. Image-guided linear accelerator-based spinal radiosurgery for hemangioblastoma. Surgical Neurology International. 2012;3(1) (no pagination).	Outcomes
Serban D, Exergian F. Intramedullary hemangioblastoma - Local experience of a tertiary clinic. Chirurgia (Romania). 2013;108(3):325-330.	Population
Shen H, Yang P, Liu Q, Tian Y. Nuclear expression and clinical significance of phosphohistidine phosphatase 1 in clear-cell renal cell carcinoma. Journal of International Medical Research. 2015;43(6):747-757.	Population
Shepard MJ, Bugarini A, Edwards NA, et al. Repurposing propranolol as an antitumor agent in von Hippel-Lindau disease. J Neurosurg. 2018:1-9.	Study design
Shuin T, Shinohara N, Yao M, Yamasaki I, Tamura K, Kamada M. [The current clinical status of kidney cancers in patients with the VHL disease in Japan: a nationwide epidemiological survey]. Nippon Hinyokika Gakkai Zasshi. 2012;103(3):552-556.	Other
Siller S, Szelenyi A, Herlitz L, Tonn JC, Zausinger S. Spinal cord hemangioblastomas: Significance of intraoperative neurophysiological monitoring for resection and long-term outcome. Journal of Neurosurgery: Spine. 2017;26(4):483-493.	Population

Study	Reason
Slater A, Moore NR, Huson SM. The natural history of cerebellar hemangioblastomas in von Hippel-Lindau disease. American Journal of Neuroradiology. 2003;24(8):1570-1574.	Outcomes
Smith GL, Kenney PA, Lee Y, Libertino JA. Non-clamped partial nephrectomy: techniques and surgical outcomes. BJU Int. 2011;107(7):1054-1058.	Population
Soczomski P, Jurecka-Lubieniecka B, Krzywon A, et al. A Direct Comparison of Patients With Hereditary and Sporadic Pancreatic Neuroendocrine Tumors: Evaluation of Clinical Course, Prognostic Factors and Genotype-Phenotype Correlations. Frontiers in Endocrinology. 2021;12 (no pagination).	Population
Song C. Unmoderated Posters. Urology. 2014;84(4):S171-S387.	Outcomes
Sorbellini M, Kattan MW, Snyder ME, et al. A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma. J Urol. 2005;173(1):48-51.	Outcomes
Srinivasan R, Donskov F, Iliopoulos O, et al. Phase II study of the oral HIF-2alpha inhibitor MK-6482 for Von Hippel-Lindau (VHL) disease-associated clear cell renal cell carcinoma (ccRCC): Update on RCC and non-RCC disease. Annals of Oncology. 2020;31(Supplement 4):S1158.	Other
Sultan A, Hassan T, Aboul-Enein H, Mansour O, Ibrahim T. The value of preoperative embolization in large and giant solid cerebellar hemangioblastomas. Interventional Neuroradiology. 2016;22(4):482-488.	Population
Sun HI, Ozduman K, Usseli MI, Ozgen S, Pamir MN. Sporadic spinal hemangioblastomas can be effectively treated by microsurgery alone. World neurosurgery. 2014;82(5):836-847.	Outcomes
Szalat A, Oleinikov K, Nahmias A, et al. Vhl-related neuroendocrine neoplasms and beyond: An Israeli specialized center real-life report. Endocrine Practice. 2020;26(10):1131-1142.	Outcomes
Tago M, Terahara A, Shin M, et al. Gamma knife surgery for hemangioblastomas. J Neurosurg. 2005;102 Suppl:171-174.	Outcomes
Takai K, Taniguchi M, Takahashi H, Usui M, Saito N. Comparative analysis of spinal hemangioblastomas in sporadic disease and Von Hippel-Lindau syndrome. Neurol Med Chir (Tokyo). 2010;50(7):560-567.	Outcomes
Takami H, Burns TC, Parney IF. Central nervous system hemangioblastoma: Difference in clinical picture of sporadic cases and von-hippel lindau disease in 184 cases. Journal of Neurological Surgery, Part B: Skull Base Conference: 29th Annual Meeting North American Skull Base Society Orlando, FL United States. 2019;80(Supplement 1).	Outcomes
Taylor SRJ, Singh J, Sagoo MS, Lightman SL. Clinical and molecular features associated with cystic visceral lesions in von hippel-lindau disease. Open Ophthalmology Journal. 2012;6:83-85.	Outcomes
Thrasher JB, Robertson JE, Paulson DF. Expanding indications for conservative renal surgery in renal cell carcinoma. Urology. 1994;43(2):160-168.	Population
Tordjman M, Dbjay J, Chamouni A, et al. Clear cell papillary renal cell carcinoma: A recent entity with distinct imaging patterns. American Journal of Roentgenology. 2020;214(3):579-587.	Population
Toy BC, Agron E, Nigam D, Chew EY, Wong WT. Longitudinal analysis of retinal hemangioblastomatosis and visual function in ocular von Hippel-Lindau disease. Ophthalmology. 2012;119(12):2622-2630.	Outcomes

Study	Reason
Uscanga-Yepez J, Rodriguez-Covarrubias F, Morales J, Granados J, Gabilondo-Navarro F. [Surgical treatment of renal-cell carcinoma in von Hippel-Lindau disease]. Rev Invest Clin. 2013;65(4):318-322.	Population
van Overdam KA, Missotten T, Kilic E, Spielberg LH. Early surgical treatment of retinal hemangioblastomas. Acta Ophthalmologica. 2017;95(1):97-102.	Population
Van Velthoven V, Reinacher PC, Klisch J, Neumann HP, Glasker S. Treatment of intramedullary hemangioblastomas, with special attention to von Hippel-Lindau disease. Neurosurgery. 2003;53(6):1306-1313; discussion 1313-1304.	Outcomes
Vergauwen E, Steiert C, Kruger MT, et al. Cumulative surgical morbidity in patients with multiple cerebellar and medullary hemangioblastomas. Clinical Neurology and Neurosurgery. 2020;197 (no pagination).	Outcomes
Violaris K, Siozos T, Skoulios N, Sakellariou P. A case report of a family with 7 patients of the Von Hippel-Lindau disease. Surgical Neurology. 2007;68(6):650-654.	Outcomes
Vougioukas VI, Glasker S, Hubbe U, et al. Surgical treatment of hemangioblastomas of the central nervous system in pediatric patients. Childs Nerv Syst. 2006;22(9):1149-1153.	Outcomes
Walther MM, Choyke PL, Hayes W, Shawker TH, Alexander RB, Linehan WM. Evaluation of color Doppler intraoperative ultrasound in parenchymal sparing renal surgery. J Urol. 1994;152(6 I):1984-1987.	Study design
Walther MM, Choyke PL, Weiss G, et al. Parenchymal sparing surgery in patients with hereditary renal cell carcinoma. J Urol. 1995;153(3 Pt 2):913-916.	Other
Wanebo JE, Lonser RR, Glenn GM, Oldfield EH. The natural history of hemangioblastomas of the central nervous system in patients with von Hippel-Lindau disease. J Neurosurg. 2003;98(1):82-94.	Outcomes
Wang EM, Pan L, Wang BJ, et al. The long-term results of gamma knife radiosurgery for hemangioblastomas of the brain. J Neurosurg. 2005;102 Suppl:225-229.	Outcomes
Wang EM, Wang BJ, Zhang N, et al. Analysis of the results of gamma knife radiosurgery for hemangioblastomas of the brain and the factors related to the tumor recurrence. [Chinese]. Zhonghua yi xue za zhi. 2004;84(10):813-817.	Other
Wang J, Qi F, Zhang P, et al. Clinical characteristics and genetic testing of an atypical familial von Hippel-Lindauzon renal cell carcinoma. Annals of Translational Medicine. 2019;7(22) (no pagination).	Outcomes
Wang J, Zhang L, Qiu J, et al. Natural history of Von Hippel-Lindau disease-associated and sporadic clear cell renal cell carcinoma: a comparative study. Journal of Cancer Research and Clinical Oncology. 2021.	Outcomes
Wang Q, Cheng J, Zhang S, Ju Y, Liu W, Hui X. Central nervous system hemangioblastomas in the elderly (over 65 years): Clinical characteristics and outcome analysis. Clinical Neurology and Neurosurgery. 2020;189 (no pagination).	Outcomes
Wang Q, Meng S, Cheng J, et al. Central nervous system hemangioblastomas: An age-stratified analysis. Clinical Neurology and Neurosurgery. 2020;199 (no pagination).	Outcomes
Wang Q, Zhang S, Cheng J, Liu W, Hui X. Radiologic Features and Surgical Strategy of Hemangioblastomas with Enhanced Cyst Wall. World Neurosurgery. 2017;108:143-150.	Outcomes

Study	Reason
Wang Y, Abu-Asab MS, Shen D, Zhuang Z, Chew EY, Chan CC. Upregulation of hypoxia-inducible factors and autophagy in von Hippel-Lindau-associated retinal hemangioblastoma. Graefe's Archive for Clinical and Experimental Ophthalmology. 2014;252(8):1319-1327.	Population
Weil RJ, Lonser RR, DeVroom HL, Wanebo JE, Oldfield EH. Surgical management of brainstem hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg. 2003;98(1):95-105.	Outcomes
Weisbrod A, Kitano M, Thomas F, et al. Pancreatic neuroendocrine tumors in von hippel lindau syndrome: An assessment of tumor growth and radiographic density. Annals of Surgical Oncology. 2013;1):S83.	Outcomes
Weisbrod AB, Kitano M, Thomas F, et al. Assessment of tumor growth in pancreatic neuroendocrine tumors in von Hippel Lindau syndrome. Journal of the American College of Surgeons. 2014;218(2):163-169.	Outcomes
Wiley H, Coleman HR, Jonasch E, et al. Oral HIF-2alpha inhibitor belzutifan for ocular von Hippel-Lindau (VHL) disease. Investigative Ophthalmology and Visual Science Conference: Annual Meeting Association for Research in Vision and Ophthalmology, ARVO. 2021;62(8).	Other
Wind JJ, Bakhtian KD, Sweet JA, et al. Long-term outcome after resection of brainstem hemangioblastomas in von Hippel-Lindau disease. J Neurosurg. 2011;114(5):1312-1318.	Outcomes
Wong WT, Agron E, Coleman HR, et al. Clinical Characterization of Retinal Capillary Hemangioblastomas in a Large Population of Patients with von Hippel-Lindau Disease. Ophthalmology. 2008;115(1):181-188.	Outcomes
Wong WT, Liang KJ, Hammel K, Coleman HR, Chew EY. Intravitreal Ranibizumab Therapy for Retinal Capillary Hemangioblastoma Related to von Hippel-Lindau Disease. Ophthalmology. 2008;115(11):1957-1964.e1953.	Population
Yano M, Misra S, Carpenter DH, Salter A, Hildebolt CF. Pancreatic Neuroendocrine Tumors: Computed Tomography Enhancement, but Not Histological Grade, Correlates with Tumor Aggression. Pancreas. 2017;46(10):1366-1372.	Population
Ye DY, Bakhtian KD, Asthagiri AR, Lonser RR. Effect of pregnancy on hemangioblastoma development and progression in von Hippel-Lindau disease. J Neurosurg. 2012;117(5):818-824.	Outcomes
Zhang J, Pan JH, Dong BJ, Xue W, Liu DM, Huang YR. Active surveillance of renal masses in von Hippel-Lindau disease: growth rates and clinical outcome over a median follow-up period of 56 months. Fam Cancer. 2012;11(2):209-214.	Outcomes
Zhang L, Wang H, Nan Y, Ma Q. Spinal hemangioblastoma: surgical procedures, outcomes and review of the literature. Acta Neurologica Belgica. 2021;121(4):973-981.	Population
Zhou J, Li NY, Zhou XJ, et al. [Clinicopathologic study of von Hippel-Lindau syndrome-related and sporadic hemangioblastomas of central nervous system]. Chung-hua Ping Li Hsueh Tsa Chih. 2010;39(3):145-150.	Outcomes
Zhou LF, Du G, Mao Y, Zhang R. Diagnosis and surgical treatment of brainstem hemangioblastomas. Surgical Neurology. 2005;63(4):307-315.	Outcomes
Zibly Z, Cohen ZR, Peled A, et al. Linear accelerator stereotactic radiosurgery can modulate the clinical course of Hemangioblastoma: Case series and review of the literature. J Clin Neurosci. 2020;82(Pt A):162-165.	Outcomes

Study	Reason
Zimmermann M, Seifert V, Schreyer F, Stolke D, Dietz H. Das Hämangioblastom: Beschreibung des Krankheitsbildes und Bericht über 41 Fälle [Hemangioblastoma: description of a disease picture and report of 41 cases]. Zentralbl Neurochir. 1997;58(1):1-6. German. PMID: 9235816.	Outcomes

Summary of trials used for indirect or mixed treatment comparisons

No indirect or mixed treatment comparisons were conducted as part of this submission.

Methods and outcomes of studies included in indirect or mixed treatment comparisons

Not applicable, no indirect or mixed treatment comparisons were conducted as part of this submission.

Methods of analysis of studies included in the indirect or mixed treatment comparison

Not applicable, no indirect or mixed treatment comparisons were conducted as part of this submission.

Programming language for the indirect or mixed treatment comparison

Not applicable, no indirect or mixed treatment comparisons were conducted as part of this submission.

Risk of bias of studies included in indirect or mixed treatment comparisons

Not applicable, no indirect or mixed treatment comparisons were conducted as part of this submission.

D1.2 Participant flow in the relevant randomised control trials

Participant flow data from the 15-JUL-2021 database cut-off date of the MK-6482-004 study are presented in this section (and also presented in Document B section B.2.6).

MK-6482-004 study patient disposition and follow-up, 15-JUL-2021 database cut-off date

A total of 61 participants were allocated and received at least 1 dose of belzutifan. As of the 15-JUL-2021 database cut-off date, 50 participants (82.0%) were receiving belzutifan, 11 participants (18.0%) had discontinued belzutifan, and 3 participants (4.9%) had discontinued from the study (Table 12). The primary reasons for discontinuation of belzutifan were progressive disease and participant decision. The median duration of follow-up among the 61 participants with RCC in the APaT population was 29.2 months (range: 4.2 to 37.5 months) (Table 13).

Table 111 MK-6482-004 summary of patient disposition (safety analysis set)

	Belzutifan (N = 61) n (%)
• Treatment Ongoing at Data Cut-Off Date	50 (82.0)
• Discontinued Treatment	11 (18.0)
Reason for Treatment Discontinuation	
• Disease progression per RECIST 1.1 for VHL disease-associated RCC tumours	4 (6.6)
• Disease progression due to symptomatic deterioration of the patient's health status	0
• Adverse event that in the opinion of the investigator or medical monitor would lead to undue risk if study treatment were continued	2 (3.3)
• Study drug interruption for more than 3 consecutive weeks due to a grade 3-4 or intolerable toxicity that is attributed to study drug	0
• Gross noncompliance with protocol	0
• Pregnancy in a female patient during the study	0
• Death	1 (1.6)
• Lost to follow-up	0
• Patient decision to discontinue study drug	4 (6.6)
• Sponsor discontinuation of study	0
• Other	0
• On Study at Data Cut-Off Date [1]	58 (95.1)
• Off Study	3 (4.9)
Reason for Study Discontinuation	
• Death	1 (1.6)
• Informed Consent Withdrawn	2 (3.3)
• Lost To Follow-up	0
• Sponsor discontinuation of study	0

	Belzutifan (N = 61) n (%)
• Other	0

[1] Patients are still on study treatment or in long term follow-up as of the cutoff date.
Date of Data Cut-off: 15JUL2021

Table 112 MK-6482-004 summary of follow-up duration (safety analysis set)

Follow-up duration (months)	Belzutifan (N = 61)
Median (Range)	29.2 (4.2-37.5)
Mean (SD)	29.7 (4.08)

Follow-up duration is defined as the time from first dose to the date of death or the database cutoff date if the subject is still alive.
Date of Data Cut-off: 15JUL2021

D1.3 Critical appraisal for each study

The quality of RCTs and non-randomised trials was assessed using the Cochrane Collaboration Risk of Bias (RoB) tool (129). This tool assessed risk of bias for the following six domains:

- Sequence generation
- Allocation concealment
- Blinding of participants, personnel and outcome assessors
- Incomplete outcome data
- Selective outcome reporting
- Other sources of bias

The RoB tool was used to assign summary assessments of within-study bias; low risk of bias (low risk of bias for all key domains), unclear risk of bias (unclear risk of bias for one or more key domains), or high-risk of bias (high-risk of bias for one or more key

domains). Study quality assessment of the MK-6482-004 (Jonasch et al. 2021 (35)) study via the RoB tool is summarised in Table 113.

Table 113 Cochrane risk of bias assessment for the study included from the systematic literature review

Study	MK-6482-004 (35)
NCT number	NCT03401788
Sequence generation	High risk
Allocation concealment	High risk
Blinding of participants, personnel and outcome assessors	High risk
Incomplete outcome reporting	Low risk
Selective outcome reporting	Low risk
Other sources of bias	High risk

Appendix E: Subgroup analysis

Relevant subgroup analyses are presented in Document B section B.2.7.

Appendix F: Adverse reactions

MK-6482-004 study

Details of the adverse events reported in the MK-6482-004 study (01-APR-2022 database cut-off date) are reported below.

Most frequently reported adverse events

[redacted] in the Safety Analysis Set (APaT population) had at least 1 AE. The most frequently reported AEs (in >25% of participants) were [redacted], as summarised in Table 114.

Table 114 MK-6482-004 patients with adverse events by decreasing incidence (incidence ≥10%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Anaemia	[redacted]
Fatigue	[redacted]
Headache	[redacted]
Dizziness	[redacted]
Nausea	[redacted]
Dyspnoea	[redacted]
Myalgia	[redacted]
Constipation	[redacted]
Arthralgia	[redacted]
Vision blurred	[redacted]
Abdominal pain	[redacted]
Alanine aminotransferase increased	[redacted]
Back pain	[redacted]
Diarrhoea	[redacted]
Upper respiratory tract infection	[redacted]
Weight increased	[redacted]
Hypertension	[redacted]
Insomnia	[redacted]
Oedema peripheral	[redacted]
COVID-19	[redacted]
Disturbance in attention	[redacted]
Urinary tract infection	[redacted]
Anxiety	[redacted]

Aspartate aminotransferase increased		[redacted]	
Blood creatinine increased		[redacted]	
Cough		[redacted]	
Muscle spasms		[redacted]	
Vomiting		[redacted]	

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Drug-related adverse events

[redacted] reported 1 or more AEs considered related to belzutifan by the investigator. Most drug-related AEs were Grade 1 and Grade 2 in severity. [redacted] reported drug-related AEs with CTCAE Grade 3 and above (described in Table 115). [redacted] reported a drug-related Grade 5 AE.

Table 115 MK-6482-004 patients with drug-related grade 3-5 by decreasing incidence (incidence ≥0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Anaemia	[redacted]
Fatigue	[redacted]
Hypoxia	[redacted]
Urinary tract infection	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Grade 3 to 5 adverse events

[redacted] reported 1 or more Grade 3 to 5 AEs. The most frequently reported Grade 3 to 5 AEs were [redacted]

[redacted] (Table 116). All other Grade 3 to 5 AEs were reported for [redacted].

Table 116 MK-6482-004 patients with grade 3-5 by decreasing incidence (incidence ≥0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Anaemia	[redacted]
Hypertension	[redacted]
Fatigue	[redacted]
Syncope	[redacted]
Anaphylactic reaction	[redacted]
Azoospermia	[redacted]
COVID-19	[redacted]
COVID-19 pneumonia	[redacted]
Cholecystectomy	[redacted]
Coronary artery dissection	[redacted]
Diarrhoea	[redacted]
Dyspnoea	[redacted]
Embolism	[redacted]
Haemorrhage intracranial	[redacted]
Hyperglycaemia	[redacted]
Hypotension	[redacted]
Hypoxia	[redacted]
Musculoskeletal pain	[redacted]
Myalgia	[redacted]
Non-small cell lung cancer	[redacted]
Otitis media chronic	[redacted]
Pneumonia	[redacted]
Retinal detachment	[redacted]
Retinal vein occlusion	[redacted]
Skin laceration	[redacted]
Suicide attempt	[redacted]
Toxicity to various agents	[redacted]
Urinary tract infection	[redacted]
Vitreous haemorrhage	[redacted]
Weight increased	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Deaths due to adverse events

[redacted] due to AE occurred during the study. [redacted].

Serious adverse events

[redacted] reported 1 or more SAEs (Table 117). The most frequently reported SAE were [redacted]

[redacted] reported an SAE that was considered related to belzutifan by the investigator. [redacted] (Table 118).

Table 117 MK-6482-004 patients with serious adverse events by decreasing incidence (incidence >0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Embolism	[redacted]
Haemorrhage intracranial	[redacted]
Abdominal pain	[redacted]
Anaemia	[redacted]
Anaphylactic reaction	[redacted]
COVID-19	[redacted]
COVID-19 pneumonia	[redacted]
Cholecystectomy	[redacted]
Coronary artery dissection	[redacted]
Cystitis	[redacted]
Dyspnoea	[redacted]
Hypertension	[redacted]
Hypotension	[redacted]
Hypoxia	[redacted]
Non-small cell lung cancer	[redacted]
Pneumonia	[redacted]

Retinal detachment	[redacted]
Retinal vein occlusion	[redacted]
Seizure	[redacted]
Skin laceration	[redacted]
Suicide attempt	[redacted]
Toxicity to various agents	[redacted]
Urinary tract infection	[redacted]
Vitreous haemorrhage	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Table 118 MK-6482-004 patients with drug-related serious adverse events by decreasing incidence (incidence >0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n (%)
61	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Anaemia	[redacted]
Haemorrhage intracranial	[redacted]
Hypoxia	[redacted]
Urinary tract infection	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Adverse events leading to discontinuation of study treatment

A total of [redacted] discontinued belzutifan due to the following AEs, which occurred in [redacted] each: [redacted] 119[redacted].

Table 119 MK-6482-004 adverse events leading to study drug discontinuation

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n (%)
Subjects with any Adverse Events Leading to Study Drug Discontinuation	[redacted]

Nervous system disorders	[redacted]
• Dizziness	[redacted]
• Haemorrhage intracranial	[redacted]
Injury, poisoning and procedural complications	[redacted]
• Toxicity to various agents	[redacted]
Psychiatric disorders	[redacted]
• Suicide attempt	[redacted]

Adverse events up to 30 days of last dose are included. Subject with two or more adverse events in the same system organ class (or with the same preferred term) is counted only once for that system organ class (or preferred term). Adverse Events were coded using MedDRA version 25.0. Uncoded preferred terms are presented in their verbatim terms.
Date of Data Cut-off: 01APR2022

Adverse events resulting in treatment interruption

A total of [redacted] experienced 1 or more AEs leading to interruption of belzutifan (Table 120). The most frequently reported AEs leading to interruption of belzutifan included [redacted].

[redacted] reported AEs leading to interruption of belzutifan that were considered drug-related by the investigator including [redacted] (Table 121).

Table 120 MK-6482-004 patients with adverse events resulting in treatment interruption by decreasing incidence (incidence >0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Fatigue	[redacted]
Nausea	[redacted]
Headache	[redacted]
Dizziness	[redacted]
Influenza like illness	[redacted]
Abdominal pain	[redacted]
Anaemia	[redacted]
COVID-19	[redacted]
Haemorrhage intracranial	[redacted]
Syncope	[redacted]
Vomiting	[redacted]
Arthralgia	[redacted]
COVID-19 pneumonia	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Cellulitis	[redacted]
Cholecystectomy	[redacted]
Cystitis	[redacted]
Diarrhoea	[redacted]
Dyspepsia	[redacted]
Embolism	[redacted]
Hypersensitivity	[redacted]
Nasal congestion	[redacted]
Pericardial effusion	[redacted]
Pyrexia	[redacted]
Retinal detachment	[redacted]
Retinal vein occlusion	[redacted]
Sensation of foreign body	[redacted]
Skin laceration	[redacted]
Tremor	[redacted]
Upper respiratory tract infection	[redacted]
Upper-airway cough syndrome	[redacted]
Urinary tract infection	[redacted]
Vertigo	[redacted]
Viral infection	[redacted]
Vision blurred	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Table 121 MK-6482-004 patients with drug-related adverse events resulting in treatment interruption by decreasing incidence (incidence >0%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more adverse events	[redacted]
with no adverse event	[redacted]
Fatigue	[redacted]
Nausea	[redacted]
Anaemia	[redacted]
Headache	[redacted]
Abdominal pain	[redacted]
Arthralgia	[redacted]
Dizziness	[redacted]
Influenza like illness	[redacted]
Sensation of foreign body	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Urinary tract infection	[redacted]
Vertigo	[redacted]
Vomiting	[redacted]

Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
Date of Data Cut-off: 01APR2022

Appendix G: Published cost-effectiveness studies

G1.1 Objectives

This SLR was conducted to understand the economic burden and summarize economic evaluations in patients with VHL disease-associated disease with a focus on VHL-associated RCC, CNS Hb and pNET.

G1.2 Systematic review methodology

Identification of studies

In July 2020, a comprehensive search was conducted including biomedical databases, conference proceedings, and other online resources. Details of the search strategy are provided in **Table 123-Table 125**. An update to the searches was conducted in July 2022. Details of the search strategy are provided in **Table 126**.

Information sources

Biomedical databases

The databases listed in Table 122 were searched.

Table 122: Databases Searched for the Literature Review and the Search Platform

Data sources	Platform
MEDLINE®	Embase.com; http://www.embase.com/
Embase®	Embase.com; http://www.embase.com/
MEDLINE® In-Process	Pubmed.com
CENTRAL	Cochrane library; http://mrw.interscience.wiley.com/cochrane/cochrane_search_fs.html

CENTRAL: Cochrane Central Register of Controlled Trials; Embase®: Excerpta Medica Database; MEDLINE®: Medical Literature Analysis and Retrieval System Online

Embase® and MEDLINE® were searched using the embase.com interface. MEDLINE® In-Process was searched using the Pubmed.com interface, while Cochrane was searched using Cochrane® library.

A systematic search strategy was designed for each of the electronic databases searched; the search terms used included keywords and medical subject headings (MeSH terms) focused on Population (P), Intervention (I), Comparator (C), Outcomes (O), and study design (S).

Conference proceedings

Conference abstracts were hand searched to retrieve studies that have not yet been published in full-text articles or abstracts reporting supplementary results of previously published studies. Abstracts from the following conference proceedings were searched for the last three years (2016-2018).

- International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (Europe and International)
- American Society of Clinical Oncology (ASCO)
- American Association for Cancer Research (AACR)
- European Cancer Congress (ECC)/European Society for Medical Oncology (ESMO)
- National Comprehensive Cancer Network (NCCN)
- ASCO genitourinary (GU) Cancers Symposium
- Academy of Managed Care Pharmacy (AMCP) Annual Meeting
- AMCP Nexus

Other source of information

Bibliographic searching of systematic reviews, reviews, and included relevant publications was conducted.

Search strategies

Search strategy for the search conducted in July 2020

Table 123: Economic Review: Embase® search strategy (searched through Embase.com)

S. No.	Search terms	Hits
1	'von hippel lindau disease'/exp OR 'von hippel lindau disease'/syn OR 'hippel angiomatosis' OR 'hippel disease' OR 'hippel lindau disease' OR 'hippel lindau syndrome' OR 'hippel-lindau disease' OR 'lindau disease' OR 'lindau tumor' OR 'lindau tumour' OR 'von hippel disease' OR 'von hippel lindau syndrome' OR 'von hippel-lindau disease' OR vhl OR 'von hippel-lindau'	12294
2	'economics'/de OR 'economic aspect'/de OR 'cost'/de OR Cost OR 'health care cost'/de OR 'drug cost'/de OR 'hospital cost'/de OR 'socioeconomics'/de OR 'health economics'/de OR 'pharmacoeconomics'/de OR 'fee'/exp OR 'budget'/exp OR 'hospital finance'/de OR 'financial management'/de OR 'health care financing'/de OR 'low cost' OR 'high cost' OR (health*care NEXT/1 cost*) OR ('health care' NEXT/1 cost*) OR fiscal OR funding OR financial OR finance OR (cost NEXT/1 estimate*) OR 'cost variable' OR (unit NEXT/1 cost*) OR economic*:ab,ti OR pharmacoeconomic*:ab,ti OR price*:ab,ti OR pricing:ab,ti OR (health*care NEXT/1 (utilisation OR utilization)) OR ('health care' NEXT/1 (utilisation OR utilization)) OR (resource NEXT/1 (utilisation OR utilization OR use)) OR ((cost* NEAR/3 (treat* OR therap*)):ab,ti) OR (((direct OR indirect) NEAR/2 cost*):ab,ti) OR 'cost effectiveness analysis'/syn OR 'cost benefit analysis'/syn OR 'cost utility analysis'/syn OR 'cost minimization analysis'/syn OR 'economic evaluation'/syn OR ((economic OR 'cost-benefit' OR 'cost-effectiveness' OR 'cost-utility') NEXT/1 (evaluation* OR analys* OR model* OR intervention*)) OR (('cost minimization' OR 'cost minimisation') NEXT/1 (analys* OR model*)) OR (economic NEXT/1 (evaluation* OR model))	1749620
3	#1 AND #2	190

Table 124: Economic Review: MEDLINE®In-Process searched via PubMed®

S. No.	Search terms	Hits
#1	von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR "lindau tumor" OR Lindau Disease	6,608

	OR "Lindau syndrome" OR "Lindau Disease" OR Angiomatosis Retinae OR "Angiomatosis Retinae" OR "Lindau's syndrome" OR "Lindau's Disease" OR "VHL disease" OR "VHL Syndrome" OR "Hippel Lindau syndrome" OR "Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau Disease" OR Hippel-Lindau Disease OR "Hippel-Lindau Disease" OR "Familial Cerebello-Retinal Angiomatosis" OR Familial Cerebello-Retinal Angiomatosis OR "von Hippel-Lindau Disease"[Mesh]	
#2	("health resources"[MeSH Terms] OR resource utilization OR "resource utilization" OR pharmacoeconomic OR "budget impact" OR economic OR cost OR "cost minimization" OR cost minimization OR cost comparison OR cost efficiency OR Marginal Analysis OR economic evaluation OR cost utility OR "cost utility" OR cost benefit OR Cost Effectiveness OR "Cost Effectiveness" OR "cost-benefit" OR "Cost-Benefit Analysis"[Mesh]) OR ((("health"[All Fields] OR "health care"[All Fields] OR "healthcare"[All Fields]) AND ("resource"[All Fields] OR "resources"[All Fields] OR "resourced"[All Fields] OR "resourceful"[All Fields])) AND ("utilization"[All Fields] OR "utilisation"[All Fields] OR "utilisations"[All Fields] OR "utilise"[All Fields] OR "utilised"[All Fields] OR "utilises"[All Fields] OR "utilising"[All Fields] OR "utilities"[All Fields] OR "utility"[All Fields] OR "utilizations"[All Fields] OR "utilize"[All Fields] OR "utilized"[All Fields] OR "utilizer"[All Fields] OR "utilizers"[All Fields] OR "utilizes"[All Fields] OR "utilizing"[All Fields]))	1441406
#3	#1 AND #2	**50**

Table 125: Economic Review: CRD York platform

S. No.	Search terms	Hits
#1	von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR "lindau tumor" OR Lindau Disease OR "Lindau syndrome" OR "Lindau Disease" OR Angiomatosis Retinae OR "Angiomatosis Retinae" OR "Lindau's syndrome" OR "Lindau's Disease" OR "VHL disease" OR "VHL Syndrome" OR "Hippel Lindau syndrome" OR "Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau Disease" OR Hippel-Lindau Disease OR "Hippel-Lindau Disease" OR "Familial Cerebello-Retinal Angiomatosis" OR Familial Cerebello-Retinal Angiomatosis OR "von Hippel-Lindau Disease"[Mesh]	**3**

Search strategy for the search conducted in July 2022

Table 126: Summary of search hits retrieved from Embase® and MEDLINE®

(From July 1, 2020 to July 26, 2022; Embase.com)

No.	Query	Results
#1	'von hippel lindau disease'/exp OR 'von hippel lindau disease'/syn OR 'hippel angiomatosis' OR 'hippel disease' OR 'hippel lindau disease' OR 'hippel lindau syndrome' OR 'hippel-lindau disease' OR 'lindau disease' OR 'lindau tumor' OR 'lindau tumour' OR 'von hippel disease' OR 'von hippel lindau syndrome' OR 'von hippel-lindau disease' OR vhl OR 'von hippel-lindau'	13,245
#2	'kidney carcinoma'/syn OR 'kidney tumor'/syn OR 'kidney adenoma'/exp OR ((renal*:ab,ti OR kidney*:ab,ti OR grawit*:ab,ti OR hypernephroid*:ab,ti OR nephroid*:ab,ti) AND (carcinoma*:ab,ti OR cancer*:ab,ti OR neoplasm*:ab,ti OR adeno*:ab,ti OR pyelocarcinoma*:ab,ti OR oncocytoma:ab,ti OR tumor*:ab,ti OR tumour*:ab,ti)) OR 'hypernephroma':ab,ti OR rcc OR 'clear cell renal cell carcinoma' OR 'cc-rcc'	297,995
#3	#1 AND #2	8,625
#4	('von hippel-lindau disease' OR vhl OR 'von hippel-lindau') NEAR/4 ('kidney carcinoma' OR 'kidney tumor' OR 'kidney adenoma' OR 'renal carcinoma' OR 'hypernephroma' OR rcc OR 'clear cell renal cell carcinoma' OR 'cc-rcc' OR 'renal cell carcinoma')	939
#5	#3 OR #4	8,625
#6	'central nervous system'/syn OR 'cns':ab,ti OR 'cns lesion' OR 'cns lesions' OR 'hemangioblastoma'/syn OR 'hemangioblastomas' OR 'cerebellar hemangioblastoma' OR 'spinal cord hemangioblastoma' OR 'brainstem hemangioblastoma' OR 'pancreas lesion' OR 'pancreatic lesion' OR 'pancreas':ab,ti OR 'pancreas cyst' OR 'pancreatic tumor' OR 'pancreatic cystic lesion' OR 'serous cystadenomas' OR 'pancreatic cyst' OR 'pancreatic cysts' OR 'pancreatic neuroendocrine tumors' OR 'pnet' OR 'pnets' OR 'necrotizing pancreatitis'	2,578,883
#7	#1 AND #6	6,232
#8	#5 OR #7	9,191
#9	'economics'/exp OR 'economics'	436,393
#10	'cost'/exp OR 'cost'	1,021,560
#11	'health economics'/exp OR 'health economics'	1,003,625

#12	'budget'/exp OR 'budget'	50,570
#13	budget:ti,ab,kw	31,988
#14	economic*:ti,kw OR cost:ti,kw OR costs:ti,kw OR costly:ti,kw OR costing:ti,kw OR price:ti,kw OR prices:ti,kw OR pricing:ti,kw OR pharmacoeconomic*:ti,kw OR 'pharmaco economic*':ti,kw OR expenditure:ti,kw OR expenditures:ti,kw OR expense:ti,kw OR expenses:ti,kw OR financial:ti,kw OR finance:ti,kw OR finances:ti,kw OR financed:ti,kw OR ((disease:ti,ab,kw OR illness:ti,ab,kw) AND burden:ti,ab,kw)	493,370
#15	(value NEAR/2 (money OR monetary)):ti,ab,kw	3,748
#16	'statistical model'/exp OR 'statistical model'	640,536
#17	economic AND model*:ab,kw	86,248
#18	'probability'/exp OR 'probability'	327,720
#19	markov:ti,ab,kw	34,465
#20	'monte carlo method'/exp OR 'monte carlo method'	48,074
#21	monte AND carlo:ti,ab,kw	57,747
#22	'monte carlo method'/exp OR 'monte carlo method'	48,074
#23	'decision tree'/exp OR 'decision tree':ti,ab,kw OR 'discrete event simulation':ti,ab,kw OR 'discrete choice experiment':ti,ab,kw	26,314
#24	(decision* NEAR/2 (tree* OR analy* OR model*)):ti,ab,kw	43,334
#25	(health*care NEAR/2 (utilisation OR utilization)) OR ('health care' NEAR/2 (utilisation OR utilization)) OR ((resource NEAR/2 (utilisation OR utilization OR use)):ti,ab,kw)	121,524
#26	'resource use':ti,ab,kw OR 'healthcare resources':ti,ab,kw OR 'resource utilization':ti,ab,kw OR 'resource':ti,ab,kw OR 'health resource':ti,ab,kw OR 'healthcare resource':ti,ab,kw	239,530
#27	'patient readmission':ti,ab,kw OR 'patient admission':ti,ab,kw OR 'length of stay':ti,ab,kw OR readmi*:ti,ab,kw OR rehosp*:ti,ab,kw OR 'hospital readmission':ti,ab,kw OR 'reoperation':ti,ab,kw OR 'emergency room':ti,ab,kw	290,811

#28	((outpatient* OR 'clinic' OR physician* OR 'office' OR specialist* OR 'professional' OR 'practitioner') NEAR/2 (visit* OR 'care')):ti,ab,kw	150,183
#29	'cost effectiveness analysis'/syn OR (('cost effectiveness' OR 'cost effective') NEAR/3 (method* OR analys* OR model* OR simulation* OR assessment*)) OR 'cost effectiveness ratio' OR 'cost effectiveness' OR 'cost-effectiveness' OR 'cea':ab,ti,kw OR 'cer':ab,ti,kw	249,177
#30	'cost utility analysis'/syn OR (('cost utility' OR 'cost utilities') NEAR/3 (method* OR analys* OR model* OR simulation* OR assessment*)) OR 'cost utility' OR 'cost-utility' OR 'cua'	16,473
#31	absenteeism OR presenteeism OR productivity:ti,kw,ab	104,006
#32	cost*:ti,ab,kw AND adj2:ti,ab,kw AND (effective*:ti,ab,kw OR utilit*:ti,ab,kw OR benefit*:ti,ab,kw OR minimi*:ti,ab,kw OR analy*:ti,ab,kw OR outcome:ti,ab,kw OR outcomes:ti,ab,kw)	1
#33	#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32	3,434,359
#34	#8 AND #33	369
#35	#34 AND [01-06-2020]/sd	63

Study selection

Two independent reviewers manually screened all citations based on the title and abstract to identify potentially relevant studies. A third independent reviewer resolved any conflicts in their decisions. After the first screening, the full texts of relevant studies were examined in more detail to determine studies eligible for the final inclusion. Two independent reviewers screened the full-text articles, and a third independent reviewer resolved conflicts.

Key eligibility criteria

Key eligibility criteria used during the search conducted in July 2020 are presented in Table 127.

Table 127: Eligibility Criteria (Search Date: July 2020)

Inclusion Criteria	Description
Population(s)	Patients with VHL syndrome (clinically diagnosed or genetically confirmed) including all tumor types
Interventions	No restriction
Comparators	No restriction
Outcomes	Health resource utilization, cost of illness, cost-effectiveness, cost-utility, budget-impact, cost-minimization, cost-comparison
Time	No restriction
Study design	No restriction
Other (Language)	There will be no exclusion on language, non-English studies were categorized separately and shared with Merck team

VHL, Von Hippel-Lindau

Key eligibility criteria used during the update conducted in July 2022 are presented in Table 128.

Table 128: Eligibility Criteria (Search Date: July 2022)

Inclusion Criteria	Description
Population(s)	All patients with VHL disease-associated RCC, CNS Hb, and pNET. VHL disease should be determined with genetically confirmed VHL germline mutation or clinically diagnosed VHL disease. In absence of a clear reporting about the diagnosis method, studies that mention 'VHL disease' will also be included. Studies that did not report VHL diagnostic criteria and included patients with a hereditary/familial cancer (RCC, pNET, CNS Hb) will be included only if there is an overwhelming majority of the patients with VHL disease. Patients with sporadic disease, somatic or epigenetic alterations of VHL gene will be excluded
Interventions	No restriction
Comparators	No restriction
Outcomes	• Study details (objectives, setting, patient population, disease state and model description including cycle length and half cycle corrections etc., perspective, time horizon, source of clinical data, clinical effectiveness) • Cost effectiveness/utility analysis (cost effectiveness and/or cost utility, ICER/ICUR, cost/QALY, cost/LYG, cost/DALY, sensitivity analyses results)

	• Direct/indirect costs, resource use data reported in economic evaluations • QALY, DALY, life years gained • Source of unit costs • Cost year • Source of resource use data • Direct and indirect costs associated with disease • Costs of absenteeism • Costs of end-of-life care • Volume of resources use • Resource utilization due to interventions associated complications • Cost due to interventions associated complications
Time	July 2020 to present (last 2 years)
Study design	• Cost effectiveness analysis • Cost utility analysis • Cost benefit analysis • Cost consequence studies • Cost studies/surveys/analysis • Burden of illness studies • Resource use studies • CE studies reporting cost • Cost minimization analysis • Budget impact models
Regions	Global
Other (Language)	There will be no exclusion on language, non-English studies will be categorized separately and shared with Merck team. Studies with English abstract and non-English full-texts will be flagged separately

CE, Cost-Effectiveness; CNS, Central Nervous System; DALY, Disability Adjusted Life Years; ICER, Incremental Cost Effectiveness Ratio; Hb, Hemangioblastoma; ICUR, Incremental Cost Utility Ratio; LYG, Life years gained; pNET, Pancreatic Neuroendocrine Tumors; RCC, Renal Cell Carcinoma; QAL, Disability Adjusted Life Years; VHL, Von Hippel-Lindau

Table 129: Summary of search hits retrieved from MEDLINE® Inprocess (From July 1, 2020 to July 26, 2022; Pubmed)

No.	Query	Results
1	von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR "lindau tumor" OR Lindau Disease OR "Lindau syndrome" OR "Lindau Disease" OR "Lindau's syndrome" OR "Lindau's Disease" OR "VHL disease" OR "VHL Syndrome" OR "Hippel Lindau syndrome" OR "Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau Disease" OR Hippel-Lindau Disease OR "Hippel-Lindau Disease" OR "von Hippel-Lindau Disease"[Mesh]	6,959
2	(((renal cell carcinoma[MeSH Terms] OR "kidney carcinoma" OR "kidney tumor" OR "kidney adenoma" OR "kidney cancer") OR ((Renal OR kidney OR grawit* OR hypernephroid* OR nephroid*) AND (carcinoma OR cancer OR neoplasm OR adenoma OR pyelocarcinoma OR oncocytoma OR tumor OR tumour))) OR (hypernephroma OR "clear cell renal cell carcinoma")) OR (clear cell renal carcinoma[MeSH Terms])	232,928
3	"central nervous system" OR "CNS"[Title/Abstract] OR "CNS lesion" OR "CNS lesions" OR "hemangioblastoma"OR "hemangioblastomas" OR "cerebellar hemangioblastoma" OR "spinal cord hemangioblastoma" OR "spinal cord hemangioblastoma" OR "brainstem hemangioblastoma" OR "pancreas lesion" OR "pancreatic lesion" OR "Pancreas"[Title/Abstract] OR "Pancreas cyst" OR "pancreatic tumor" OR "pancreatic cystic lesion" OR "serous cystadenomas" OR "pancreatic cyst" OR "pancreatic cysts" OR "pancreatic neuroendocrine tumors" OR "pNET" OR "pNETs" OR "Necrotizing pancreatitis"	454,578
4	#1 AND (#2 OR #3)	3,971
5	"Economics" OR "Costs and Cost Analysis"[mh] OR "Economics, Nursing"[mh] OR "Economics, Medical"[mh] OR "Economics, Pharmaceutical"[mh] OR "Economics, Hospital"[mh] OR "Economics, Dental"[mh] OR "Fees and Charges"[mh] OR "Budgets"[mh] OR budget*[tiab] OR economic*[tiab] OR cost[tiab] OR costs[tiab] OR costly[tiab] OR costing[tiab] OR price[tiab] OR prices[tiab] OR pricing[tiab] OR pharmacoeconomic*[tiab] OR pharmaco-economic*[tiab] OR expenditure[tiab] OR expenditures[tiab] OR expense[tiab] OR expenses[tiab] OR financial[tiab] OR finance[tiab] OR finances[tiab] OR financed[tiab] OR value for money[tiab] OR monetary value*[tiab] OR "models, economic"[mh] OR economic model*[tiab] OR "markov chains"[mh] OR markov[tiab] OR "monte carlo method"[mh] OR monte carlo[tiab] OR "Decision Theory"[mh] OR decision tree*[tiab] OR decision analy*[tiab] OR decision model*[tiab] (burden[tiab] AND	33,127

No.	Query	Results
	(illness[tiab] OR disease[tiab])) OR "discrete event simulation"[tiab] OR "discrete event experiment"[tiab]	
6	#4 AND #5	2
7	#6 AND (inprocess[sb] OR pubstatusaheadofprint)	0

Table 130: Summary of search hits retrieved from NHS EED, DARE and HTA (From July 1, 2020 to July 26, 2022; York CRD)

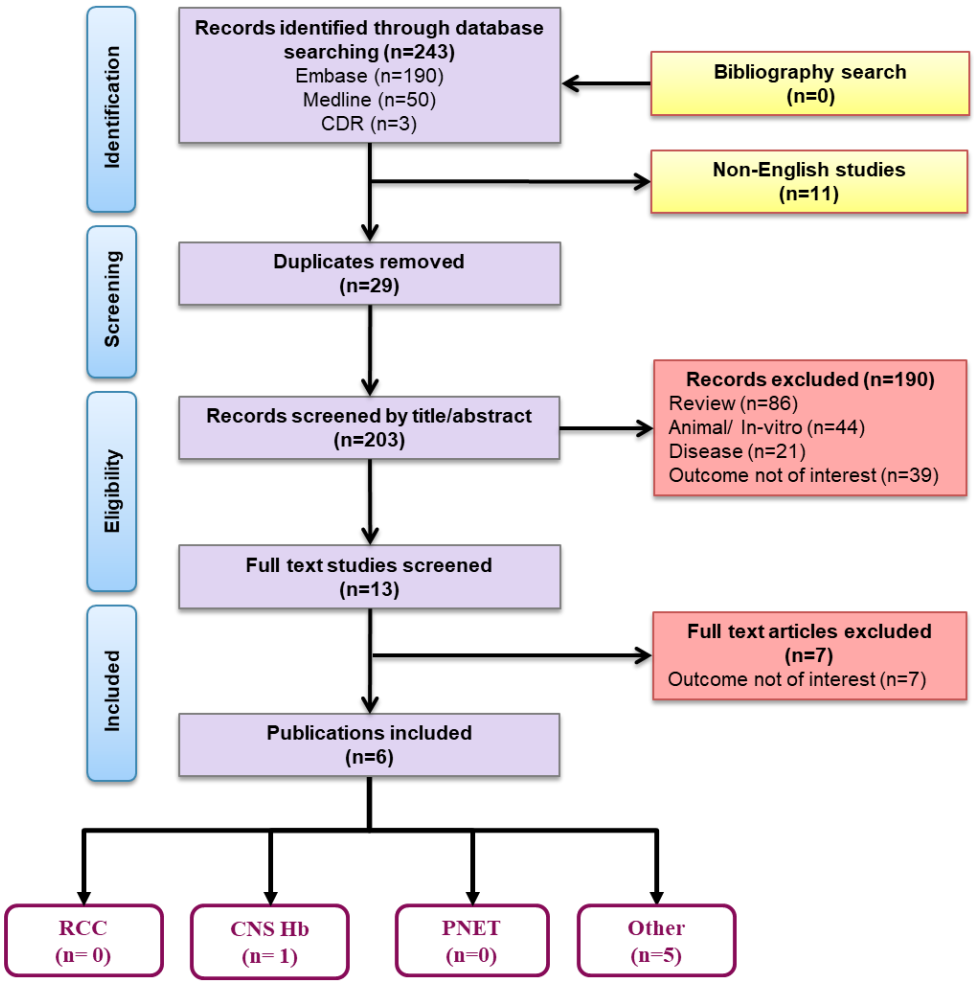
No.	Query	Results
#1	Results for: (von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR lindau tumor OR Lindau Disease OR Lindau syndrome OR Lindau Disease OR Lindau"s syndrome OR Lindau"s Disease OR VHL disease OR VHL Syndrome OR Hippel Lindau syndrome OR Hippel Lindau Disease OR von Hippel Lindau syndrome OR von Hippel Lindau Disease OR Hippel-Lindau Disease OR Hippel-Lindau Disease OR von Hippel-Lindau Disease):TI FROM 2020 TO 2022	0

G1.3 Results for economic review

PRISMA flow diagram

Of 203 unique records, six studies were identified reporting data on economic burden of genetic analysis for germline mutation in VHL gene. The data was reported from Italy, France, US, UK, and India (n=5 studies) (Figure 35).

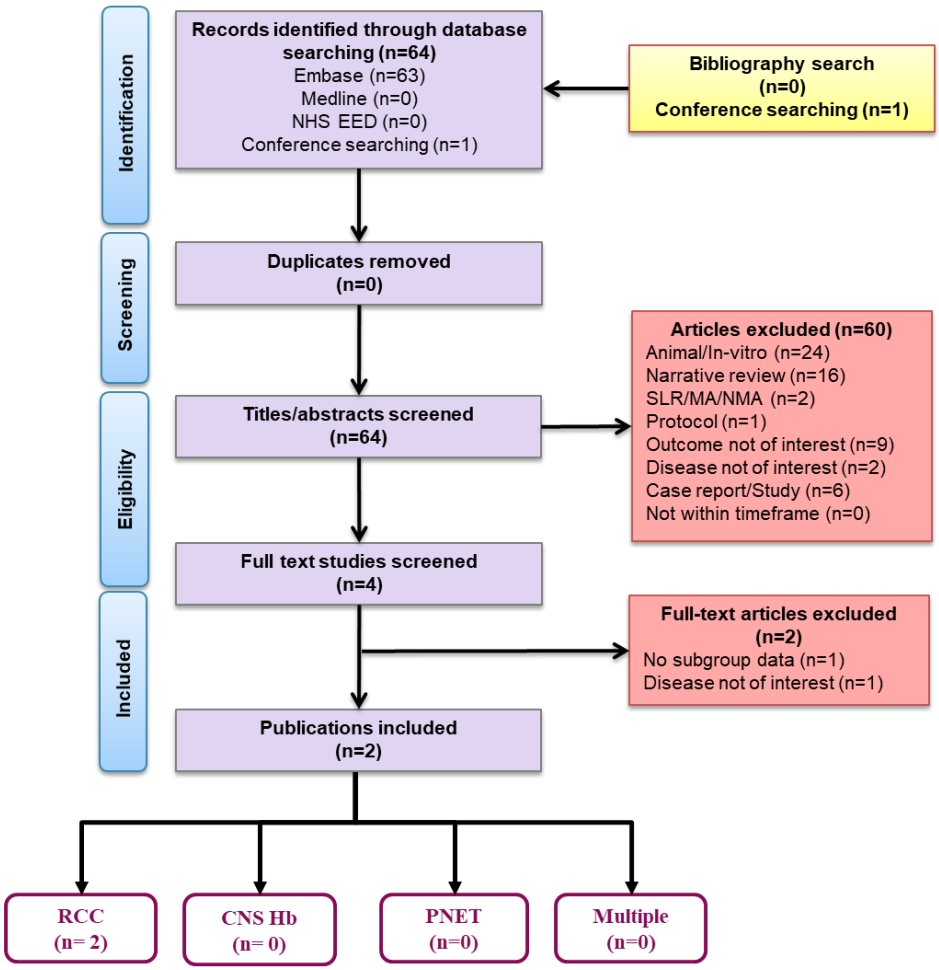
Figure 35: PRISMA flow diagram (Search date: July 2020)



CRD: Center for Reviews and Dissemination

As part of the 2022 update, a total of 64 records were retrieved through electronic literature searches – 63 through database searches (Embase [n=63], PubMed [n=0], and NHS EED [n=0]), and 1 record was identified through conference searches. No duplicates were identified. Following the screening of all 64 studies, 4 studies were selected for full text review, 2 of which met the pre-defined eligibility criteria, and were included in the update (Figure 36).

Figure 36: PRISMA flow diagram (July 2022)



MA: Metanalysis; NHS EED: National Health Service Economic Evaluation Database; NMA: Network Metanalysis; SLR: Systematic Literature Review

Study characteristics

Following the searches conducted in July 2020, a total of six studies were included in the economic evaluations, of which four studies were published as journal article, and two studies were published as conference abstracts. Four studies were retrospective in nature, while one study was prospective in nature. Five studies each reported data from Italy, France, US, UK, and India. Country was not reported in one study (Garrett 2017). Three studies included patients <50 patients, while three studies included patients ≥100 patients. Overall, sample size of the patient ranged from 7 patients to 280 patients across six studies.

Additional two conference poster abstracts were identified during an update conducted in July 2022. Both studies were retrospective in nature and used the

Optum Clinformatics claims database to collect information on HRU and costs associated with the treatment of VHL-associated RCC in the US (Jonasch 2022, Sundaram 2022).

Detailed summary of study characteristics is presented in Table 131.

Table 131: Summary of Study Characteristics

Study name	Publication type	Study design	Country	Sample size	Time frame	Setting
(Jonasch 2022)	Abstract	Retrospective	US	160	NR	NR
(Sundaram 2022)	Abstract	Retrospective	US	160	NR	NR
(Rattenberry 2013)	Journal	Prospective	UK	205	NR	West Midlands Regional Genetics Laboratory, UK
(Catapano 2005)	Journal	Retrospective	Italy	14	1993-2002	Casa Sollievo della Sofferenza Hospital, Italy
(Pigny 2009)	Journal	Retrospective	France	100	2002-2007	CHRU de Lille, France
(McInerney-Leo 2014)	Abstract	NR	US	7	NR	NR
(Pai 2014)	Journal	Retrospective	India	44	January 2010-June 2012	Christian Medical College, Vellore, India
(Garrett 2017)	Abstract	Retrospective	NR	280	20142016	Surgical pathology electronic medical record

CHRU, Regional University Hospital Center (Centre Hospitalier Régional Universitaire); NR: Not Reported; sample size is the total number of patients included in the study

Population characteristics

The mean age of the patients was reported in three studies ranging from 37.8 years to 51.5 years. The age of the patient who were positive for VHL gene varied across the three studies ranging from 14 years to 45 years. The type of tumor reported across the studies included pheochromocytomas (4 studies), paragangliomas (3 studies), CNS hemangioblastoma (1 study) (Catapano 2005) and metastatic liver tumor (Garrett 2017) (1 study). The detailed summary of population characteristic is presented in Table 132.

Table 132: Summary of Population Characteristics

Study name	Country	Sample size	VHL gene mutation, n	Mean Age (years)	Tumour type	Genetic analysis	Gene panel
(Jonasch 2022)	US	160	NR	51.5 years	RCC	N/A	N/A
(Sundaram 2022)	US	160	NR	NR	RCC	N/A	N/A
(Rattenberry 2013)	UK	205	NR	NR	PCC, PGL	Next generation sequencing	MAX, RET, SDHA, SDHB, SDHC, SDHD, SDHAF2, TMEM127, and VHL
(Catapano 2005)	Italy	14	2	43.5 (Age of 2 VHL patients: 32 and 45 years)	CNS Hb	PCR	VHL
(Pigny 2009)	France	100 (8 with gene mutation)	2	37.8 (8 patients) (Age of 2 VHL patients: 13 and 43 years)	PCC	Systematic genetic testing, target genetic testing	RET, VHL, SDHD, SDHB

(McInerney-Leo 2014)	US	7	NR	NR	PCC, PGL	Illumina/TruSeq exome, Roche/NimbleGen SeqCapEZ	RET, NF1, VHL, SDHD, SDHB, SDHC, SDHA, DHAF2, KIF1B, TMEM127, EGLN1 and MAX)
(Pai 2014)	India	44 (13 with gene mutations)	4	49 (13 patients) (Age of 4 VHL patients: 14, 18,21,33 years)	PCC, PGL	PCR; Succinate dehydrogenase B immunohistochemistry	RET, VHL, SDHB
(Garrett 2017)	NR	280	NA	NR	Metastatic tumor to liver	Immunohistochemistry	VHL

Hb: Hemangioblastoma; NR: Not Reported; NA: Not Applicable, PCR: Polymerase Chain Reaction; PCC: Pheochromocytomas; PGL: Paragangliomas; SDHB: Succinate Dehydrogenase B; VHL: Von Hippel Lindau

Summary of Economic Review Results

Summary of results from cost burden studies on genetic testing

- Clinically guided genetic testing would lead to cost reduction by 24% (42.700 euros per 100 patients attending the outpatient visit) (Pigny 2009).

- Whole exome sequencing appeared to be cost-effective method of detecting germline mutations in patients with phaeochromocytoma or paraganglioma (McInerney-Leo 2014).

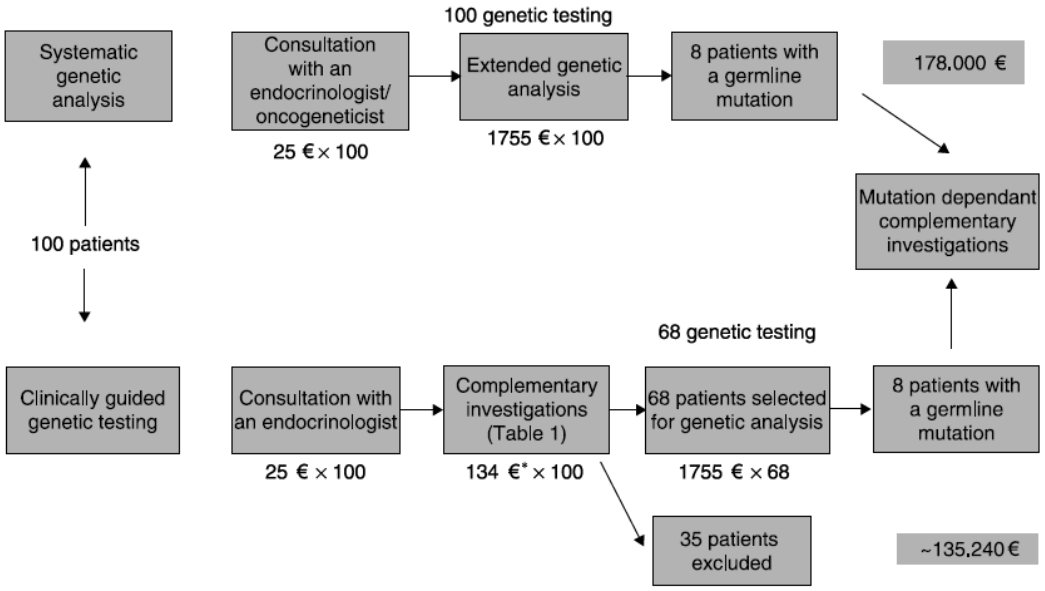
- Cost analysis showed that triaging with SDHB immunohistochemistry could lead to cost saving by USD $320–500 per patient (Pai 2014).

- NGS-based genetic testing strategy provided a cost-effective approach to increase detection of individuals with inherited PPGL and HNPGL (Rattenberry 2013).
- Antibody staining analysis of the VHL gene product is not cost effective strategy in primary testing (Garrett 2017).
- Clinical-instrumental monitoring program for VHL testing costs approximately 1400 Euros (Italian Health Care Service) (Catapano 2005)

Five studies reported the data on cost-effectiveness in terms of cost comparison of different techniques used for the detection of VHL gene mutation.

In a study by Pigny et al, of eight patients with pheochromocytoma, germline VHL Y156C mutation was reported in two patients, one male and one female. Male patient had bilateral pheochromocytoma and age of onset was 13 years; while female patient had unilateral pheochromocytoma and age of onset was 43 years. The cost for searching for a germline VHL mutation by nucleotide sequencing (three exons) is 235 euros that includes individual costs of each technical step performed (i.e. DNA extraction plus PCR-sequencing of each exon on both DNA strands) and this cost does not include the search of complex molecular events such as deletion(s)/insertion(s) or gene rearrangements. In this study cost analysis of the two algorithms- systematic genetic testing and target genetic testing were compared (Figure 37). Target genetic testing required a more detailed clinical survey for selection the patients eligibility for genetic testing. Patients with an age of onset <50 years and/or bilateral pheochromocytoma will be eligible for genetic testing. No patients with a hereditary tumor would be missed with type of genetic analysis Clinically guided genetic testing would lead to cost reduction by 24% (42.700 euros per 100 patients attending the outpatient visit) (Pigny 2009).

Figure 37: Cost analysis of the two algorithms, systematic genetic analysis, and clinically guided genetic analysis



Source: (Pigny 2009)

Mcinerney-Leo et al, determined whether whole exome sequencing using an off-the-shelf exome chip was cost effective method of detecting causative mutations. The gene panel included RET, NF1, VHL, SDHD, SDHB, SDHC, SDHA, SDHAF2, KIF1B, TMEM127, EGLN1 and MAX. Whole exome sequencing was performed on samples from 7 unrelated individuals with phaeochromocytoma or paraganglioma. The cost of exome sequencing of all known PCC/PGL genes (13 genes) was approximately $AU1,500 per sample; while cost of sequencing only the four most common genes was $AU4,100; hence, whole exome sequencing with Roche/NimbleGen SeqCap EZ appeared to be cost-effective method of detecting germline mutations in patients with phaeochromocytoma or paraganglioma (McInerney-Leo 2014).

Pai et al, evaluated the cost benefit of succinate dehydrogenase B immunohistochemistry triaging before genetic analysis in 44 cases with pheochromocytomas and paragangliomas. Of 44 cases, four cases were of VHL gene mutations. The age of the patients ranged from 17–33 years and had adrenal tumor.

The cost of gene sequencing ranged between $150 to $600 (the number of genes sequenced per case) as per genetic analysis algorithm for suggested by Benn et al. In this study, cost of immunohistochemistry was cheaper ($14 per test). Initially 35 cases underwent succinate dehydrogenase B mutational analysis. SDHB immunohistochemistry was used prior to genetic screening. Mutational analysis helped to detect mutations in three cases with succinate dehydrogenase B and three cases with succinate dehydrogenase D; hence, mutational analysis was not required for the remaining 29 patients who were eventually found to be negative for mutations. This led to significant cost reduction and cost saving between $320 to $500 per patient (Pai 2014).

Rattenberry et al, developed a novel diagnostic assay- NGS-based genetic testing that increased the detection of germline mutations in patients with pheochromocytomas and paragangliomas (PPGL) and Head and neck paragangliomas (HNPGL) at a substantially lower cost (70% cost reduction) than conventional (Sanger-based) molecular genetic analysis. The diagnostic test was locally priced at £500 ~750) per sample to test for 9 PPGL/HNPGL genes (MAX, RET, SDHA, SDHB, SDHC, SDHD, SDHAF2, TMEM127, and VHL). The price charged locally for testing 4 genes (VHL, RET, SDHB, and SDHD) by conventional Sanger sequencing technology and MLPA was ~£1800 (~$2700). Hence, NGS-based genetic testing appeared to be cost-effective mutation detection technique (Rattenberry 2013).

During 2014-2016, a study by Garette et al, identified 280 cases of metastatic liver tumor. As the prevalence of primary site carcinomas positive for VHL was low, incremental cost of $150 per case for incorporation of VHL in preliminary basic panel was determined to be inappropriate; hence, antibody staining analysis of the VHL gene product in primary testing was not cost effective strategy; however, incorporating antibody staining analysis in second tier testing would be most cost-effective strategy (Garrett 2017).

One study evaluated the cost burden of clinical instrumental monitoring and molecular genetic screening programs for VHL gene. In a study by Catapano et al, peripheral blood of 14 patients (6 female and 8 male) with symptomatic CNS hemangioblastoma was tested for VHL gene. Of 14 patients, germline mutations of the VHL gene were identified in two (14%) patients (one male patient of age 32 years and one female of age 45 years). The costs of the clinical instrumental monitoring and molecular genetic screening programs was evaluated. The costs of periodically performed clinical-instrumental monitoring program was ~1400 Euros (Italian Health Care Service) that includes pedigree analysis, physical examination, 24-h urine test, ophthalmological, upper abdominal ultrasound, MRI neuraxis, audiogram and MRI inner ear. The cost of one time performed molecular screening was 250 Euros per patient, cost of effect sequence was approximately 280 Euros. Cost of the genetic screening after identification of mutation carrier was ~120 Euros per family member. Other gene sequencing techniques such as automated sequencing, Southern blotting, and fluorescence in situ hybridization would cost approximately 750 Euros per patient (Table 133) (Catapano 2005).

Table 133: Cost Associated with Genetic Analysis of VHL Gene

Test	Cost
Clinical-instrumental monitoring program	1400 Euros
Molecular screening (based on the use of dHPLC methodology)	250 Euros per patient
Effect sequence	280 Euros per patient
Genetic screening	120 Euros for each family
Automated sequencing, Southern blotting, and fluorescence in situ hybridization	750 Euros per patient

Source: (130)
dHPLC: Denaturing High-Performance Liquid Chromatography

Table 134: Overall Cost Burden Reported Across Six Studies

Study	Technique	Cost
(Pigny 2009)	Clinically guided genetic testing	42.700 euros per 100 patients

	Consultation with endocrinologist/oncogeneticist	25 Euros per patient
	Extended genetic analysis	1755 Euros per patient
	Complementary investigations	134 Euros per patient
(McInerney-Leo 2014)	Exome sequencing of all known phaeochromocytoma/ paraganglioma genes (13 genes) with Roche/nimblegen seqcap EZ	$AU1,500 per sample
	Sequencing 4 genes	$AU4,100
(Pai 2014)	Gene sequencing with Benn et al	$150–$600
	Cost of immunohistochemistry	$14 per test
(Rattenberry 2013)	NGS-based genetic testing	£500 (~$750) per sample
	Sanger sequencing technology and MLPA	~£1800 (~$2700)
(Garrett 2017)	antibody staining analysis	$150 per case
(Catapano 2005)	Clinical-instrumental monitoring program	1400 Euros
	Molecular screening (based on the use of dHPLC methodology)	250 Euros per patient
	Effect sequence	280 Euros per patient
	Genetic screening	120 Euros for each family
	Automated sequencing, Southern blotting, and fluorescence in situ hybridization	750 Euros per patient

dHPLC: Denaturing High-Performance Liquid Chromatography; MLPA: Multiplex Ligation-Dependent Probe Amplification; NGS: Next Generation Sequencing

Table 135: Overall Cost Burden of Different Techniques Used for VHL Mutation Detection

Study	Technique	Cost saving
(Pigny 2009)	Target genetic testing versus Systematic genetic testing	✓
(McInerney-Leo 2014)	Whole exome sequencing with Roche/NimbleGen SeqCap EZ	✓
(Pai 2014)	Immunohistochemistry versus Genetic analysis algorithm for suggested by Benn et al	✓
(Rattenberry 2013)	Next-generation sequencing -based genetic testing versus Conventional (Sanger-based) molecular genetic analysis	✓
(Garrett 2017)	Antibody staining analysis (primary screening)	×

Summary of results from HRU and healthcare cost studies

- VHL-RCC is associated with significant HRU and healthcare costs due to the needs of managing RCC (Jonasch 2022)
- Inpatient costs were the main driver of all-cause healthcare costs (Jonasch 2022)
- Patients with VHL-associated RCC require costly tumor-reduction procedures (the mean cost of a nephrectomy and ablation/cryotherapy was $29,313 and $18,290, respectively) (Sundaram 2022)
- Repeated tumor-reduction procedures put patients at risk of surgery-related complications which further add to the healthcare costs; the most expensive complications were related to impaired renal function: acute renal failure ($21,013 over 4 weeks), CKD ($26,032 over 6 months) and end stage renal disease ($65,338 over 6 months) (Sundaram 2022)

Two studies reported data on HRU and healthcare costs associated with treatment of VHL-associated RCC in the US setting using the Optum Clinformatics claims database. Study by Jonasch et al. reported the HRU and associated costs (Jonasch 2022) whereas study by Sundaram focused on the costs of tumor reduction procedures and complication associated with them (Sundaram 2022).

Jonasch et al. analyzed data of 160 patients with VHL-associated RCC. During the study period, patients with VHL-associated RCC incurred 0.10 hospitalizations (0.52 inpatient days), 1.49 outpatient visits, 0.09 emergency department visits, and 0.17 other medical visits per person-month on average (Table 136). This translated to a monthly all-cause healthcare cost of $4,276, which included $2,222 inpatient, $1,318 outpatient, $188 emergency department visits, $63 other medical visits, and $485 pharmacy costs. $1,627 out of the all-cause cost per month was RCC-related, which was mainly driven by inpatient costs for RCC ($1,184/month). There were also notable costs associated

with other tumor types: $2,159/month for CNS Hb, and $3,306/month for pNETs on average. Among patients with VHL-associated RCC who received surgical procedures for different tumors, the average hospitalization or outpatient visit costs were $28,356 for nephrectomies, $70,515 for CNS Hb surgeries and $81,825 for pNET surgeries (Table 137) (Jonasch 2022).

Table 136: Summary of HRU Results

Study name	Country	Source	Sample size	Patient population	HRU outcomes (mean, per person-month)
(Jonasch 2022)	US	Optum informatics claims data	160	Patients with VHL-associated RCC	Hospitalisations: 0.1
					In patient days: 0.52
					Outpatient visits: 1.49
					Emergency department visits: 0.09
					Other medical visits: 0.17

HRU: Healthcare Utilization; RCC: Renal Cell Carcinoma; VHL: Von Hippel Lindau

Additionally, as demonstrated by Sundaram et al., patients with VHL-associated RCC require costly tumor-reduction procedures (the mean cost of a nephrectomy and ablation/cryotherapy was $29,313 and $18,290, respectively). Repeated tumor-reduction procedures put patients at risk of surgery-related complications which further add to the healthcare costs; the most expensive complications were related to impaired renal function: acute renal failure ($21,013 over 4 weeks), CKD ($26,032 over 6 months) and end stage renal disease ($65,338 over 6 months) (Table 137) (Sundaram 2022).

Table 137: Summary of healthcare cost results

Study	Country	Source	Patient population	Total mean cost (USD)
	US			Healthcare cost:

(Jonasch 2022)		Optum informatics claims data	Patients with VHL-associated RCC	• All cause: 4,276 per person per month • RCC-related: 1,627 per person per month
				Inpatient cost • All cause: 2,222 per person per month • RCC-related: 1,184 per person per month
				All cause outpatient cost: 1,318 per person per month
				All cause emergency department visit: 188 per person per month
				All cause other medical visits: 63 per person per month
				All cause pharmacy costs: 485 per person per month
				Average hospitalization or outpatient visit costs: • Per nephrectomy: 28,356 • Per CNS Hb surgery: 70,515 • Per pNET surgery: 1,825
(Sundaram 2022)	US	Optum informatics claims data	Patients with VHL-associated RCC	Treatment-reduction procedure cost: • Per nephrectomy: 29,313 • Per ablation/cryotherapy: 18,290
				Surgery-related complication cost: • Acute renal failure: 21,013 over 4 weeks • Chronic kidney disease: 26,032 over 6 months • End stage renal disease: 65,338 over 6 months

CNS: Central Nervous System; Hb: Hemangioblastoma; pNET: Pancreatic Neuroendocrine Tumors; RCC: Renal Cell Carcinoma; VHL: Von Hippel Lindau

G1.4 References for economic burden SLR

Aufforth, R. D., Ramakant, P., Sadowski, S. M., Mehta, A., Trebska-McGowan, K., et al. (2015). Pheochromocytoma Screening Initiation and Frequency in von Hippel-Lindau Syndrome. J Clin Endocrinol Metab 100(12): 4498-4504.

Binderup, M. L. M. (2018). von Hippel-Lindau disease: Diagnosis and factors influencing disease outcome. Dan Med J 65(3).

Blansfield, J. A., Choyke, L., Morita, S. Y., Choyke, P. L., Pingpank, J. F., et al. (2007). Clinical, genetic and radiographic analysis of 108 patients with von Hippel-Lindau disease (VHL) manifested by pancreatic neuroendocrine neoplasms (PNETs). Surgery 142(6): 814-818; discussion 818.e811-812.

Catapano, D., Muscarella, L. A., Guarnieri, V., Zelante, L., D'Angelo, V. A., et al. (2005). Hemangioblastomas of central nervous system: molecular genetic analysis and clinical management. Neurosurgery 56(6): 1215-1221.

Garrett, E., Mohammad, A., McCormick-Baw, C., Zreik, R. and Rao, A. (2017). Modeling of Effectiveness of Addition of New Antibodies in Panel Immunohistohistochemistry. LABORATORY INVESTIGATION, NATURE PUBLISHING GROUP 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013-1917 USA.

Jafri, M. and Maher, E. R. (2012). The genetics of phaeochromocytoma: using clinical features to guide genetic testing. Eur J Endocrinol 166(2): 151-158.

Jafri, M., Whitworth, J., Rattenberry, E., Vialard, L., Kilby, G., et al. (2013). Evaluation of SDHB, SDHD and VHL gene susceptibility testing in the assessment of individuals with non-syndromic phaeochromocytoma, paraganglioma and head and neck paraganglioma. Clinical endocrinology 78(6): 898-906.

Jilg, C. A., Neumann, H. P., Gläsker, S., Schäfer, O., Ardelt, P. U., et al. (2012). Growth kinetics in von Hippel-Lindau-associated renal cell carcinoma. Urol Int 88(1): 71-78.

Jonasch, E., Song, Y., Freimark, J., Berman, R., Nguyen, H., et al. (2022). Healthcare resource utilization (HRU) and costs among patients with Von Hippel-Lindau disease (VHL)-associated renal cell carcinoma (RCC): A retrospective administrative claims analysis. Journal of Clinical Oncology 40(6_suppl): 305-305.

Kim, E. and Zschiedrich, S. (2018). Renal Cell Carcinoma in von Hippel-Lindau Disease-From Tumor Genetics to Novel Therapeutic Strategies. Front Pediatr 6: 16.

Lonser, R. R., Butman, J. A., Huntoon, K., Asthagiri, A. R., Wu, T., et al. (2014). Prospective natural history study of central nervous system hemangioblastomas in von Hippel-Lindau disease. Journal of neurosurgery 120(5): 1055-1062.

Lonser, R. R., Glenn, G. M., Walther, M., Chew, E. Y., Libutti, S. K., et al. (2003). von Hippel-Lindau disease. Lancet 361(9374): 2059-2067.

Lonser, R. R. B., J. A.: Kiringoda, R.: Song, D.: Oldfield, E. H. (2009). Pituitary stalk hemangioblastomas in von Hippel-Lindau disease. J Neurosurg 110(2): 350-353.

McInerney-Leo, A. M., Marshall, M. S., Gardiner, B., Benn, D. E., McFarlane, J., et al. (2014). Whole exome sequencing is an efficient and sensitive method for detection of germline mutations in patients with phaeochromcytomas and paragangliomas. Clinical endocrinology 80(1): 25-33.

Medline Plus. "Von Hippel-Lindau syndrome." 2020, from https://medlineplus.gov/genetics/condition/von-hippel-lindau-syndrome/#statistics.

Pai, R., Manipadam, M. T., Singh, P., Ebenazer, A., Samuel, P., et al. (2014). Usefulness of S uccinate dehydrogenase B (SDHB) immunohistochemistry in guiding mutational screening among patients with pheochromocytoma-paraganglioma syndromes. Apmis 122(11): 1130-1135.

Pigny, P., Cardot-Bauters, C., Do Cao, C., Vantyghem, M. C., Carnaille, B., et al. (2009). Should genetic testing be performed in each patient with sporadic

pheochromocytoma at presentation? European journal of endocrinology 160(2): 227-231.

Rattenberry, E., Vialard, L., Yeung, A., Bair, H., McKay, K., et al. (2013). A comprehensive next generation sequencing–based genetic testing strategy to improve diagnosis of inherited pheochromocytoma and paraganglioma. The Journal of Clinical Endocrinology & Metabolism 98(7): E1248-E1256.

Schmid, S., Gillessen, S., Binet, I., Brändle, M., Engeler, D., et al. (2014). Management of von hippel-lindau disease: an interdisciplinary review. Oncol Res Treat 37(12): 761-771.

Sundaram, M., Song, Y., Freimark, J., Berman, R., Nguyen, H., et al. (2022). Real-world treatment patterns in von Hippel-Lindau (VHL) disease-associated renal cell carcinoma (RCC): Costs of tumor reduction procedures and their complications. Journal of Clinical Oncology 40(16_suppl): 4539-4539.

Zbar, B., Kishida, T., Chen, F., Schmidt, L., Maher, E. R., et al. (1996). Germline mutations in the Von Hippel-Lindau disease (VHL) gene in families from North America, Europe, and Japan. Human mutation 8(4): 348-357.

Zhou, B., Wang, J., Liu, S., Peng, X., Hong, B., et al. (2019). Hemangioblastoma Instead of Renal Cell Carcinoma Plays a Major Role in the Unfavorable Overall Survival of Von Hippel-Lindau Disease Patients. Frontiers in Oncology 9.

Appendix H: Health-related quality-of-life studies

H1.1 Objectives

This systematic review was conducted for synthesizing evidence related to humanistic burden of VHL disease with a focus on VHL-associated RCC, CNS hemangioblastomas and pNETs.

H1.2 Systematic review methodology

Identification of studies

In July 2020, a comprehensive search was conducted including biomedical databases, conference proceedings, and other online resources. Details of the search strategy are provided in Table 139-Table 142. An update to the searches was conducted in July 2022. Details of the search strategy are provided in Table 143-Table 145.

Information sources

Biomedical databases

The databases listed in **Table 138** were searched.

Table 138: Databases Searched for the Literature Review and the Search Platform

Data sources	Platform
MEDLINE®	Embase.com; http://www.embase.com/
Embase®	Embase.com; http://www.embase.com/
MEDLINE® In-Process	Pubmed.com
CENTRAL	Cochrane library; http://mrw.interscience.wiley.com/cochrane/cochrane_search_fs.html

CENTRAL: Cochrane Central Register of Controlled Trials; Embase®: Excerpta Medica Database; MEDLINE®: Medical Literature Analysis and Retrieval System Online

Embase® and MEDLINE® were searched using the embase.com interface. MEDLINE® In-Process was searched using the Pubmed.com interface, while Cochrane was searched using Cochrane® library.

A systematic search strategy was designed for each of the electronic databases searched; the search terms used included keywords and medical subject headings (MeSH terms) focused on Population (P), Intervention (I), Comparator (C), Outcomes (O), and study design (S).

Conference proceedings

Conference abstracts were hand searched to retrieve studies that have not yet been published in full-text articles or abstracts reporting supplementary results of previously published studies. Abstracts from the following conference proceedings were searched for the last three years (2016-2018).

- International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (Europe and International)
- American Society of Clinical Oncology (ASCO)
- American Association for Cancer Research (AACR)
- European Cancer Congress (ECC)/European Society for Medical Oncology (ESMO)
- National Comprehensive Cancer Network (NCCN)
- ASCO genitourinary (GU) Cancers Symposium
- Academy of Managed Care Pharmacy (AMCP) Annual Meeting
- AMCP Nexus
- American Academy of Ophthalmology (AAO)
- Association for Research in Vision and Ophthalmology (ARVO)

Other source of information

Bibliographic searching of systematic reviews, reviews, and included relevant publications was conducted.

Search strategies

Search strategy for the search conducted in July 2020

Table 139 ***Embase® search strategy for PRO's and quality of life review***

S. No.	Search terms	Hits
• #1	• 'von hippel lindau disease'/exp OR 'von hippel lindau disease'/syn OR 'hippel angiomatosis' OR 'hippel disease' OR 'hippel lindau disease' OR 'hippel lindau syndrome' OR 'hippel-lindau disease' OR 'lindau disease' OR 'lindau tumor' OR 'lindau tumour' OR 'von hippel disease' OR 'von hippel lindau syndrome' OR 'von hippel-lindau disease' OR vhl OR 'von hippel-lindau'	12,294
• #2	• ((utilit* NEAR/2 (measure* OR outcome* OR state* OR health OR score* OR weight* OR analysis)):ab,ti) OR 'health utility index' OR 'hui':ab,ti OR 'hrqol':ab,ti OR 'hqol':ab,ti OR 'quality of life'/exp OR 'quality of life' OR 'quality-of-life'/exp OR 'quality-of-life' OR qol:ab,ti OR (utilit* NEXT/1 (score* OR value* OR evaluation*)) OR (health NEXT/2 utilit*) OR (('health'/exp OR 'health') AND (state NEXT/1 utilit*)) OR hui:ab,ti OR ((health NEXT/1 state*) AND (state* NEXT/1 preference*)) OR 'quality adjusted life year'/exp OR 'quality adjusted life year' OR 'quality adjusted life' OR ('quality adjusted' NEXT/1 survival*) OR qaly:ab,ti OR qald:ab,ti OR qale*:ab,ti OR qtime*:ab,ti OR 'disability adjusted life' OR daly*:ab,ti OR 'health survey'/exp OR 'health survey' OR hye*:ab,ti OR health*year*equivalent OR (health NEAR/2 utility*) OR 'wellbeing'/exp OR 'wellbeing':ab,ti OR (quality NEAR/2 well*being) OR qwb:ab,ti OR (willingness NEAR/2 pay) OR (standard NEAR/2 gamble) OR disutili*:ab,ti OR (time NEAR/2 trade*off) OR tto:ab,ti OR ('discrete choice' NEXT/1 experiment*) OR 'short form 36'/exp OR 'short form 36' OR 'sf36':ab,ti OR 'sf-36':ab,ti OR 'sf 36':ab,ti OR 'short form 12'/exp OR 'short form 12' OR 'sf12':ab,ti OR 'sf-12':ab,ti OR 'sf 12':ab,ti OR 'short form 6' OR 'sf6':ab,ti OR 'sf-6':ab,ti OR 'sf 6':ab,ti OR 'euroqol' OR 'euro-qol'	4,626,352

	OR 'euro qol' OR 'eq5d':ab,ti OR 'eq-5d':ab,ti OR 'eq 5d':ab,ti OR rosser OR ((visual NEXT/1 analog*) AND (analog* NEXT/1 scale*)) OR ((patient OR self OR clinician OR observer OR investigator) NEAR/1 (reported OR assessed)) OR ('patient-reported' NEAR/2 outcome*) OR ('patient reported' NEAR/2 outcome*) OR pro:ab,ti OR suicide:ab,ti OR pain OR symptom OR depression:ab,ti OR disfigurement:ab,ti OR 'psycholog*':ab,ti OR social:ab,ti OR 'disfigur*':ab,ti OR 'productiv*':ab,ti OR psychosocial:ab,ti OR (patient* NEAR/2 (survey OR interview)) OR 'patient symptom*' OR 'fact':ab,ti OR 'functional assessment of cancer therapy' OR 'eortc qlq' OR 'european organization for research and treatment of cancer questionnaire' OR 'routine electronic monitoring of hrqol' OR remoqol OR 'functional assessment of cancer therapy-kidney symptom index' OR 'fact-ksi' OR 'prom' OR 'patient reported outcome measure*' OR coa OR 'clinical outcome* assessment*'	
• #3	• #1 AND #2 NOT (([animals]/lim NOT ([humans]/lim AND [animals]/lim) OR ([conference review]/lim OR [editorial]/lim OR [erratum]/lim OR [letter]/lim OR [note]/lim OR [review]/lim) NOT ([cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim))	773

*Table 140 Search **strategy for Cochrane database***

S. No.	Search terms	Hits
#1	MeSH descriptor: [von Hippel-Lindau Disease] explode all trees OR von Hippel Lindau Disease OR von Hippel Lindau Disease OR von Hippel Lindau syndrome OR von Hippel Lindau syndrome OR ("von" AND "hippel-lindau") OR ("von" AND "hippel" AND "lindau") OR ("hippel" AND "lindau") OR ("Angiomatosis Retinae") OR (Angiomatosis AND Retinae) OR ("Familial Cerebello-Retinal Angiomatosis") OR (Familial AND "Cerebello-Retinal" AND Angiomatosis)	34
#2	“hrqol” OR “hqol” OR MeSH descriptor: [Quality of Life] explode all trees OR ”quality of life” OR ”quality-of-life” OR ”qol” OR ”short form 36” OR ”sf36” OR ”sf-36” OR ”sf 36” OR ”short form 12” OR ”sf12” OR ”sf-12” OR ”sf 12” OR ”short form 6” OR ”sf6” OR ”sf-6” OR ”sf 6” OR ”euroqol” OR euro*qol OR ”eq5d” OR ”eq-5d” OR ”eq 5d” OR ”Rosser” OR (”visual” NEXT/1 analog*) AND (analog*	156,902

	NEXT/1 scale*) OR "fact-g" OR "eortc qlq-c30" OR "fact-c" OR "eortc qlq-cr29"	
#3	#1 AND #2	6

Table 141 Search strategy for MEDLINE®In-Process searched via PubMed® platform

S. No.	Search terms	Hits
#1	von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR "lindau tumor" OR Lindau Disease OR "Lindau syndrome" OR "Lindau Disease" OR Angiomatosis Retinae OR "Angiomatosis Retinae" OR "Lindau's syndrome" OR "Lindau's Disease" OR "VHL disease" OR "VHL Syndrome" OR "Hippel Lindau syndrome" OR "Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau Disease" OR Hippel-Lindau Disease OR "Hippel-Lindau Disease" OR "Familial Cerebello-Retinal Angiomatosis" OR Familial Cerebello-Retinal Angiomatosis OR "von Hippel-Lindau Disease"[Mesh]	6,608
#2	("effect" OR "impact") AND (social OR educational OR professional OR vocational OR psychosocial OR life OR living OR psychological OR cultural) OR "impact on daily living" OR (impact AND daily AND living) OR “hrqol” OR “hqol” OR "Quality of Life"[MeSH Terms] OR "quality of life” OR "health related quality of life”	1143123
#3	#1 AND #2	186

Table 142 Search strategy for clinicaltrials.gov

• S. No.	• Search terms	• Hits
• #1	• “Von Hippel-Lindau Disease”	• 38

Search strategy for the search conducted in July 2022

Table 143 Embase® and MEDLINE® search strategy for PRO’s and quality of life review

No.	Query	Results
#1	'von hippel lindau disease'/exp OR 'von hippel lindau disease'/syn OR 'hippel angiomatosis' OR 'hippel disease' OR 'hippel lindau disease' OR 'hippel lindau syndrome' OR 'hippel-lindau disease' OR 'lindau disease' OR 'lindau tumor' OR 'lindau tumour' OR 'von hippel disease' OR 'von hippel lindau syndrome' OR 'von hippel-lindau disease' OR vhl OR 'von hippel-lindau'	13,093

#2	((utilit* NEAR/2 (measure* OR outcome* OR state* OR health OR score* OR weight* OR analysis)):ab,ti) OR 'health utility index' OR 'hui':ab,ti OR 'hrqol':ab,ti OR 'hqol':ab,ti OR 'quality of life'/exp OR 'quality of life' OR 'quality-of-life'/exp OR 'quality-of-life' OR qol:ab,ti OR (utilit* NEXT/1 (score* OR value* OR evaluation*)) OR (health NEXT/2 utilit*) OR (('health'/exp OR 'health') AND (state NEXT/1 utilit*)) OR hui:ab,ti OR ((health NEXT/1 state*) AND (state* NEXT/1 preference*)) OR 'quality adjusted life year'/exp OR 'quality adjusted life year' OR 'quality adjusted life' OR ('quality adjusted' NEXT/1 survival*) OR qaly:ab,ti OR qald:ab,ti OR qale*:ab,ti OR qtime*:ab,ti OR 'disability adjusted life' OR daly*:ab,ti OR 'health RW '/exp OR 'health survey' OR hye*:ab,ti OR health*year*equivalent OR (health NEAR/2 utility*) OR 'wellbeing'/exp OR 'wellbeing':ab,ti OR (quality NEAR/2 well*being) OR qwb:ab,ti OR (willingness NEAR/2 pay) OR (standard NEAR/2 gamble) OR disutili*:ab,ti OR (time NEAR/2 trade*off) OR tto:ab,ti OR ('discrete choice' NEXT/1 experiment*) OR 'short form 36'/exp OR 'short form 36' OR 'sf36':ab,ti OR 'sf-36':ab,ti OR 'sf 36':ab,ti OR 'short form 12'/exp OR 'short form 12' OR 'sf12':ab,ti OR 'sf-12':ab,ti OR 'sf 12':ab,ti OR 'short form 6' OR 'sf6':ab,ti OR 'sf-6':ab,ti OR 'sf 6':ab,ti OR 'euroqol' OR 'euro-qol' OR 'euro qol' OR 'eq5d':ab,ti OR 'eq-5d':ab,ti OR 'eq 5d':ab,ti OR rosser OR ((visual NEXT/1 analog*) AND (analog* NEXT/1 scale*)) OR ((patient OR self OR clinician OR observer OR investigator) NEAR/1 (reported OR assessed)) OR ('patient-reported' NEAR/2 outcome*) OR ('patient reported' NEAR/2 outcome*) OR pro:ab,ti OR suicide:ab,ti OR pain OR symptom OR depression:ab,ti OR disfigurement:ab,ti OR 'psycholog*':ab,ti OR social:ab,ti OR 'disfigur*':ab,ti OR 'productiv*':ab,ti OR psychosocial:ab,ti OR (patient* NEAR/2 (survey OR interview)) OR 'patient symptom*' OR 'fact':ab,ti OR 'functional assessment of cancer therapy' OR 'eortc qlq' OR 'european organization for research and treatment of cancer questionnaire' OR 'routine electronic monitoring of hrqol' OR remoqol OR 'functional assessment of cancer therapy-kidney symptom index' OR 'fact-ksi' OR 'prom' OR 'patient reported outcome measure*' OR coa OR 'clinical outcome* assessment*'	5,043,672
#3	#1 AND #2	1,212
#4	#1 AND #2 AND [2020-2022]/py	230

Table 144 Search strategy for MEDLINE®In-Process searched via PubMed® platform

No.	Query	Results
1	von Hippel Lindau Disease OR Hippel Lindau Disease OR VHL Syndrome OR lindau tumor OR "lindau tumor" OR Lindau Disease OR "Lindau syndrome" OR "Lindau	6,959

	Disease" OR Angiomatosis Retinae OR "Angiomatosis Retinae" OR "Lindau's syndrome" OR "Lindau's Disease" OR "VHL disease" OR "VHL Syndrome" OR "Hippel Lindau syndrome" OR "Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau Disease" OR Hippel-Lindau Disease OR "Hippel-Lindau Disease" OR "von Hippel-Lindau Disease"[Mesh]	
2	(((renal cell carcinoma[MeSH Terms] OR "kidney carcinoma" OR "kidney tumor" OR "kidney adenoma" OR "kidney cancer") OR ((Renal OR kidney OR grawit* OR hypernephroid* OR nephroid*) AND (carcinoma OR cancer OR neoplasm OR adenoma OR pyelocarcinoma OR oncocytoma OR tumor OR tumour))) OR (hypernephroma OR "clear cell renal cell carcinoma")) OR (clear cell renal carcinoma[MeSH Terms])	232,928
3	"central nervous system" OR "CNS"[Title/Abstract] OR "CNS lesion" OR "CNS lesions" OR "hemangioblastoma"OR "hemangioblastomas" OR "cerebellar hemangioblastoma" OR "spinal cord hemangioblastoma" OR "spinal cord hemangioblastoma" OR "brainstem hemangioblastoma" OR "pancreas lesion" OR "pancreatic lesion" OR "Pancreas"[Title/Abstract] OR "Pancreas cyst" OR "pancreatic tumor" OR "pancreatic cystic lesion" OR "serous cystadenomas" OR "pancreatic cyst" OR "pancreatic cysts" OR "pancreatic neuroendocrine tumors" OR "pNET" OR "pNETs" OR "Necrotizing pancreatitis"	454,578
4	#1 AND (#2 OR #3)	3,971
5	"Value of Life"[mh] OR "Quality of Life"[mh] OR quality of life[tiab] OR "Quality-Adjusted Life Years"[mh] OR quality adjusted life[tiab] OR qaly*[tiab] OR qald*[tiab] OR qale*[tiab] OR qtime*[tiab] OR life year[tiab] OR life years[tiab] OR disability adjusted life[tiab] OR daly*[tiab] OR "sf36"[tiab] OR "sf 36"[tiab] OR "short form 36"[tiab] OR "shortform 36"[tiab] OR "short form36"[tiab] OR "shortform36"[tiab] OR "sf thirtysix"[tiab] OR "sfthirtysix"[tiab] OR "sfthirty six"[tiab] OR "sf thirty six"[tiab] OR "shortform thirtysix"[tiab] OR "shortform thirty six"[tiab] OR "short form thirtysix"[tiab] OR "short form thirty six"[tiab] OR "sf6"[tiab] OR "sf 6"[tiab] OR "short form 6"[tiab] OR "shortform 6"[tiab] OR "shortform6"[tiab] OR "short form6"[tiab] OR "sf6d"[tiab] OR "sf 6d"[tiab] OR "short form 6d"[tiab] OR "shortform 6d"[tiab] OR "sf six"[tiab] OR "sfsix"[tiab] OR "shortform six"[tiab] OR "short form six"[tiab] OR "sf8"[tiab] OR "sf 8"[tiab] OR "short form	995,102

	8"[tiab] OR "shortform 8"[tiab] OR "shortform8"[tiab] OR "short form8"[tiab] OR "sf eight"[tiab] OR "sfeight"[tiab] OR "shortform eight"[tiab] OR "short form eight"[tiab] OR "sf12"[tiab] OR "sf 12"[tiab] OR "short form 12"[tiab] OR "shortform 12"[tiab] OR "short form12"[tiab] OR "shortform12"[tiab] OR "sf twelve"[tiab] OR "sftwelve"[tiab] OR "shortform twelve"[tiab] OR "short form twelve"[tiab] OR "sf16"[tiab] OR "sf 16"[tiab] OR "short form 16"[tiab] OR "shortform 16"[tiab] OR "short form16"[tiab] OR "shortform16"[tiab] OR "sf sixteen"[tiab] OR "sfsixteen"[tiab] OR "shortform sixteen"[tiab] OR "short form sixteen"[tiab] OR "sf20"[tiab] OR "sf 20"[tiab] OR "short form 20"[tiab] OR "shortform 20"[tiab] OR "short form20"[tiab] OR "shortform20"[tiab] OR "sf twenty"[tiab] OR "sftwenty"[tiab] OR "shortform twenty"[tiab] OR "short form twenty"[tiab] OR "hql"[tiab] OR "hqol"[tiab] OR "h qol"[tiab] OR "hrqol"[tiab] OR "hr qol"[tiab] OR "hye"[tiab] OR "hyes"[tiab] OR "healthy year equivalent*"[tiab] OR "healthy years equivalent*"[tiab] OR "pqol"[tiab] OR "qls"[tiab] OR "quality of wellbeing"[tiab] OR "quality of well being"[tiab] OR "index of wellbeing"[tiab] OR "index of well being"[tiab] OR "qwb"[tiab] OR "nottingham health profile*"[tiab] OR "sickness impact profile"[tiab] OR "health status indicators"[mh] OR "health utilit*"[tiab] OR "health status"[tiab] OR "disutilit*"[tiab] OR "rosser"[tiab] OR "willingness to pay"[tiab] OR "standard gamble*"[tiab] OR "time trade off"[tiab] OR "time tradeoff"[tiab] OR "tto"[tiab] OR "hui"[tiab] OR "hui1"[tiab] OR "hui2"[tiab] OR hui3[tiab] OR eq[tiab] OR euroqol[tiab] OR euro qol[tiab] OR eq5d[tiab] OR eq 5d[tiab] OR euroqual[tiab] OR euro qual[tiab] OR duke health profile[tiab] OR functional status questionnaire[tiab] OR dartmouth coop functional health assessment*[tiab] OR (utilit*[tiab] AND (valu*[tiab] OR measur*[tiab] OR health[tiab] OR life[tiab] OR estimat*[tiab] OR elicit*[tiab] OR disease[tiab] OR score*[tiab] OR weight[tiab])) OR (preference*[tiab] AND (valu*[tiab] OR measur*[tiab] OR health[tiab] OR life[tiab] OR estimat*[tiab] OR elicit*[tiab] OR disease[tiab] OR score*[tiab] OR instrument[tiab] OR instruments[tiab]))	
6	#4 AND #5	77
7	#6 AND (inprocess[sb] OR pubstatusaheadofprint)	2

Table 145 Search strategy for Cochrane database

No.	Query	Results
#1	“kidney carcinoma” OR “kidney tumor” OR “kidney adenoma” OR “kidney cancer” OR (Renal*:ab,ti OR kidney*:ab,ti OR grawit*:ab,ti OR hypernephroid*:ab,ti OR nephroid*:ab,ti) AND (carcinoma*:ab,ti OR cancer*:ab,ti OR neoplasm*:ab,ti OR adeno*:ab,ti OR pyelocarcinoma*:ab,ti OR oncocytoma:ab,ti OR tumor*:ab,ti OR tumour*:ab,ti) OR 'hypernephroma':ab,ti OR rcc OR 'clear cell renal cell carcinoma' OR 'cc-rcc'	77,081
#2	MeSH descriptor: [Carcinoma, Renal Cell] explode all trees	1,030
#3	MeSH descriptor: [von Hippel-Lindau Disease] explode all trees	2
#4	von Hippel Lindau Disease OR "von Hippel Lindau Disease" OR "von Hippel Lindau syndrome" OR "von Hippel Lindau syndrome" OR ("von" AND "hippel-lindau") OR ("von" AND "hippel" AND "lindau") OR ("hippel" AND "lindau")	42
#5	'central nervous system'OR ‘CNS’:ab,ti OR ‘CNS lesion’ OR ‘CNS lesions’ OR 'hemangioblastoma' OR ‘hemangioblastomas' OR ‘cerebellar hemangioblastoma' OR ‘spinal cord hemangioblastoma' OR ‘spinal cord hemangioblastoma' OR ‘brainstem hemangioblastoma' OR ‘pancreas lesion’ OR ‘pancreatic lesion’ OR ‘Pancreas’:ab,ti OR ‘Pancreas cyst’ OR ‘pancreatic tumor’ OR ‘pancreatic cystic lesion’ OR ‘serous cystadenomas’ OR ‘pancreatic cyst’ OR ‘pancreatic cysts’ OR ‘pancreatic neuroendocrine tumors’ OR ‘pNET’ OR ‘pNETs’ OR ‘Necrotizing pancreatitis’	6,945
#6	#1 OR #2 OR #5	83,275
#7	#3 OR #4	42
#8	#6 AND #7	35
#9	MeSH descriptor: [Quality of Life] explode all trees	27,526
#10	("Quality of Life" OR "Quality-Adjusted Life Years" OR qaly* OR qald* OR "disability adjusted life" OR daly* OR "sf36" OR "short form 36" OR "shortform 36" OR "short form36" OR "shortform36" OR "sf thirtysix" OR "sfthirtysix" OR "sfthirty six" OR "sf thirty six" OR "shortform thirtysix" OR "shortform thirty six" OR "short form thirtysix" OR "short form thirty six" OR "sf6" OR "sf 6" OR "short form 6" OR "shortform 6" OR "shortform6" OR "short form6" OR "sf6d" OR "sf 6d" OR "short form 6d" OR "shortform 6d" OR "sf six" OR "sfsix" OR "shortform six" OR "short form six" OR "sf8" OR "sf 8" OR "short form 8" OR "shortform 8" OR "shortform8" OR "short form8" OR "sf eight" OR "sfeight" OR "shortform eight" OR "short form eight" OR "sf12" OR "sf 12" OR "short form 12" OR "shortform 12" OR "short form12" OR "shortform12" OR "sf twelve" OR "sftwelve" OR "shortform twelve" OR "short form twelve" OR "sf16" OR "sf 16" OR "short form 16" OR "shortform 16" OR "short form16" OR "shortform16" OR "sf sixteen" OR "sfsixteen" OR "shortform sixteen" OR "short form sixteen" OR "sf20" OR "sf 20" OR "short form 20" OR "shortform 20" OR "short form20" OR "shortform20" OR "sf twenty" OR "sftwenty"	165,388

No.	Query	Results
	OR "shortform twenty" OR "short form twenty" OR "hql" OR "hqol" OR "h qol" OR "hrqol" OR "hr qol" OR "hye" OR "hyes" OR "healthy year equivalent*" OR "healthy years equivalent*" OR "pqol" OR "qls" OR "quality of wellbeing" OR "quality of well being" OR "index of wellbeing" OR "index of well being" OR "qwb" OR "nottingham health profile*" OR "sickness impact profile" OR "health utilit*" OR "health status" OR "disutilit*" OR "rosser" OR "willingness to pay" OR "standard gamble*" OR "time trade off" OR "time tradeoff" OR "tto" OR "hui" OR "hui1" OR "hui2" OR hui3 OR eq OR euroqol OR euro qol OR eq5d OR eq 5d OR euroqual OR euro qual OR duke health profile OR functional status questionnaire OR dartmouth coop functional health assessment* OR (utilit* AND (valu* OR measur* OR health OR life OR estimat* OR elicit* OR disease OR score* OR weight)) OR (preference* AND (valu* OR measur* OR health OR life OR estimat* OR elicit* OR disease OR score* OR instrument OR instruments))):ti,ab,kw	
#11	MeSH descriptor: [Quality-Adjusted Life Years] explode all trees	1,286
#12	MeSH descriptor: [Health Status Indicators] explode all trees	23,272
#13	#9 OR #10 OR #11 OR #12	182,915
#14	#8 AND #13: with Cochrane Library publication date from Jan 2020 to July 2022	2

Study selection

Two independent reviewers manually screened all citations based on the title and abstract to identify potentially relevant studies. A third independent reviewer resolved any discrepancies in their decisions. After the first screening, the full texts of relevant studies were examined in more detail to determine studies eligible for the final inclusion. Two independent reviewers screened the full-text articles, and a third independent reviewer resolved discrepancies.

Key eligibility criteria

Key eligibility criteria used during the original literature review conducted in July 2020 are presented in Table 127.

Table 146: Eligibility Criteria (Search Date: July 2020)

Inclusion Criteria	Description

Population(s)	Patients with VHL syndrome (clinically diagnosed or genetically confirmed) including all tumor types
Interventions	No restriction on the interventions
Comparators	No restriction on the comparators
Outcomes	The inclusion of studies will not be restricted by outcome types
Time	No restrictions on timeframe
Study design	Studies reporting HRQoL and PROs
Other (Language)	No exclusion on language, non-English studies will be categorized separately and shared with Merck team

HRQoL: Health-related quality of life; PROs: Patient-reported outcomes; VHL: Von Hippel-Lindau

Key eligibility criteria used during the present update (July 2022) are presented in Table 147.

Table 147: Eligibility Criteria (Search Date: July 2022)

Inclusion Criteria	**Description**
Population(s)	All patients with VHL disease-associated RCC, CNS Hb, and pNET. VHL disease should be determined with genetically confirmed VHL germline mutation or clinically diagnosed VHL disease. In absence of a clear reporting about the diagnosis method, studies that mention 'VHL disease' will also be included. Studies that did not report VHL diagnostic criteria and included patients with a hereditary/familial cancer (RCC, pNET, CNS Hb) will be included only if there is an overwhelming majority of the patients with VHL disease.
Interventions	No restriction on the interventions
Comparators	No restriction on the comparators
Outcomes	Utility/disutility data associated with disease and adverse events including EQ-5D, time to trade off, standard gamble, etc. PRO instruments Endpoint hierarchy Symptoms of interest Duration of therapy PRO completion rate Methods used to analyze PROs HRQoL data • PRO data
Time	July 2020 to July 2022
Study design	No restriction on study design provided relevant outcomes are reported

Other (Language)	No exclusion on language, non-English studies will be categorized separately and shared with Merck team

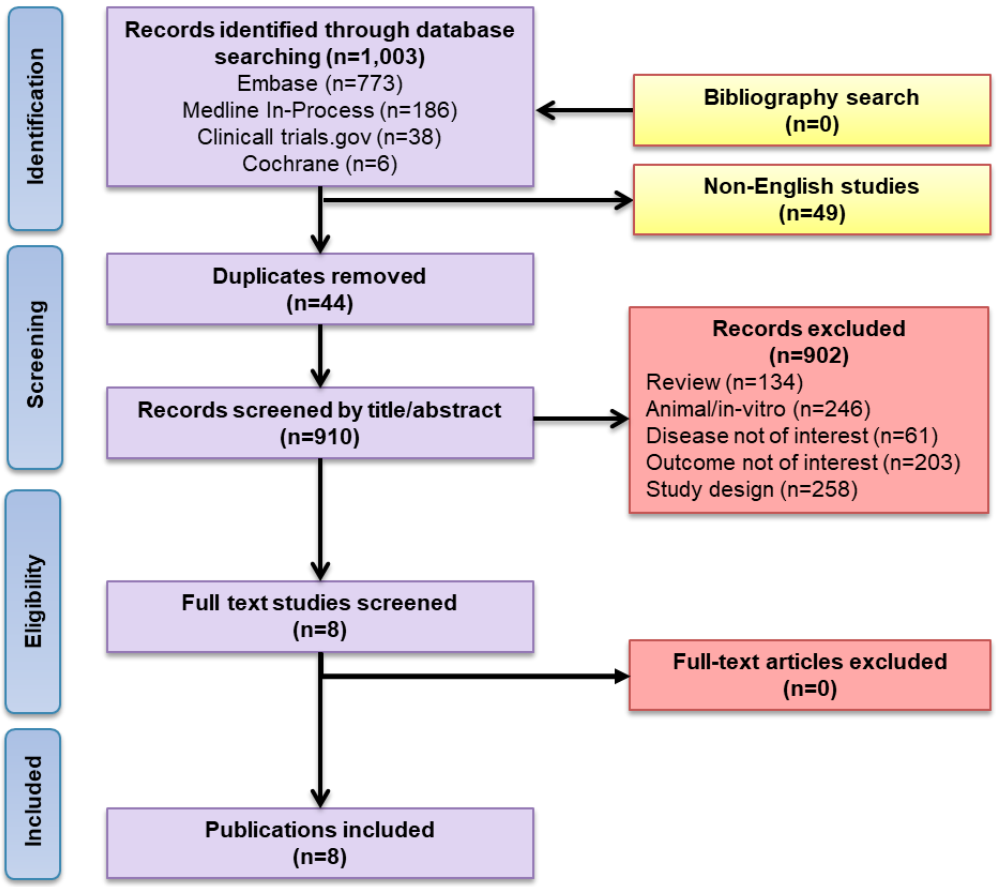
CNS: Central nervous system; EQ-5D: EuroQol Questionnaire - 5 Dimensions; Hb: Hemangioblastoma; HRQoL: Health-related quality of life; pNET: Pancreatic neuroendocrine tumor; PRO: Patient-reported outcomes; RCC: Renal cell carcinoma; VHL: Von Hippel-Lindau

H1.3 Results

Study Selection and PRISMA flow diagram

In July 2020, 1,003 unique records were identified, following removal of duplicates and non-English studies, 910 records were screened by title and abstract and of those, eight met the inclusion criteria and are reported in the present document. Among these, majority of the studies (n=6) were conducted across Europe, while the remaining two studies were conducted in Australia and Mexico (Figure 38).

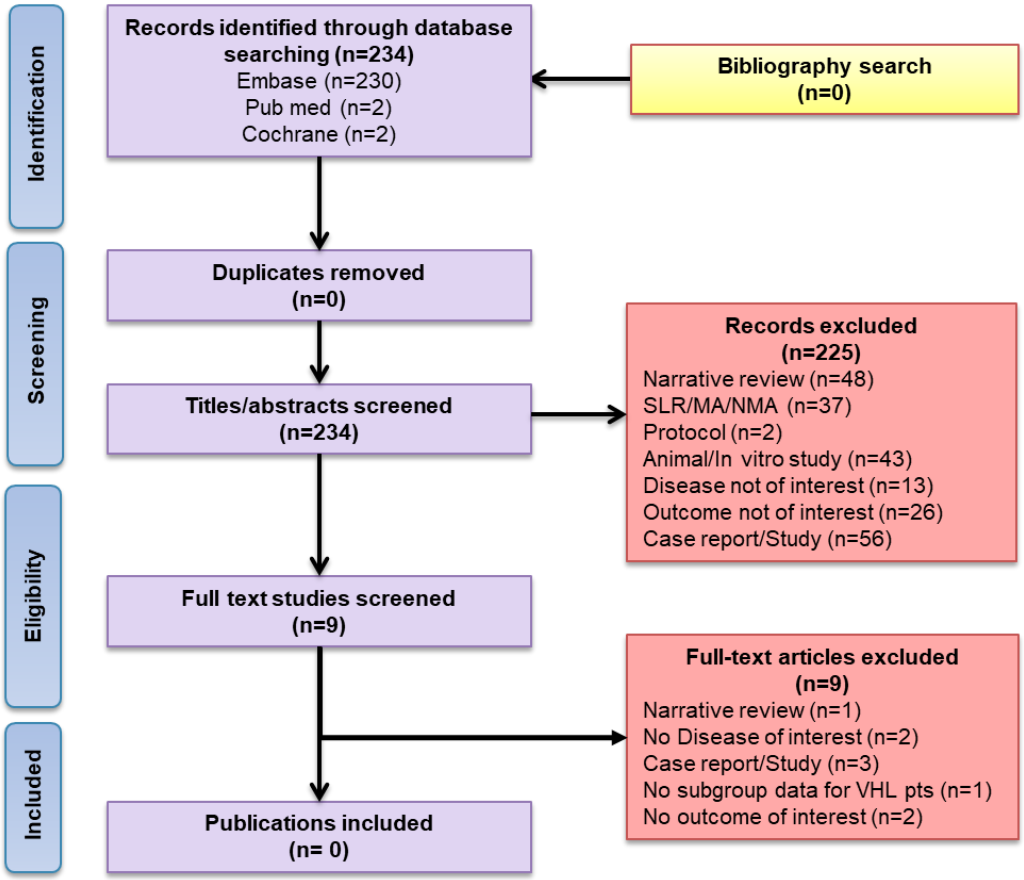
Figure 38: PRISMA Flow and Study Selection (Search Date: July 2020)



Embase: Excerpta Medica Database; Medline: Medical Literature Analysis and Retrieval System Online

As part of the 2022 update, a total of 234 records were retrieved through electronic literature searches conducted over Embase (n = 230), Medline (n = 2), and Cochrane (n = 2), no deduplicates were identified. Following the screening of all 234 studies, 9 studies were selected for full-time review, following which none met the pre-defined eligibility criteria, and consequently, no additional studies were included in the update (Figure 39).

Figure 39: PRISMA Flow and Study Selection (Search Date: July 2022)



Embase: Excerpta Medica Database; MA: Meta-analysis; NMA: Network meta-analysis; SLR: Systematic literature review; VHL: Von Hippel-Lindau

Study Characteristics

Table 148 summarizes the characteristics of the included studies. Majority of the studies provided QoL data for overall VHL syndrome rather than focusing on specific tumor types, except two studies that reported QoL data for CNS hemangioblastomas (Siller, Szelényi et al. 2017, Rochette, Baumstarck et al. 2018). A total of four studies reported QoL data in patients with VHL disease while the remaining four studies reported QoL data in patients with VHL and their partners, caregivers or other family members. A wide variety of QoL scales were utilized across the reported studies, including: SF-36 Health Survey, Barthel Index (BI), ODI, State-Trait Anxiety Inventory (SF-STAI), Hospital Anxiety and Depression Scale (HADS), Impact of Event Scale (IES), and Cancer Worry Scale (CWS).

Table 148: Study Characteristics

Study name	Study design and setting	Publication type	Country	Sample size	Outcomes assessed	QoL scales utilized
(Siller, Szelényi et al. 2017)	Single-center study	Journal article	Germany	24	QoL, General performance evaluations	PSI, KPS, BI, ODI, SF-36 v2 Health Survey
(Lammens, Bleiker et al. 2009)	Nationwide, cross-sectional study	Journal article	The Netherlands	117	Psychosocial issues and attitudes of VHL family members towards PGD	Self-report questionnaire
(Fraser, Watts et al. 2007)	Descriptive study	Journal article	England and Scotland	n=54 at one stop clinic n=18 at ad hoc clinic	Symptom distress, psychological morbidity and quality of care experienced at One stop and Ad Hoc clinics	SF-STAI, HADS, the VHL Worry Scale, National Outpatient Customer Satisfaction Survey
(Rochette, Baumstarck et al. 2018)	Longitudinal study	Journal article	France	12	Psychosocial impact of VHL genetic testing	Self-reported questionnaire that assessed: anxiety,

						depression, and QoL; SF-36 questionnaire, PCQ
(Rasmussen, Alonso et al. 2010)	Retrospective study	Journal article	Mexico	109	The uptake of diagnostic and pre-symptomatic genetic testing	Sociodemographic and psychological tests
(Lammens 2010)	Nationwide, cross-sectional study	Journal article	The Netherlands	123	VHL-related worry and distress, and Health-related QoL	IES, CWS, SF-36 Health Survey; PCS
(Lammens, Bleiker et al. 2011)	Nationwide, cross-sectional study	Journal article	The Netherlands	50	VHL-related psychological worry and distress	IES, CWS, SF-36 Health Survey
(Kasparian, Rutstein et al. 2015)	Qualitative study	Journal article	Australia	N=23 (15 patients and 8 caregivers)	Patients' and caregivers' experiences in relation to VHL	IES, HADS, Brief COPE, Caregiver burden scale

*if not reported, address mentioned in corresponding author's academic affiliations was used as surrogate

BDI: Beck Depression Inventory; BI: Barthel Index; CWS: Cancer Worry Scale; HADS: Hospital Anxiety and Depression Scale; FAD: Family Functioning Scale; NR: Not reported; IES: Impact of Event Scale; KPS: Karnofsky Performance Scale; ODI: Oswestry Disability Index; PCQ: Psychological Consequences Questionnaire; PCS: Physical Component Summary; PGD: Pre-implantation genetic diagnosis; PSI: Patient Satisfaction Index; QoL: Quality of life; SF-STAI: State-Trait Anxiety Inventory; VHL: Von Hippel-Lindau

Patient characteristics

Table 149: Patient Characteristics

Study name	Patient characteristics	VHL diagnosis	Country	Sample size	Organ involvement	Age (years)	Sex (male)
(Siller, Szelényi et al. 2017)	Sporadic and VHL disease patients (n=2.4:1) underwent microsurgical treatment for intraspinal Hb	NR	Germany	24 (17 were included in the study)	Spinal cord	Mean at first surgery (range) 36.8 (15–62)	12 (50%)
(Lammens, Bleiker et al. 2009)	Family members with a hereditary cancer predisposition (VHL disease = 48)	Known germline mutation	The Netherlands	117	NR	Mean (SD) 39.9 (14)	91 (51%)
(Fraser, Watts et al. 2007)	Patients aged ≥ 18 years with VHL disease or were confirmed carriers	Clinically diagnosed or clinically confirmed	England and Scotland	n=54 at one stop clinic n=18 at ad hoc clinic	CNS Hb, RA, ccRCC, pheochromocytoma, PNET, pancreatic cyst, epididymal cyst	Range: 18-76	32 (45%)
(Rochette, Baumstarck et al. 2018)	Patients underwent surgery for CNS Hb	Clinically diagnosed or clinically confirmed	France	12	CNS Hb	Median (IQR) 58 (44–68)	6 (50%)

(Rasmussen, Alonso et al. 2010)	Families with VHL mutation	Clinical diagnosis	Mexico	109	CNS Hb, kidney cysts, pancreatic cysts, epididymal cysts, RCC, pheochromocytoma, ELST, pancreatic cancer	Median (range) 20.5 (1-45)	NR
(Lammens 2010)	Families with VHL mutation (carriers, 50% at-risk, non-carriers)	NR	The Netherlands	123	NR	Mean (SD) 41 (14)	19 (42%)
(Lammens, Bleiker et al. 2011)	partners of individuals diagnosed with or at high risk of Li Fraumeni syndrome or VHL disease	NR	The Netherlands	50	NR	Mean (range) Partners: 42.2 (21-71) Spouse: 42.0 (20-60)	Partners: 52% Spouse: 56%
(Kasparian, Rutstein et al. 2015)	Families of VHL patients	NR	Australia	N=23 (15 patients and 8 caregivers	CNS Hb, RA, RCC, pheochromocytoma, ELST	NR	NR

ccRCC: Clear cell renal cell carcinoma; CNS: Central nervous system; ELST: Endolymphatic sac tumor; Hb: Hemangioblastoma; IQR: Inter quartile range; NR: Not reported; pNET: Pancreatic neuroendocrine tumor; RA: Retinal angioma; RCC: Renal cell carcinoma; SD: Standard deviation; VHL, von Hippel Lindau

Study details

Some inconsistencies in the reported data can be observed as most of the results were qualitative. Given the small number of studies included, short paragraphs summarizing individual studies are provided below.

Quality of life of patients who had undergone microsurgical resection of spinal cord hemangioblastomas

Siller and colleagues conducted a single-center study in Germany evaluating 24 patients who had undergone microsurgical resection of spinal cord hemangioblastomas. Among the 24 patients, seven patients (29.2%) met the criteria for VHL disease, (including six patients with multiple CNS hemangioblastomas, one patient with a family history of VHL disease, one with pancreatic cysts, one with multiple kidney cysts, and one with both pheochromocytoma and renal cell carcinoma). However, all patients were analyzed in a single group as no significant difference in epidemiological and clinical characteristics was observed between patients with sporadically developing tumor or as part of VHL (Siller, Szelényi et al. 2017).

A total of 17 (70.8%) patients were available for long-term follow-up interviews and questionnaires regarding quality of life and general performance evaluations (including the Patient Satisfaction Index [PSI], WHO/ECOG Performance Status Scale, Karnofsky Index [also known as the Karnofsky Performance Scale [KPS], BI, ODI, and SF-36 v2 Health Survey). Mean follow-up was 7.9 ± 4.0 years (range 1.1–14.1 years). Results showed a high rate of satisfaction for surgical outcome, with 12 patients (70.6%) reporting a PSI score of 1 and one patient (5.9%) reporting a PSI score of 2. However, one patient with a PSI score of 3 and three patients with a score of 4 would not undergo surgery again. Long-term performance in everyday life was excellent in 12 patients (70.6%) with a BI of 100, indicating no disability. Furthermore, 15 patients (88.2%) showed a WHO/ECOG Performance Status grade ≤1, and 13 patients (76.5%) had a KPS score ≥80, indicating an ability to carry on normal activities and to work with no need of special care during the long-term follow-up. The mean ODI score on the long-term questionnaire was 11.4% ± 12.5% (range 0%–32%), showing only minimal disability. However, three patients (17.6%)

reported moderate disability (BI: 60–95) and two patients (11.8%) reported severe disability (BI: < 60) in carrying out daily living activities. At the same time, two patients (11.8%) had a WHO/ECOG Performance Status Grade 4 and one patient (5.9%) had a KPS score < 50, indicating severe disability and the need for institutional or hospital care. According to the SF-36v2 Health Survey, the mean physical health composite score (PCS) on the long-term questionnaire evaluation was 49.1 ± 9.1 (range 30.3–61.7) and the mean mental health composite score (MCS) was 45.5 ± 10.9 (range 32.5–68.6), indicating slightly worse physical and mental health than the average population (Siller, Szelényi et al. 2017).

QoL results from this study showed that resection provided satisfying long-term outcome with excellent performance in everyday life (more than three-quarters of the patients having a BI ≥ 95, WHO/ECOG Performance Status grade ≤ 1, and KPS score ≥ 80, while the mean long-term ODI score indicated only minimal disability). Although, VHL is a risk factor for a poorer long-term functional performance, four of 17 patients stated that they would not undergo surgery again for the same results (Siller, Szelényi et al. 2017).

Psychosocial impact of VHL genetic testing in patients with previously diagnosed CNS hemangioblastomas

Rochette and colleagues conducted a longitudinal study to assess the psychosocial impact of VHL genetic testing in patients with previously diagnosed CNS hemangioblastomas at a hospital in France. Patients were asked to complete a self-reported questionnaire that assessed: anxiety, depression, and quality of life during the initial visit (visit 1) and just before the final visit (visit 3). Level of anxiety was assessed with the State-Trait Anxiety Inventory (STAI) scale (including 20 items assessing trait anxiety [general anxiety level] and 20 items assessing state anxiety [anxiety level for a specific time]) ranging between 20–80, with higher scores corresponding to higher levels of anxiety. Depression was assessed using the Beck Depression Inventory (BDI) with scores ranging from 0–39, with higher scores indicating worsening depression (score 0–<4: no depression, 4–<8: mild depression, 8–<30: moderate depression; score >30: severe depression). The QoL was

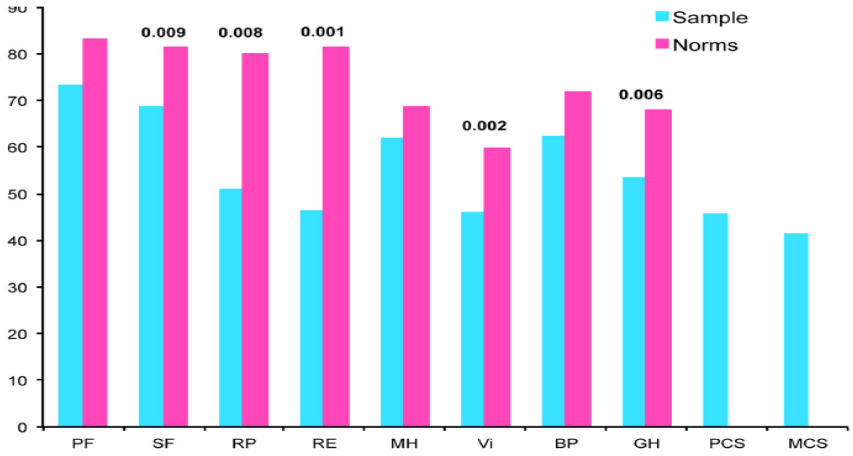
assessed using the SF-36 questionnaire, a generic questionnaire containing eight subscales (physical function, social functioning, role physical, role emotional, mental health, vitality, bodily pain, and general health, with scores ranging from 0 indicating low QoL level to 100 indicating high QoL level). Two component summary measures of SF-36, namely PCS and MCS were calculated. The Psychological Consequences Questionnaire (PCQ, part 1, negative consequences of screening) was used to evaluate the psychological consequences of the genetic screening (Rochette, Baumstarck et al. 2018).

A total of 24 patients were included in the study. At baseline, the mean state anxiety level was 42.9 (standard deviation [SD]±10.1) and the mean trait anxiety level was 39.8 (SD±14.4). According to the BDI, 48%, 30%, and 22% of patients presented with no, mild, or moderate mood disorders, respectively. The lowest and highest scores were role-physical and physical functioning, respectively. Compared with French age-sex adjusted control cohort, significantly lower scores for social functioning, role-physical, role-emotional, vitality, and general health dimensions were reported (Figure 40). The mean (+/−SD) PCQ was 4.8 (±1.8).

For the 12 patients who attended visit 3, the state anxiety level did not differ between the two evaluation times, but the trait anxiety level significantly decreased from a score of 41 at baseline, to 38 at visit 3. Two scores of QoL, vitality and general health dimensions, were significantly improved between baseline and visit 3. The mood disorder score and the PCQ score did not differ between the two evaluations (Table 150)

The state anxiety level at visit 3 was significantly correlated to the anxiety levels and mood disorder levels at baseline, the mental health dimension, and the mental composite score of QoL at baseline, while the trait anxiety level at visit 3 was only significantly correlated to the ‘physical-like’ QoL dimension at visit 1 (bodily pain and physical composite score of SF-36) (Rochette, Baumstarck et al. 2018).

Figure 40: Quality of Life Scores of the Study Population, and of French Age-Sex Crossed Norms



BP: Bodily pain; G: General health; MCS: Mental component summary; MH: Mental health; PCS: Physical component summary; PF: Physical functioning; RE: Role emotional; RP: Role physical; SF: Social functioning; Vi: Vitality

Table 150: Changes in Psychological Profile in Patients Who Attended Visit 3

		Baseline Visit 1 N=12	**Communication of results Visit 3 N=12**	**P-value***
Anxiety (Spielberger)	Trait (20-80)*	33.0 [25.3;40.5]	36 [23.8;40.0]	0.722
	State (20-80)*	42.5 [31.9;47.0]	40.5 [27.5;45.8]	0.028
Mood disorders (BDI)	Total score (0-39)**	4.5 [1.3;6.0]	2.5 [1.3;5.0]	0.572
Quality of life (SF-36)***	Physical functioning	85 [72.5;95]	90 [56.3;95.0]	0.287
	Social functioning	56.3 [50;97.0]	81.3 [62.5; 100]	0.135
	Role physical	25 [0;91.7]	87.5 [6.0; 100]	0.078
	Role emotional	41.7 [0;91.7]	100 [8.3;100]	0.206
	Mental health	64.0 [48.0;83.0]	72.0 [57.0;88.0]	0.133

	Vitality	45.0 [35.0;60.0]	65.0 [36.3;75.0]	0.005
	Bodily pain	61.5 [43.5;81.5]	73.0 [46.0;81.5]	0.760
	General health	57.0 [37.8;79.5]	77.0 [52.0;82.0]	0.026
	Mental composite score	43.6 [30.8;52.9]	48.8 [39.9;52.1]	0.060
	Physical composite score	47.8 [38.6;51.9]	49.3 [39.0;58.5]	0.754
Psychological consequences of testing (PCQ)	Total score (0-10)****	4.6 [3.6-5.4]	4.0 [2.8-4.7]	0.168

*Higher score signifies higher anxiety level
** Higher score signifies higher mood disorder level
*** Higher score signifies better quality of life
****higher score signifies higher negative psychological consequence
BDI: Beck depression inventory; SF-36: Short-form 36; PCQ: Psychological consequences of screening

Some VHL patients were clinically anxious and depressed irrespective of the screening service received

Fraser and colleagues conducted a descriptive study comparing two types of screening service One stop clinic (OSC; n=54 from four clinics) and *ad hoc* clinic (AH; n=18 from four clinics) at eight regional genetics centers in England and Scotland using a questionnaire-based approach. The questionnaire included measures of perceived VHL disease severity and symptom distress, psychological morbidity and the quality of care experienced at the screening clinic attended. Psychological morbidity was assessed with the SF-STAI; scores ranging from 6 to 24 (score ≥ 18 = 'clinical case') and the Hospital Anxiety and Depression Scale (HADS; including two subscales: anxiety and depression), each with a maximum score of 21 (0–7 normal, 8–10 mild, 11–14 moderate, 15–21 severe; ≥11 = 'clinical case'). Symptom distress was assessed with a modified version of the Breast Cancer Worry Scale adapted for VHL; scores ranging from 6–24. Patients' perceptions of the quality of care provided were assessed with items from the National Outpatient Customer Satisfaction Survey: three items from the access/timeliness sub-scale, five

from the patient education/information sub-scale and all items from the emotional support and co-ordination of care subscales, giving a maximum total score of 18.

The reliability of measures of psychological distress and symptom distress exhibited high internal consistency: HADS α = 0.9377, SF-STAI α = 0.9055 and VHL Worry Scale = 0.9258. Levels of worry about VHL symptoms were similar between the OS and AH groups. Median anxiety scores on both the SF-STAI and HADS anxiety subscale were higher in the OSC group; however, these results were not statistically significant. Depression scores in the OSC screening group were significantly higher on the HADS depression subscale (P=0.03) compared with the AH group. There was a positive correlation between volunteers' subjective ratings of disease severity and anxiety scores on the SF-STAI (P<0.05), VHL worry scores (P<0.0001) and also the HADS scores (P<0.0001), where an increased subjective disease rating indicates higher levels of psychological distress. Clinician's objective disease severity ratings were found to be positively correlated with the HADS scores (P<0.05) and the VHL worry scores (P<0.05), but not the SF-STAI scores (P=0.33). Using a cut-off value of 11 on the HADS anxiety and depression subscales 29% of volunteers were defined as exhibiting moderate levels of clinical anxiety and 13% moderate levels of depression. A larger proportion of these were from the OSC group. Similarly, using a cut-off value of 18 on the SF-STAI, 25% were defined as clinically anxious. Again, a larger proportion of these were from the OSC group. The scores for quality of health care between the OSC and AH groups did not differ significantly (median difference: 1 [95% CI: 0, 3]).

Results showed that patients screened at OS clinics received more comprehensive surveillance and attend more regularly than patients screened at AH clinics. Over a quarter (29%) of respondents were clinically anxious and 13% clinically depressed. Psychological morbidity was higher in the OSC group, evidenced by significantly higher depression scores (P=0.03).

Regarding factors contributing to the levels of depression, disease severity per se was an unlikely candidate, as levels were similar across both screening groups. One possible explanation is that a person who knows they are carrying the mutated gene identified as causing VHL in their family may become more depressed about the risk

to their future health. Significantly more confirmed mutation carriers were noted in the OSC group (*P*=0.05) where levels of depression were higher (Fraser, Watts et al. 2007).

Patients with significant pre-test anxiety tended to prematurely abandon the surveillance program

Rasmussen and colleagues conducted a five-year follow-up of a series of patients (n=109) that underwent presymptomatic genetic testing for VHL disease to evaluate the uptake of diagnostic and pre-symptomatic

genetic testing. In order to identify the factors influencing their adherence to a long-term follow-up program for hereditary cancer the clinical, psychological and socio-demographic features of the mutation carriers were analyzed. The complete set of sociodemographic and psychological tests was available for 17 adults. Patients who were symptomatic before the molecular test were five times more likely to continue the surveillance program (OR = 5; 95% CI: 1.2 - 20.3; *P*=0.02), which was maintained even after adjustment for the clustering of observations (OR = 5.0, CI 95%= 1.37- 18.29; *P*= 0.02) while significant pre-test anxiety was more common amongst the individuals that prematurely dropped out of surveillance (64.7% vs. 35.3%; *P* = 0.01). Follow-up was not found to be associated with having or not having children, the mutation status or affectedness of the children, and there was also no relationship with pre-test depression.

Most socio-demographic and psychiatric variables analyzed did not correlate with follow-up adherence. However, individuals who were already symptomatic at the time of testing had a higher likelihood of adhering to long-term follow-up (OR = 5; 95% CI 1.2 - 20.3; *P*=0.02), and those who had significant pre-test anxiety tended to abandon the follow-up program (*P*=0.01) (Rasmussen, Alonso et al. 2010).

Attitude of VHL family members towards pre-implantation genetic diagnosis

Lammens and colleagues conducted a nationwide, cross-sectional study evaluating the psychosocial issues and attitudes of VHL and LFS family members towards pre-implantation genetic diagnosis (PGD), and to identify characteristics significantly

associated with a positive attitude towards its use in VHL in the Netherlands. Questions on PGD were posed to the high-risk family members (proven carriers, clinical diagnosis, individuals at 50% risk) ≥16 years of age, and to their partners. Patients were asked to complete a self-report questionnaire including questions on sociodemographic, personal and family medical history, psychosocial variables, and attitude towards the use of PGD. Respondents were asked 'Would you consider the use of PGD if this would be/ would have been available to you?'' Additionally, respondents were asked to rate the advantages and disadvantages of PGD. In total, 117 respondents from 36 VHL families completed the questionnaire. None of the participants had used PGD. Thirty-five percent of the 129 VHL/LFS family members indicated that they would consider the use of PGD if this would be/would have been a possibility for them, 27% were uncertain, and 38% would not use PGD. Approximately one-third of the family members and their partners expressed a positive attitude towards the use of PGD. A current desire to have children was related significantly to a positive attitude towards PGD, with those with such plans being more likely to express an intention to use PGD than those without such plans (48 vs.25%, respectively; *P*=0.01). Individuals within the childbearing age range tended to have a more positive attitude towards PGD than those over 40 years of age (41 vs. 26%, respectively; *P*=0.10), as did those without vs. with children (43 vs. 29%, respectively; *P*=0.09). None of the medical (e.g., personal history of VHL/LFS, number of affected first degree relatives) or psychosocial variables (e.g., cancer worries, syndrome-related distress, feelings of guilt towards (future) children) were associated significantly with attitude towards PGD. Of the 50 partners, one-third would consider the use of PGD if this would be/would have been a possibility for them, 11 (22%) were not sure, and 22 (45%) would not use PGD. None of the sociodemographic or medical variables were associated significantly with attitudes toward PGD, although male partners tended to be more positive than female partners (44 vs. 21%, respectively; *P*=0.08). Consistent with the high-risk family members, none of the psychosocial variables were significantly associated with a positive attitude towards PGD. The most frequently rated perceived advantage of PGD was avoiding the possibility of a selective pregnancy termination (32%). The most frequently rated disadvantage of PGD was the fact that the long-term effects of PGD are unknown (18%) (Lammens, Bleiker et al. 2009).

Clinically relevant levels of VHL-related distress in VHL families

Lammens conducted a cross-sectional study in the Netherlands in 48 families with a VHL germline mutation between August 2006 and February 2008. Participants completed a self-reported questionnaire including questions on sociodemographic, personal and family medical history, and psychosocial variables, including VHL-related worry and distress, and Health-related QoL. The intrusion subscale (a 7-item questionnaire assessing intrusive thoughts and feelings about VHL during the past seven days) of the Impact of Event Scale (IES) was used to assess VHL-specific with scores ranging between 0–35. A score between 0–8 is defined as no to low levels of distress (no additional help indicated), 9-18 as moderate levels of distress (some additional specialized psychosocial help may be indicated), and 19 or higher as high levels of distress (professional psychosocial help is indicated). A score of 9 or higher is considered clinically relevant. VHL-related worries were assessed with an 8-item questionnaire adapted from the Cancer Worry Scale (CWS) to refer to VHL-related tumors. Scores range from 8–32, with higher scores indicating more frequent worries about VHL. HRQoL was assessed with the SF-36 Health Survey, composed of eight multi-item scales. Scale scores range from 0–100, with higher scores indicating higher levels of functioning and wellbeing. Respondents were asked whether there was at least one person in their social network with whom they could share personal problems. Furthermore, impact of VHL on the family was measured with the 12-item General Family Functioning (GFF) subscale of the McMaster Family Functioning Scale (FAD). The total GFF score ranged between 12 and 48, with a higher score indicating better general family functioning (Lammens 2010).

In total, 123 family members (72%) from 37 families completed the questionnaire. Sixty-eight individuals were carriers of a VHL germline mutation. Sixteen individuals were at 50% risk and 39 were proven non-carriers. Sixty-six family members were diagnosed with/treated for one or more VHL manifestations. The mean score on the ‘intrusion’ subscale of the IES was 7.7 (SD = 8.5). Thirty-nine percent (n=48) of the respondents scored 9 or higher on the IES, indicating clinically relevant, moderate to

severe levels of VHL-related distress. Thirteen percent (n=16) scored above 19, indicating severe levels of distress. The mean score on the CWS was 13.9 (SD=4.8). Overall, 38% of the carriers expressed frequent concerns about the possibility of developing (an additional) VHL tumor, and 46% expressed concerns about the possibility that they would again require surgery due to VHL (Table 151). Forty-one percent of the total sample reported frequent worries about the chance of family members developing a VHL-related tumor (carriers 35%; at-risk 44%; non-carriers 50%) (Lammens 2010).

Table 151: VHL Related Worries

	Total group (n=123)	**(a)symptomatic carrier group (n=68) %**	**At-risk group (n=16)**	**Non-Carrier group**	**p-value**
1. How often have you thought about your chances of getting a tumor (again)?	15	25	0	3	.00
2. Have these thoughts affected your mood?	8	15	0	0	.01
3. Have these thoughts interfered with your ability to do daily activities?	4	6	6	0	.32
4. How concerned are you about the possibility of getting a tumor one day?	25	38	19	5	.00
5. How often do you worry about developing a tumor?	20	28	19	5	.01
6. How much of a problem is this worry?	8	12	6	3	.29
7. How often do you worry about the chance of family members developing a tumor?	41	35	44	50	.33

8. How concerned are you about the possibility that you will ever need surgery (again)?	27	46	13	0	.00

Factors associated with VHL-related distress included (Table 152), female gender (OR 4.1; 95% CI 1.5 - 10.6), actual risk, a high perceived risk of developing tumors, and having experienced the death of a close relative due to VHL during adolescence were associated significantly with clinically relevant levels of VHL-related distress (OR: 11.2; 95% CI 1.4 - 86.9) (Lammens 2010).

Table 152: Variables Associated with VHL Related Worries (CWS) and Distress (IES)

Variable	**CWS, mean (SD)**	**IES Low distressed, n (%) (n=75)**	**IES Moderate to high distressed, n (%) (n=48)**
Actual risk Carrier	15.6 (5.3)**	36 (53%)*	32 (47%)*
50% at-risk	13.5 (2.9)**	14 (88%)*	2 (12%)*
Non-carrier	10.9 (2.3)**	25 (64%)*	14 (36%)*
Age °	-.087	41.6 ± 14.3	38.9 ±15.1
Gender male	13.9 (4.4)	39 (76%)**	12 (24%)**
female	13.9 (5.1)	36 (50%)**	36 (50%)**
Level of education low	14.6 (5.7)	20 (56%)	16 (44%)
moderate	13.7 (4.3)	38 (59%)	27 (41%)
high	13.2 (4.5)	17 (77%)	5 (23%)
Children	13.8 (5.3)	47 (60%)	31 (40%)

yes			
no	13.9 (3.8)	28 (62%)	17 (38%)
Diagnosed with/treated for VHL yes	16.0 (5.2)**	39 (59%)	27 (41%)
no	11.4 (2.6)**	36 (63%)	21 (37%)
Social support yes	13.4 (4.6)*	64 (65%)	35 (35%)
no	15.7 (5.1)*	10 (44%)	13 (56%)
Social constraints 0	.283**	16.1 ± 3.3	16.8 ± 2.6
Risk perception of developing a tumor low	11.9 (3.5)**	31 (63%)**	18 (37%)**
moderate	12.9 (4.1)**	28 (78%)**	8 (22%)**
high	17.9 (4.7)**	14 (41%)**	20 (59%)**
Knowledge of VHL 0	.104	6.7 ± 2.3	6.4 ± 1.9
Family functioning 0	-.026	30.2 ± 2.2	29.9 ± 2.3
Received psychosocial support yes	14.9 (4.9)	18 (53%)	16 (47%)
no	13.5 (4.7)	57 (64%)	32 (36%)
1st degree relative diagnosed with VHL no VHL	13.9 (6.2)	9(69%)	4(31%)
childhood (< 13 y)	14.4 (4.7)	28 (58%)	20 (42%)
adolescence (13-20 y)	14.4 (8.6)	11 (52%)	10 (48%)
adulthood (> 20 y)	12.8 (3.8)	26 (65%)	14 (35%)
Nr of 1st degree relatives with VHL none	12.9 (7.0)	6 (75%)	2 (25%)

1-2	13.7 (4.0)	32 (62%)	20 (38%)
3 or more	14.3 (5.0)	33 (58%)	24 (42%)
1st degree relative died due to VHL			
no loss	13.3 (4.4)**	38 (65%)*	20 (35%)*
childhood (< 13 y)	13.8 (4.2)**	12 (67%)*	6 (33%)*
adolescence (13-20 y)	18.6 (5.5)**	2 (20%)*	8 (80%)*
Nr of 1st degree relatives who died of VHL			
none	13.5 (4.4)	37 (65%)	20 (35%)
1-2	14.9 (5.2)	30 (56%)	24 (44%)
3 or more	11.3 (2.7)	6 (60%)	4 (40%)

* statistically significant associated (p ≤ .05), ** statistically significant associated (p ≤ .01)
º For these variables Pearson correlations are reported for the CWS and mean scores and SD's for the IES.
CWS: Cancer worry scale; IES: Impact of event scale; SD: Standard deviation; VHL: Von Hippel-Lindau

Individuals diagnosed with/treated for VHL, with a higher level of social constraints, with a high perceived risk of developing tumors, and those who experienced the loss of a close relative due to VHL during adolescence were significantly more prone to experience higher levels of VHL-related worries. In total, this set of variables explained 45% of the variance (R^2) in VHL-related worries (Table 153).

Table 153: Linear Regression Analysis VHL Related worries

Variable	**B (SE) (unstandardized)**	**p-value**
Actual risk		
Non-carrier		
At-risk	2.2 (1.3)	0.08
Carrier	1.8 (1.4)	0.19

Diagnosed with/treated for VHL (no)	-2.5 (1.1)	0.02
Social support (yes)	-1.1 (0.91)	0.25
Social constraints	0.26 (0.12)	0.03
Perceived risk of developing a tumor		
Low		
Moderate	-1.6 (1.1)	0.13
High	2.5 (1.2)	0.04
1st degree relative died due to VHL		
No loss		
Childhood (<13 y)	-0.12 (1.0)	0.91
Adolescence (13-20 y)	3.2 (1.3)	0.02
Adulthood (>20 y)	1.4 (.81)	0.09

The total variance explained (R^2) with this model is 45%
B (SE) refers to comparison with the category listed first
SE: Standard error; VHL: Von Hippel-Lindau

Mean scores on the SF-36 scales, and on the PCS and MCS scales were comparable to those of the age- and gender-matched reference group from the general Dutch population, with the exception of the 'general health' scale, where the VHL family members scored significantly worse (*P*=0.007). The carrier group scored significantly worse than the non-carrier and at-risk groups on the 'general health perceptions' scale. Additionally, the carrier group also scored significantly worse on the 'physical functioning' and the 'role limitations due to physical health problems' scales, as compared with the non-carrier group. The carrier group also scored significantly worse on the PCS scale (*P*= 0.003). Twenty-eight percent (n=34) of the respondents had received specialized professional psychosocial support (33% of those who were moderately to severely distressed). Carriers had more frequently received professional psychosocial support than those at 50% risk and non-carriers (*P*=0.04) (Lammens 2010).

Partners of high-risk VHL individuals experienced heightened levels of psychological distress

In the same nationwide, cross-sectional study, Lammens and colleagues evaluated the prevalence of LFS- or VHL-related psychological worry and distress among partners of individuals diagnosed with or at high risk for LFS and VHL; (2) factors associated significantly with LFS- and VHL-related worry and distress; (3) the impact of LFS and VHL on HRQoL; (4) the impact of LFS and VHL on the spousal relationship; and (5) the need for and use of professional psychological support. Study participants were asked to complete a self-reported questionnaire, including questions on sociodemographic, personal and family medical history, and psychosocial variables, including LFS- or VHL-related worry and distress, and HRQoL. Similar to other studies conducted by Lammens, this study utilized IES scale to measure VHL specific distress. VHL related worries were measured using CWS and HRQoL was utilized using SF-36 health survey.

This study focuses on the partners of individuals clinically and/or genetically diagnosed with (symptomatic

and asymptomatic carriers) or at 50% risk for VHL referred to as high-risk spouses. In total, 91% (N=50) of the partners who were approached to participate in the study completed the self-report questionnaire. Approximately two-thirds of the respondents were partners of VHL high-risk spouses. There were approximately equal number of men and women, with the partners' mean age 42.2 years; SD=13.1 years). Most of the respondents had children (60%), and most were partners of symptomatic family members (68%). The mean score for the partners on the intrusion subscale of the IES was 6.74 (SD=8.06). Twenty eight percent (n=14) scored 9 or higher, indicating clinically relevant, moderate-to-severe levels of LFS- or VHL-related distress. Twelve percent (n=6) scored above 19, indicating severe levels of distress. The mean score of the partners on the CWS was 13.9 (SD=3.5). The mean scores of the high-risk spouses and their partners on the IES and CWS did not differ significantly and were statistically significantly correlated (r=0.386 and 0.550), respectively (Table 154).

Table 154: Psychosocial Impact of LFS-and VHL Among Partners and Their High-risk Spouses

	Partners, Mean (SD)	**High-risk spouses, Mean (SD)**	**t**	**t**
SF-36[a]				
Physical functioning	91 (18.6)	87.8 (19.3)	1.11	0.425**
Role functioning physical	94 (22.3)	76.0 (37.4)	3.25**	0.221
Bodily pain	86.7 (20.9)	80.5 (25.3)	1.40	0.086
General health	77.8 (18.5)	63.8 (22.9)	4.14**	0.347**
Vitality	74.6 (18.2)	65.6 (20.7)	2.67**	0.259
Social functioning	91.8 (18.0)	85.5 (21.3)	1.66	0.064
Role functioning emotional	90.7 (22.4)	80.0 (36.3)	1.97	0.211
Mental health	84.1 (13.7)	77.8 (17.2)	2.36*	0.259
CWS[b]	13.9 (3.5)	13.4 (3.5)	1.09	0.550*
IES`	6.7 (8.1)	6.7 (7.9)	0.02	0.386*
Syndrome-related distress, IES[d] IES(binary),IES IES[d]				
Low	34 (72%)	32 (68%)	N/A	N/A
Moderate-to-high	14 (28%)	16 (32%)	N/A	N/A

[a]SF-36 measuring health-related quality of life; Possible range for the SF-36 scales: 0 (poor functioning) to 100 (best functioning).
[b]CWS measuring syndrome-related worries: a higher score indicates a higher level of worries.
[c]IES measuring syndrome-related distress: a higher score indicates a higher level of distress.
[d]IES as a dichotomous variable: low score 0–8 and moderate to high score 9 or higher.
*Statistically significant at the 0.05 level
**Statistically significant at the 0.01 level
CWS: Cancer worry scale; IES: Impact of event scale; SD: Standard deviation; SF-36: Short-form 36; VHL: Von Hippel-Lindau

Younger age and the distress level of the high-risk spouse were significantly associated with partners' levels of LFS- or VHL-related distress. The high-risk

spouses' level of distress (OR 1.1; 95% CI 1.03–1.24;p=0.01), and the partners' age (OR 0.93; 95% CI 0.87–0.99; p=0.02) were associated significantly with the partners' LFS- or VHL-related distress at the multivariate level.

In general, the partners' mean scores on most of the SF-36 scales were comparable to those of an age and

gender-matched sample from the general Dutch population apart from 'bodily pain', 'role functioning physical', and 'mental health' scales, where the partners scored significantly better than the comparison group. The partners' scores were significantly better than those of the high-risk spouses for four of the eight SF-36 scales: namely the 'role functioning physical', 'VT', 'general health' and 'mental health' scales.

A total of 14% (n=7) of the partners indicated that LFS/VHL had had a negative influence on their relationship. Relationship problems mentioned included: (1) difficulty in talking about the syndrome and its impact (n=2); stress due to feelings of nervousness and anxiety, especially around the time of test results (n=2); arguments about non-adherence to recommended screening programs (n=1); and practical problems surrounding frequent hospitalization. Conversely, 52% (n=26) of the partners stated that LFS/VHL had had a positive effect on their relationship. Positive effects mentioned most often were that the situation had taught them to enjoy and appreciate life and each other more, (n=12), and that it had brought them closer together (n=6). Overall, 76% believed that professional psychosocial support should be offered routinely, not only to those at high risk but also to their partners (Lammens, Bleiker et al. 2011).

Emotional, social, and practical challenges of VHL patients and their caregivers

Kasparian and colleagues conducted a qualitative study in Australia to explore both patients' and caregivers' experiences in relation to VHL. A semi-structured interview was utilized covering several life domains such as self-identity and self-esteem, interpersonal relationships, education and career opportunities, family communication, physical health and emotional well-being, and supportive care needs. Quantitative measures were also used to examine the prevalence of anxiety, depression, and disease-specific distress in this sample utilizing IES, HADS, BRIEF COPE. Additionally, VHL-related clinical characteristics and screening practices where participants were asked about their (or their family member's) diagnosis, genetic testing, current tumors, family history of disease and VHL-related deaths, and screening practices (14 items). Beliefs about and barriers to screening (e.g., 'I do not believe that screening applies to me') were assessed via 16 items, with response options ranging from 1 ('Strongly agree') to 5 ('Strongly disagree'). Furthermore, caregivers also completed the 12-item Caregiver Burden Scale, with scores >17 indicating significant caregiving-related burden.

A total of 23 participants (15 patients, 8 caregivers; a response rate of 75%) were included. The mean age of patients was 37 years (SD=14.5; range: 18–59 years), and the mean age of caregivers was 57 years (SD=16.9; range: 37–75 years). The majority of patients were men (n=9, 60%), whereas five caregivers were mothers, one a father, and three the wife or long-term partner of a person with VHL. Almost all caregivers (n=8, 89%), and just over half of patients (n=8, 53%), were married. Eight patients (53%) had children (M=12.8 years, range: 0–25 years).

Medical aspects of living with VHL

- Although all patients perceived regular screening as effective in managing VHL and detecting tumors at an early stage, it was considered as a necessary yet anxiety-provoking burden
- Majority of the participants described ongoing difficulties in attempting to locate interested, knowledgeable general practitioners to provide personalized medical care

Genetic testing, family planning, and perceptions of preimplantation genetic diagnosis

- Most patients reported wanting their children to have genetic testing for VHL either at birth (n=3) or before the age of 10 years (n=10). Of the eight patients with children, six had requested genetic testing for their children (mean age of child at the time of genetic test=10.0 years, SD=7.8, range: 0–21 years), and two had unaffected children by utilizing preimplantation genetic diagnosis. For four patients, their diagnosis had made them less willing to have children in the future

Coping with VHL:

Greater variations were observed with day-to-day coping with VHL. Some participants experienced difficulties coping with the consequences of VHL, with reminders of the disease ever-present in their lives. Participants with more severe symptoms felt that VHL weighed heavily on their life choices and opportunities. In contrast, other participants viewed VHL as having a minor influence on their lives

For other patients it was important to look to the future and to focus on their family, as well as other aspects of their life

Strategies of acceptance (M=3.39, SD=0.57), emotional support (M=2.47, SD=1.14), planning (M=2.45, SD=1.13), and positive reframing (M=2.21, SD=1.11) were the most highly endorsed coping methods, whereas denial (M=1.11, SD=0.36), behavioral disengagement (M=1.21, SD=0.48), and self-blame (M=1.32, SD=0.65) were the least endorsed

Role of family in living with VHL:

Most participants mentioned partner support as most important and beneficial

Family support was also perceived as important, particularly for younger patients. Most participants felt it was easier to talk about VHL with family members who carried a VHL mutation, compared with those who did not

Some participants described feelings of transmission guilt with an intense sense of blame about having passed VHL onto one's children

Perceived financial and career limitations:

Some participants perceived their livelihood as heavily influenced by the medical challenges associated with VHL. Career choices and opportunities were often viewed differently once symptoms started to manifest

Although most patients had not encountered workplace discrimination relating to their condition (n=14), many remained hesitant to disclose their diagnosis to employers and colleagues. The need to communicate to employers about VHL was also perceived as a significant barrier to career progression

*Psychological consequences of VHL (*Psychological well-being)

Of the 19 participants (14 patients, 5 caregivers) who completed the IES, one caregiver reported symptoms indicative of significant traumatic stress warranting clinical assessment

Of the 13 patients who completed the HADS, six patients (46%) reported anxiety, and two patients reported (15%) depressive symptoms, warranting clinical assessment. The anxiety symptom, 'I feel restless, as if I have to be on the move' was most strongly endorsed, with seven patients (54%) experiencing this symptom 'quite a lot' or 'very much indeed' in the past week. Of the five caregivers who completed the HADS, one caregiver reported anxiety, and another caregiver reported depressive symptoms, indicative of the need for clinical assessment.

For some, physical impairments associated with VHL had resulted in the loss of particular social, occupational, or sporting activities, which in turn often resulted in the loss of friendships and a narrowing of one's social network (Kasparian, Rutstein et al. 2015).

H1.3 References for HRQoL SLR

Fraser, L., S. Watts, A. Cargill, S. Sutton and S. Hodgson (2007). "Study comparing two types of screening provision for people with von Hippel-Lindau disease." Fam Cancer **6**(1): 103-111.

Kasparian, N. A., A. Rutstein, U. M. Sansom-Daly, S. Mireskandari, J. Tyler, J. Duffy and K. M. Tucker (2015). "Through the looking glass: an exploratory study of the lived experiences and unmet needs of families affected by Von Hippel-Lindau disease." Eur J Hum Genet **23**(1): 34-40.

Kim, E. and S. Zschiedrich (2018). "Renal Cell Carcinoma in von Hippel-Lindau Disease-From Tumor Genetics to Novel Therapeutic Strategies." Front Pediatr **6**: 16.

Lammens (2010). "Psychosocial impact of Von Hippel-Lindau disease: levels and sources of distress." Clin Genet.

Lammens, C., E. Bleiker, N. Aaronson, A. Vriends, M. Ausems, M. Jansweijer, A. Wagner, R. Sijmons, A. van den Ouweland, R. van der Luijt, L. Spruijt, E. Gómez García, M. Ruijs and S. Verhoef (2009). "Attitude towards pre-implantation genetic diagnosis for hereditary cancer." Fam Cancer **8**(4): 457-464.

Lammens, C. R., E. M. Bleiker, S. Verhoef, M. G. Ausems, D. Majoor-Krakauer, R. H. Sijmons, F. J. Hes, E. B. Gómez-García, T. A. Van Os, L. Spruijt, R. B. van der Luijt, A. M. van den Ouweland, M. W. Ruijs, C. Gundy, T. Nagtegaal and N. K. Aaronson (2011). "Distress in partners of individuals diagnosed with or at high risk of developing tumors due to rare hereditary cancer syndromes." Psychooncology **20**(6): 631-638.

Rasmussen, A., E. Alonso, A. Ochoa, I. De Biase, I. Familiar, P. Yescas, A. L. Sosa, Y. Rodríguez, M. Chávez, M. López-López and S. I. Bidichandani (2010). "Uptake of genetic testing and long-term tumor surveillance in von Hippel-Lindau disease." BMC Med Genet **11**: 4.

Reference, G. "Genetics Home Reference. Von Hippel-Lindau syndrome: Genetics Home Reference; 2020 [cited 2020 May]. Available from: https://ghr.nlm.nih.gov/condition/von-hippel-lindau-syndrome#statistics.".

Rochette, C., K. Baumstarck, H. Canoni-Zattara, A. E. Abdullah, D. Figarella-Branger, M. Pertuit, A. Barlier, F. Castinetti, K. Pacak, P. Metellus and D. Taïeb (2018). "Psychological impact of von Hippel-Lindau genetic screening in patients with a previous history of hemangioblastoma of the central nervous system." J Psychosoc Oncol **36**(5): 624-634.

Schmid, S., S. Gillessen, I. Binet, M. Brändle, D. Engeler, J. Greiner, C. Hader, K. Heinimann, P. Kloos, W. Krek, I. Krull, S. J. Stoeckli, M. C. Sulz, K. van Leyen, J. Weber, C. Rothermundt and T. Hundsberger (2014). "Management of von hippel-lindau disease: an interdisciplinary review." Oncol Res Treat **37**(12): 761-771.

Siller, S., A. Szelényi, L. Herlitz, J. C. Tonn and S. Zausinger (2017). "Spinal cord hemangioblastomas: significance of intraoperative neurophysiological monitoring for resection and long-term outcome." J Neurosurg Spine **26**(4): 483-493.

Appendix I: Cost and healthcare resource identification, measurement and valuation

Please see Appendix G: Published cost-effectiveness studies above for the economic burden SLR which includes cost and healthcare resource identification, measurement and valuation.

Appendix J: Clinical outcomes, base-case analysis inputs and disaggregated results from the model

J1.1 Clinical outcomes from the model

The only outcome from the MK-6482-004 trial that is directly modelled is time to surgery (TTS). Please see *Validation of transition probabilities* in B.3.3 Clinical parameters and variables for full validation of modelled transition probabilities and long-term extrapolations of OS.

J1.2 Summary of base-case analysis inputs

Table 155 Summary of variables applied in the economic model

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Time horizon & discounting			
Time horizon (years)	59.0	N/A (fixed values)	B.3.2 Economic analysis
Discount rate, annual - costs	3.5%	N/A (fixed values)	B.3.2 Economic analysis
Discount rate, annual - effectiveness	3.5%	N/A (fixed values)	B.3.2 Economic analysis
Patient characteristics			
Starting age (years)	41.0	N/A (fixed values)	Patient population
Percent female	47.5%	N/A (fixed values)	Patient population
Weight (kg)	79.7	N/A (fixed values)	Patient population

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Body surface area (m^2)	1.9	N/A (fixed values)	Patient population
Per cycle health state TPs			
Pre-surgery to surgery (VHL-RCC)	SOC: 90% - immediate transition 10% - 0.003772	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000554	Normal	Clinical efficacy: transition probabilities
Pre-surgery to metastatic disease (VHL-RCC)	SOC: 0.001187	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000174	Normal	Clinical efficacy: transition probabilities
Pre-surgery to death (VHL-RCC)	SOC: 0.000116	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000071	Normal	Clinical efficacy: transition probabilities
Pre-surgery to surgery (VHL-CNS Hb)	SOC: 100% - immediate transition	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000100	Normal	Clinical efficacy: transition probabilities
Pre-surgery to metastatic disease (VHL-CNS Hb)	SOC: 0.001123	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000056	Normal	Clinical efficacy: transition probabilities
Pre-surgery to death (VHL- CNS Hb)	SOC: 0.000276	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000145	Normal	Clinical efficacy: transition probabilities
Pre-surgery to surgery (VHL-pNET)	SOC: 90% - immediate transition 10% - 0.000167	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000006	Normal	Clinical efficacy: transition probabilities
Pre-surgery to metastatic disease (VHL-pNET)	SOC: 0.002676	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000099	Normal	Clinical efficacy: transition probabilities
Pre-surgery to death (VHL-pNET)	SOC: 0.000208	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.000120	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after	SOC: 0.0000589	Normal	Clinical efficacy: transition probabilities

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
surgery to metastatic disease (VHL-RCC)	Belzutifan: 3.223E-06	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after surgery to metastatic disease (VHL-CNS Hb)	SOC: 0.000055	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 3.89994E-06	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after surgery to metastatic disease (VHL-pNET)	SOC: 0.00013	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 1.74907E-06	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after surgery to death (VHL-RCC)	SOC: 0.00023	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.00023	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after surgery to death (VHL-CNS Hb)	SOC: 0.00028	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.00028	Normal	Clinical efficacy: transition probabilities
Exponential rate of event-free after surgery to death (VHL-pNET)	SOC: 0.00021	Normal	Clinical efficacy: transition probabilities
	Belzutifan: 0.00021	Normal	Clinical efficacy: transition probabilities
Exponential rates of OS and PFS failure with treatments of advanced RCC			
Sunitinib, OS	0.00398	Normal	Clinical efficacy: transition probabilities
Sunitinib, PFS	0.01436	Normal	Clinical efficacy: transition probabilities
No active treatment, OS	0.0076	Normal	Clinical efficacy: transition probabilities
No active treatment, PFS	0.02746	Normal	Clinical efficacy: transition probabilities
Tivozanib, HR of OS vs. sunitinib	1.33	Log-normal	Clinical efficacy: transition probabilities
Tivozanib, HR of PFS vs. sunitinib	1.19	Log-normal	Clinical efficacy: transition probabilities
Pazopanib, HR of OS vs. sunitinib	0.92	Log-normal	Clinical efficacy: transition probabilities
Pazopanib, HR of PFS vs. sunitinib	1.05	Log-normal	Clinical efficacy: transition probabilities

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Cabozantinib, HR of OS vs. sunitinib	0.8	Log-normal	Clinical efficacy: transition probabilities
Cabozantinib, HR of PFS vs. sunitinib	0.48	Log-normal	Clinical efficacy: transition probabilities
Nivolumab / ipilimumab, HR of OS vs. sunitinib	0.72	Log-normal	Clinical efficacy: transition probabilities
Nivolumab / ipilimumab, HR of PFS vs. sunitinib	0.89	Log-normal	Clinical efficacy: transition probabilities
Avelumab / axitinib, HR of OS vs. sunitinib	0.8	Log-normal	Clinical efficacy: transition probabilities
Avelumab / axitinib, HR of PFS vs. sunitinib	0.69	Log-normal	Clinical efficacy: transition probabilities
Nivolumab / cabozantinib, HR of OS vs. sunitinib	0.60	Log-normal	Clinical efficacy: transition probabilities
Nivolumab / cabozantinib, HR of PFS vs. sunitinib	0.51	Log-normal	Clinical efficacy: transition probabilities
Pembrolizumab / lenvatinib, HR of OS vs. sunitinib	0.66	Log-normal	Clinical efficacy: transition probabilities
Pembrolizumab / lenvatinib, HR of PFS vs. sunitinib	0.39	Log-normal	Clinical efficacy: transition probabilities
Exponential rates of OS and PFS failure with treatments of advanced pNET			
Streptozocin / 5-fluorouracil, OS	0.00398	Normal	Clinical efficacy: transition probabilities
Pembrolizumab / axitinib, PFS	0.01436	Normal	Clinical efficacy: transition probabilities
No active treatment, OS	0.0076	Normal	Clinical efficacy: transition probabilities
No active treatment, PFS	0.02746	Normal	Clinical efficacy: transition probabilities
Everolimus, HR of OS vs. no active treatment	0.35	Log-normal	Clinical efficacy: transition probabilities

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Everolimus, HR of PFS vs. no active treatment	0.35	Log-normal	Clinical efficacy: transition probabilities
Sunitinib, HR of OS vs. no active treatment	0.42	Log-normal	Clinical efficacy: transition probabilities
Sunitinib, HR of PFS vs. no active treatment	0.42	Log-normal	Clinical efficacy: transition probabilities
Interferon a2B, HR of OS vs. no active treatment	0.37	Log-normal	Clinical efficacy: transition probabilities
Interferon a2B, HR of PFS vs. no active treatment	0.37	Log-normal	Clinical efficacy: transition probabilities
Lanreotide, HR of OS vs. no active treatment	0.46	Log-normal	Clinical efficacy: transition probabilities
Lanreotide, HR of PFS vs. no active treatment	0.46	Log-normal	Clinical efficacy: transition probabilities
Octreotide, HR of OS vs. no active treatment	0.46	Log-normal	Clinical efficacy: transition probabilities
Octreotide, HR of PFS vs. no active treatment	0.46	Log-normal	Clinical efficacy: transition probabilities
Medical management costs by health state			
Pre-surgery cost per week (£)	VHL-RCC: 11.95	Gamma	Health state unit costs and resource use
	VHL-CNS Hb: 11.95	Gamma	Health state unit costs and resource use
	VHL-pNET: 11.95	Gamma	Health state unit costs and resource use
Event-free after surgery cost per week (£)	VHL-RCC: 11.95	Gamma	Health state unit costs and resource use
	VHL-CNS Hb: 14.16	Gamma	Health state unit costs and resource use
	VHL-pNET: 11.95	Gamma	Health state unit costs and resource use
Social care cost of CNS Hb progression in the VHL-CNS Hb cohort (per week)	VHL-CNS Hb: 20.86	Gamma	Health state unit costs and resource use

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
Entry into metastatic disease state one-time cost (£)	VHL-RCC: 2036.84	Gamma	Health state unit costs and resource use
	VHL-pNET: 459.97	Gamma	Health state unit costs and resource use
Pre-progression metastatic disease state cost per week (£)	VHL-RCC: 62.73	Gamma	Health state unit costs and resource use
	VHL-pNET: 77.31	Gamma	Health state unit costs and resource use
Post-progression metastatic disease state cost per week (£)	VHL-RCC: 62.73	Gamma	Health state unit costs and resource use
	VHL-pNET: 77.31	Gamma	Health state unit costs and resource use
Terminal care one-time cost (£)	7220.05	Gamma	Health state unit costs and resource use
Drug Administration Costs			
Unit cost of simple IV drug administration (£)	361.53	Gamma	Intervention and comparators' costs and resource use
Unit cost of complex IV drug administration (£)	426.80	Gamma	Intervention and comparators' costs and resource use
Unit cost of oral drug dispensing (£)	526.52	Gamma	Intervention and comparators' costs and resource use
Relative dose intensity			
Belzutifan	[redacted]	Normal	Intervention and comparators' costs and resource use
Cost of adverse events			
Belzutifan cost per cycle (£)	46.62	Gamma	Intervention and comparators' costs and resource use
Utilities and disutilities			
Utility in pre-surgery, surgery, and event-free after surgery states (with CR)	0.86757	Beta	Health-related quality-of-life data used in the cost-effectiveness analysis
Utility in pre-surgery, surgery, and event-free after surgery states (with PR/SD)	0.754	Beta	Health-related quality-of-life data used in the cost-effectiveness analysis
Utility in pre-surgery, surgery,	0.665	Beta	

Variable	Mean Value	Measurement of uncertainty and distribution: confidence interval (distribution)	Reference to section in submission
and event-free after surgery states (with PD)			Health-related quality-of-life data used in the cost-effectiveness analysis
Utility of pre-progression metastatic disease	0.525	Beta	Health-related quality-of-life data used in the cost-effectiveness analysis
Utility of post-progression metastatic disease	0.412	Beta	Health-related quality-of-life data used in the cost-effectiveness analysis
Disutility from AEs	-0.06417	Normal	Health-related quality-of-life data used in the cost-effectiveness analysis
Disutility associated with age	-0.00026	Normal	Health-related quality-of-life data used in the cost-effectiveness analysis
Disutility associated with age^2	-3.3E-05	Normal	Health-related quality-of-life data used in the cost-effectiveness analysis

AE: adverse event; CNS: central nervous system; CR: complete response; Hb: haemoangioblastoma; HR: hazard ratio; IV: intravenous; kg: kilogram; OS: overall survival; PFS: progression free survival; pNET: pancreatic neuroendocrine tumour; PD: progressed disease; PR: partial response; RCC: renal cell carcinoma; SD: stable disease; SOC: standard of care; VHL: Von Hippel Lindau

J1.3 Disaggregated results of the base-case incremental cost-effectiveness analysis

Table 156 Base-case disaggregated costs and effectiveness

Outcomes	VHL-RCC cohort		VHL-CNS Hb cohort		VHL-pNET cohort	
	Belzutifan	SOC	Belzutifan	SOC	Belzutifan	SOC
Costs (£)						
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Belzutifan treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for other tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Effectiveness						
Quality-adjusted life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for other tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutlity	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR: complete response; PD: progressive disease; PR: partial response; SD: stable disease; SOC: standard of care

Table 157 Summary of QALY gain by health state – VHL-RCC cohort

Health state	QALY intervention (belzutifan)	QALY comparator (SOC)	Increment	Absolute increment	% absolute increment
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 158 Summary of QALY gain by health state – VHL-CNS Hb cohort

Health state	QALY intervention (belzutifan)	QALY comparator (SOC)	Increment	Absolute increment	% absolute increment
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 159 Summary of QALY gain by health state – VHL-pNET cohort

Health state	QALY intervention (belzutifan)	QALY comparator (SOC)	Increment	Absolute increment	% absolute increment
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 160 Summary of costs by health state – VHL-RCC cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 161 Summary of costs by health state – VHL-CNS Hb cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 162 Summary of costs by health state – VHL-pNET cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 163 Summary of predicted resource use by category of cost – VHL-RCC cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for other tumors	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 164 Summary of predicted resource use by category of cost – VHL-CNS Hb cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for other tumors	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 165 Summary of predicted resource use by category of cost – VHL-pNET cohort

Health state	Cost intervention (belzutifan)	Cost comparator (SOC)	Increment	Absolute increment	% absolute increment
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Adverse event costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumor	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for other tumors	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

J1.4 Full DSA and scenario analysis results

Table 166 Tabular DSA and scenario analysis results for VHL-RCC cohort

	ICER vs. comparator: Belzutifan vs. SOC	
	(£/QALY)	
	Low input value	*High input value*
Base case	42,997	
High/low DSAs		
Efficacy and transition probabilities		
Proportion receiving immediate RCC surgery under SOC in VHL-RCC cohort: 80-100%	46,474	39,863
Proportion receiving immediate CNS Hb surgery or sequelae under SOC in VHL-CNS Hb cohort: 90-100%	42,997	42,997
Proportion receiving immediate pNET surgery under SOC in VHL-pNET cohort: 80-100%	42,997	42,997
Exponential rates of EF→MD under belzutifan +/- 10%	43,006	42,988
Exponential rates of EF→MD under SOC +/- 10%	42,980	43,015
Exponential rates of EF→Death under belzutifan +/- 10%	43,015	42,980
Exponential rates of EF→Death under SOC +/- 10%	42,742	43,245
Perioperative mortality risks +/- 10%	42,937	43,058
Exponential rates of OS and PFS failure with treatments for advanced RCC and pNET +/- 10%	42,441	43,476
Drug acquisition and administration costs		
Unit costs of IV drug administration +/- 10%	42,998	42,996
Unit cost of oral drug dispensing +/- 10%	43,002	42,992
Surgery, surgical complication, disease management, and AE costs		
Medical management costs in pre-surgery state +/- 10%	42,943	43,051
Medical management costs in event-free after surgery state +/- 10%	43,026	42,968
Social care costs of CNS Hb progression in the VHL-CNS Hb cohort +/- 10%	42,997	42,997
Costs of surgical procedures +/- 10%	43,087	42,908
Costs of short-term surgical complications +/- 10%	43,070	42,925

Costs of long-term surgical complications +/- 10%	44,621	41,373
Medical management costs in pre-progression metastatic disease state +/- 10%	43,003	42,991
Medical management costs in post-progression metastatic disease state +/- 10%	43,008	42,987
Terminal care costs (one-time cost) +/- 10%	43,006	42,988
Costs of AEs +/- 10%	42,997	42,998
Utilities		
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	46,992	56,659
Utility in pre-progression metastatic disease state +/- 10%	43,042	42,953
Utility in post-progression metastatic disease state +/- 10%	43,085	42,912
Disutilities of short-term complications +/- 10%	42,984	43,010
Disutilities of long-term complications +/- 10%	62,697	42,362
Disutility from AEs +/- 10%	42,996	42,998
Scenario analyses		
Perspective		
Include indirect costs (societal perspective)	35,597	
Time horizon and annual discount rate		
Time horizon: 20 years	57,096	
Time horizon: 30 years	59,623	
Annual discount rate: 0.0%	43,153	
Annual discount rate: 1.5%	42,642	
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	39,535	
Efficacy and transition probability scenarios		
Distribution for pre-surgery→surgery in the belzutifan arm (VHL-RCC cohort): Gamma	44,781	
Do not adjust surgery and metastases rates to account for real-world standard of care	63,178	
Drug cost scenarios		
Distribution for belzutifan ToT: Weibull	53,685	
Assume no treatment effect waning: Model efficacy and ToT separately	15,831	
Do not apply relative dose intensity	47,944	
Do not allow vial-sharing	42,987	
Include costs of first-line advanced regimens only	43,257	

Utility/disutility scenarios	
Assume same utility for CR as PR/SD	43,359
Apply caregiver disutility	41,141
Do not apply age-adjusted disutility	41,290
Do not apply AE-related disutility	42,985

*Note: The x1.7 severity weight is applied to the incremental QALYs in the ICER calculation.

Table 167 Tabular DSA and scenario analysis results for VHL-CNS Hb cohort

	ICER vs. comparator: Belzutifan vs. SOC	
	(£/QALY)	
	Low input value	*High input value*
Base case	33,490	
High/low DSAs		
Efficacy and transition probabilities		
Proportion receiving immediate RCC surgery under SOC in VHL-RCC cohort: 80-100%	33,490	33,490
Proportion receiving immediate CNS Hb surgery or sequelae under SOC in VHL-CNS Hb cohort: 90-100%	35,302	33,490
Proportion receiving immediate pNET surgery under SOC in VHL-pNET cohort: 80-100%	33,490	33,490
Exponential rates of EF→MD under belzutifan +/- 10%	33,491	33,490
Exponential rates of EF→MD under SOC +/- 10%	33,485	33,495
Exponential rates of EF→Death under belzutifan +/- 10%	33,493	33,488
Exponential rates of EF→Death under SOC +/- 10%	33,254	33,719
Perioperative mortality risks +/- 10%	33,433	33,547
Exponential rates of OS and PFS failure with treatments for advanced RCC and pNET +/- 10%	33,105	33,818
Drug acquisition and administration costs		
Unit costs of IV drug administration +/- 10%	33,493	33,488
Unit cost of oral drug dispensing +/- 10%	33,494	33,486

Surgery, surgical complication, disease management, and AE costs		
Medical management costs in pre-surgery state +/- 10%	33,436	33,544
Medical management costs in event-free after surgery state +/- 10%	33,532	33,449
Social care costs of CNS Hb progression in the VHL-CNS Hb cohort +/- 10%	33,498	33,483
Costs of surgical procedures +/- 10%	33,529	33,452
Costs of short-term surgical complications +/- 10%	33,565	33,416
Costs of long-term surgical complications +/- 10%	34,661	32,319
Medical management costs in pre-progression metastatic disease state +/- 10%	33,496	33,484
Medical management costs in post-progression metastatic disease state +/- 10%	33,502	33,478
Terminal care costs (one-time cost) +/- 10%	33,497	33,484
Costs of AEs +/- 10%	33,490	33,491
Utilities		
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	38,429	42,768
Utility in pre-progression metastatic disease state +/- 10%	33,593	33,387
Utility in post-progression metastatic disease state +/- 10%	33,717	33,262
Disutilities of short-term complications +/- 10%	33,477	33,503
Disutilities of long-term complications +/- 10%	46,917	34,620
Disutility from AEs +/- 10%	33,489	33,491
Scenario analyses		
Perspective		
Include indirect costs (societal perspective)	28,041	
Time horizon and annual discount rate		
Time horizon: 20 years	46,025	
Time horizon: 30 years	46,861	
Annual discount rate: 0.0%	31,097	
Annual discount rate: 1.5%	31,972	
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	30,644	
Efficacy and transition probability scenarios		
Distribution for pre-surgery→surgery in the belzutifan arm (VHL-RCC cohort): Gamma	33,490	
Do not adjust surgery and metastases rates to account for real-world standard of care	41,584	

Drug cost scenarios	
Distribution for belzutifan ToT: Weibull	42,057
Assume no treatment effect waning: Model efficacy and ToT separately	10,385
Do not apply relative dose intensity	37,272
Do not allow vial-sharing	33,479
Include costs of first-line advanced regimens only	33,717
Utility/disutility scenarios	
Assume same utility for CR as PR/SD	33,760
Apply caregiver disutility	31,468
Do not apply age-adjusted disutility	31,789
Do not apply AE-related disutility	33,483

*Note: The x1.7 severity weight is applied to the incremental QALYs in the ICER calculation.

Table 168 Tabular DSA and scenario analysis results for VHL-pNET cohort

	ICER vs. comparator: Belzutifan vs. SOC	
	(£/QALY)	
	Low input value	*High input value*
Base case	45,676	
High/low DSAs		
Efficacy and transition probabilities		
Proportion receiving immediate RCC surgery under SOC in VHL-RCC cohort: 80-100%	45,676	45,676
Proportion receiving immediate CNS Hb surgery or sequelae under SOC in VHL-CNS Hb cohort: 90-100%	45,676	45,676
Proportion receiving immediate pNET surgery under SOC in VHL-pNET cohort: 80-100%	49,328	42,293
Exponential rates of EF→MD under belzutifan +/- 10%	45,676	45,676
Exponential rates of EF→MD under SOC +/- 10%	45,721	45,630
Exponential rates of EF→Death under belzutifan +/- 10%	45,676	45,676

Exponential rates of EF→Death under SOC +/- 10%	45,587	45,762
Perioperative mortality risks +/- 10%	45,652	45,699
Exponential rates of OS and PFS failure with treatments for advanced RCC and pNET +/- 10%	44,740	46,487
Drug acquisition and administration costs		
Unit costs of IV drug administration +/- 10%	45,682	45,670
Unit cost of oral drug dispensing +/- 10%	45,682	45,669
Surgery, surgical complication, disease management, and AE costs		
Medical management costs in pre-surgery state +/- 10%	45,600	45,751
Medical management costs in event-free after surgery state +/- 10%	45,717	45,634
Social care costs of CNS Hb progression in the VHL-CNS Hb cohort +/- 10%	45,676	45,676
Costs of surgical procedures +/- 10%	46,044	45,307
Costs of short-term surgical complications +/- 10%	45,917	45,434
Costs of long-term surgical complications +/- 10%	46,961	44,390
Medical management costs in pre-progression metastatic disease state +/- 10%	45,686	45,666
Medical management costs in post-progression metastatic disease state +/- 10%	45,698	45,653
Terminal care costs (one-time cost) +/- 10%	45,687	45,664
Costs of AEs +/- 10%	45,675	45,676
Utilities		
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	49,433	60,169
Utility in pre-progression metastatic disease state +/- 10%	45,625	45,732
Utility in post-progression metastatic disease state +/- 10%	45,565	45,818
Disutilities of short-term complications +/- 10%	45,677	45,674
Disutilities of long-term complications +/- 10%	66,940	44,318
Disutility from AEs +/- 10%	45,674	45,677
Scenario analyses		
Perspective		
Include indirect costs (societal perspective)	35,367	
Time horizon and annual discount rate		
Time horizon: 20 years	65,994	
Time horizon: 30 years	64,661	

Annual discount rate: 0.0%	38,457
Annual discount rate: 1.5%	41,523
Annual discount rate: 3.5% for costs and 1.5% for effectiveness	39,930
Efficacy and transition probability scenarios	
Distribution for pre-surgery→surgery in the belzutifan arm (VHL-RCC cohort): Gamma	46,264
Do not adjust surgery and metastases rates to account for real-world standard of care	46,474
Drug cost scenarios	
Distribution for belzutifan ToT: Weibull	56,748
Assume no treatment effect waning: Model efficacy and ToT separately	13,560
Do not apply relative dose intensity	50,945
Do not allow vial-sharing	45,655
Include costs of first-line advanced regimens only	46,046
Utility/disutility scenarios	
Assume same utility for CR as PR/SD	47,698
Apply caregiver disutility	42,868
Do not apply age-adjusted disutility	44,119
Do not apply AE-related disutility	45,661

*Note: The x1.7 severity weight is applied to the incremental QALYs in the ICER calculation.

Appendix K: Price details of treatments included in the submission

K1.1 Price of intervention

Please see section *B.3.5 Cost and healthcare resource use identification, measurement and valuation* for details of intervention costs used in the model.

K1.2 Price of comparators and subsequent treatments

Please see section *B.3.5 Cost and healthcare resource use identification, measurement and valuation* for details of intervention costs used in the model.

Appendix L: Checklist of confidential information

The checklist of confidential information is presented in its own stand-alone "Appendix D – confidential information checklist" document provided as part of this submission.

Appendix M: MK-6482-004 study statistical analysis details

Study objectives

Primary objective

- To evaluate the efficacy of belzutifan for the treatment of von Hippel-Lindau (VHL) disease-associated renal cell carcinoma (RCC) as measured by overall response rate (ORR) per Response Evaluation Criteria in Solid Tumours Version 1.1 (RECIST 1.1)

Secondary objective

- To evaluate efficacy of belzutifan for the treatment of VHL disease-associated RCC measured as follows:
    - Duration of response (DOR)
    - Time to response (TTR)
    - Progression-free survival (PFS)
    - Time to Surgery (TTS)
- To evaluate efficacy of belzutifan for the treatment of VHL disease associated non-RCC tumours (retinal and CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)
- To evaluate safety and tolerability of belzutifan
- To assess the pharmacokinetics (PK) of belzutifan

Exploratory objective

- To evaluate changes in pharmacodynamic (PD) markers (e.g., serum erythropoietin)

Study hypotheses

No formal hypothesis testing will be performed for this study.

Sample size determination

This study will enrol approximately 50 patients. Even though no formal hypothesis testing will be performed for this study, the required sample size for this study is based on the following assumptions. The null hypothesis is that the ORR is 15% (P0 = 0.15). The alternative hypothesis is that the ORR is 30% (P1 = 0.3). A sample size of 50 patients will provide greater than 80% power to reject the null under the alternative hypothesis using a one-sided test at a 0.05 level of significance.

Analysis population

All Patients

All patients who have signed the written informed consent form. This population will be used for the summary of patient disposition and data listings.

Efficacy Analysis Set

The All Participants as Treated (APaT) population will be used for the analyses of efficacy. The APaT population consists of all allocated patients who received at least one dose of belzutifan.

Safety Analysis Set

The APaT population will be used for the analysis of safety data in this study. The APaT population consists of all allocated patients who received at least one dose of belzutifan.

Pharmacokinetic Analysis Set

Pharmacokinetic Analysis Set will include all patients who received at least 1 dose of belzutifan and have at least one post-dose pharmacokinetic sample collection.

Pharmacodynamic Analysis Set

Pharmacodynamic Analysis Set will include all patients who received at least one dose of study drug and have evaluable pharmacodynamics data above the limit of quantification.

Interim Analyses, Final Analyses, and Unblinding

Periodic review of the trial data will be performed. Any analysis for the study will only take place after all patients have had the opportunity to complete at least two imaging assessments on study or have discontinued study therapy by the time of analysis data cut-off. The final analyses for the study will utilise a data cut-off date which will be at least 36 weeks after enrolment of the last patient. This is a single-arm open-label study, no unblinding process will take place.

Pooling of sites

This is a global multicentre study which will have approximately 20 sites enrolling patients. Data from all centres will be pooled in the analysis. No adjustment of centre effect will be performed.

Statistical analysis considerations

Since this is a single-arm study, inferential statistical methodology will not be employed in any of the data analyses for treatment comparison.

In general, data will be summarised using descriptive statistics. Continuous data will be summarised with number of patients (n), arithmetic mean, standard deviation (SD), median, minimum, and maximum. Additionally, geometric mean and coefficient of variation will be included where applicable in pharmacokinetic analyses. Categorical data will be summarised using frequency counts and percentages. Time-to-event endpoints will be reported using Kaplan-Meier estimates, along with 95% confidence intervals for median time to event.

Efficacy analyses

Efficacy Analysis of VHL Associated RCC Lesions

The primary summaries of all response data will be based on the Efficacy Analysis Set using IRC assessments per organ-based modified RECIST approach. Summaries of efficacy data may also be provided based on investigator assessments.

Response of RCC tumours will be measured by contrast enhanced computerised tomography (CT) or magnetic resonance image (MRI) scan and assessed according to RECIST 1.1.

The maximum percent decrease in VHL disease associated RCC total tumour burden will be displayed using waterfall plot. Duration of treatment will be displayed using swimmer plot.

Change in sum of largest diameters of VHL-associated RCC lesions from baseline (screening) before and after study drug treatment will be displayed using spider plots. Linear growth rate (LGR) for target tumours before and after study drug treatment will be calculated for patients with at least 3 scans, including the screening scan. For each treatment phase (pre-treatment or post-treatment), linear regression model will be used to determine LGR by regressing tumour size on time as a continuous variable at tumour-level. To determine patient-level LGR for each treatment phase, linear regression model will be used to regress tumour size on time as continuous variable and individual tumour as categorical variable. LGR will be derived as the regression coefficient of time. Pre-treatment and post-treatment LGR would be reported for each tumour and each patient along with adjusted R2. Descriptive summary of LGR will be provided to compare pre-treatment and post-treatment LGR.

Best Overall Response

Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response is documented. Where confirmed responses are required, the date of the first assessment will be used in all analyses. In the case that the confirmation assessment is not available, the assessment at the next scheduled time point will be used for confirmation. If there are no assessments, the response will be taken to be unconfirmed and a non-response.

Disease control rate (DCR) is the proportion of patients who have a BOR of CR, PR or SD.

Best confirmed response and DCR for VHL disease associated RCC tumours will be summarised.

Overall Response Rate

The overall response rate (ORR) for VHL disease-associated RCC tumours will be calculated as the proportion of patients who achieve a best confirmed response of CR or PR determined by RECIST 1.1 using the IRC assessment. Patients who do not have a tumour assessment for any reason will be considered non-responders and included in the denominator when calculating the response rate.

Estimates of the RCC ORR along with the associated 90% and 95% exact binomial confidence intervals (Clopper Pearson method) will be provided. No hypothesis testing will be conducted for ORR.

Progression-Free Survival

Progression-free survival in VHL disease associated RCC tumours is defined as the time from administration of the first dose of PT2977 first evidence of documented disease progression determined by RECIST 1.1 using the IRC assessment, or death (due to any cause), whichever occurs first.

The PFS derivation rules follow the publication by the Food and Drug Administration (FDA), “Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (2018)”

Table 169 Censoring rules for PFS

Situation	Date of progression or censoring
PD or death documented after ≤ 1 missed disease assessment, and before new anticancer therapy, if any	Progressed at date of documented PD or death
PD or death documented immediately after ≥ 2 consecutive missed disease assessments or after new anticancer therapy, if any	Censored at last adequate disease assessment prior to the earlier date of ≥ 2 consecutive missed disease assessment and new anticancer therapy, if any
No PD and no death; and new anticancer therapy is not initiated	Censored at last adequate disease assessment
No PD and no death; new anticancer therapy is initiated	Censored at last disease assessment documented before new anticancer therapy

Progression-free survival will be analysed using the Kaplan-Meier estimator. PFS survival rates will be presented with two-sided 95% CIs. Median, first and third

quartiles of PFS will be reported along with 95% Brookmeyer-Crowley confidence intervals. The cumulative PFS will be plotted in a Kaplan-Meier step plot over time in Efficacy Analysis Set.

KM cumulative estimators along with the associated 95% CIs will be provided for PFS rates at specific time-points.

Duration of Response

Duration of response (DOR) in VHL disease associated RCC tumours will be calculated for patients who achieve a CR or PR determined by RECIST 1.1 using the IRC assessment. Duration of response is defined as the time from first documentation of response assessment of CR or PR until documentation of objective disease progression or death from any cause. Start time is the first instance of CR or PR that is subsequently confirmed, not the confirmatory date itself.

Table 170 Censoring rules for DOR

Situation	Date of progression or censoring	Outcome
No progression nor death, no new anti-cancer therapy initiated	Last adequate disease assessment	Censor (non-event)
No progression nor death, new anti-cancer therapy initiated	Last adequate disease assessment documented before new anti-cancer therapy initiated	Censor (non-event)
Death or progression immediately after ≥ 2 consecutive missed disease assessments or after new anti-cancer therapy, if any	Earlier date of last adequate disease assessment prior to ≥ 2 missed adequate disease assessments and new anti-cancer therapy, if any	Censor (non-event)
Death or progression after ≤ 1 missed disease assessments and before new anti-cancer therapy, if any	PD or death	End of response (Event)

A missed disease assessment includes any assessment that is not obtained or is considered inadequate for evaluation of response.

The distribution of DOR will be estimated and displayed graphically using the Kaplan-Meier method. Median, first, and third quartiles of DOR will be reported along with 95% Brookmeyer-Crowley confidence intervals. The cumulative DOR will be plotted in a Kaplan-Meier step plot over time in patients who achieve a CR or PR from Efficacy Analysis Set.

Time to Response

Time to response (TTR) in VHL disease associated RCC tumours will be calculated for patients who achieve a CR or PR determined by RECIST 1.1 using the IRC assessment. TTR is defined as the time (weeks) from the study treatment start to the first occurrence of a response of PR or better that leads to the confirmed response.

TTR will be descriptively summarised in Efficacy Analysis Set.

Time to surgery

Time to surgery (TTS) in VHL disease associated RCC is defined as the time (weeks) from the study treatment start to the date of surgery. Surgery is defined as any tumour reducing intervention including partial nephrectomy, radical nephrectomy, ablative procedure (cryoablation, thermal ablation, radioablation, etc), tumour debulking surgeries etc but excluding radiation therapy. Patients who do not undergo surgery will be censored as of the date of last known alive date.

TTS will be estimated and displayed graphically using the Kaplan-Meier method. Median, first, and third quartiles of TTS will be reported along with 95% Brookmeyer-Crowley confidence intervals. The cumulative TTS will be plotted in a Kaplan-Meier step plot over time in Efficacy Analysis Set.

Efficacy analysis of non-RCC lesions other than retinal lesions and epididyamal cystadenomas

The primary summaries of all response data will be based on the Efficacy Analysis Set. Summaries of efficacy data will be provided based on investigator assessments. For pancreatic and CNS hemangioblastomas, summary of efficacy data will be provided based on IRC assessment.

Tumour response will be measured by contrast enhanced computerized tomography (CT) or magnetic resonance image (MRI) scan and assessed according to RECIST 1.1.

Summaries of ORR, DOR and best confirmed response will be provided for non-RCC tumours. Summaries of PFS, DOR, and TTR will be provided for CNS

hemangioblastomas, and pancreas. The above analyses will be conducted by applying the same rules and methods as RCC lesions.

TTS for CNS hemangioblastomas, and pancreas will be estimated and displayed graphically using the Kaplan-Meier method. Median, first and third quartiles of TTS will be reported along with 95% Brookmeyer-Crowley confidence intervals. The cumulative TTS will be plotted in a Kaplan-Meier step plot over time in Efficacy Analysis Set.

Safety analyses

All safety analyses will be performed based on Safety Analysis Set.

Extent of exposure

Duration of belzutifan exposure in weeks (calculated as date of administration of last dose - date of administration of first dose of belzutifan + 1 divided by 7) will be summarised. Duration of belzutifan exposure in months will be summarised.

The total cumulative dose defined as the sum of the received doses across all study days will be summarised descriptively.

Number of subjects with dose decrease, and dose interruptions will be summarised.

Adverse events

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA). The severity of AEs will be coded according to NCI CTCAE 4.03. Treatment-related AEs are AEs that are not stated as unrelated to belzutifan.

AEs that occurs after first administration of study drug up to 28 days after administration of the last dose of study drug will be included in AE summary tables. Any AE occurring after 28 day follow-up period after discontinuation of study treatment will be included in listings.

Appendix N: Response Evaluation Criteria in Solid Tumours Version 1.1 (RECIST 1.1) definitions of best overall response

The following is a summary of the key criteria and definitions used for best overall response as part of RECIST 1.1 (38), that were used as part of the MK-6482-004 study. For the full details of RECIST 1.1, pleasec consult the Eisenhauer et al. 2009 publication (38).

Response criteria

Evaluation of target lesions

Complete Response (CR):

Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to <10 mm.

Partial Response (PR):

At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

Progressive Disease (PD):

At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

Stable Disease (SD):

Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

Evaluation of best overall response

The best overall response is the best response recorded from the start of the study treatment until the end of treatment taking into account any requirement for confirmation (in the case of the MK-6482-004 study specifically, best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response is documented). The patient's best overall response assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions. Depending on the nature of the study and the protocol requirements, it may also require confirmatory measurement (described later). Specifically, in non-randomised trials where response is the primary endpoint, confirmation of PR or CR is needed to deem either one the 'best overall response'. This is described further below.

Time point response

It is assumed that at each protocol specified time point, a response assessment occurs. Table 171 provides a summary of the overall response status calculation at each time point for patients who have measurable disease at baseline.

Table 171 Time point response: patients with target (+/– non-target) disease.

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and

When patients have non-measurable (therefore non-target) disease only, Table 172 is to be used (as the results of the MK-6482-004 study are reported in terms of target RCC, CNS hemangioblastoma, and pNET lesions that exist at baseline, these are less relevant in for the reporting of results in the MK-6482-004 study).

Table 172 Time point response: patients with non-target disease only.

Non-target lesions	New lesions	Overall response
CR	No	CR
Non-CR/non-PD	No	Non-CR/non-PD
Not all evaluated	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR = complete response, PD = progressive disease, and NE = inevaluable.
[a]'Non-CR/non-PD' is preferred over 'stable disease' for non-target disease since SD is increasingly used as endpoint for assessment of efficacy in some trials so to assign this category when no lesions can be measured is not advised.

Missing assessments and inevaluable designation

When no imaging/measurement is done at all at a particular time point, the patient is not evaluable (NE) at that time point. If only a subset of lesion measurements are made at an assessment, usually the case is also considered NE at that time point, unless a convincing argument can be made that the contribution of the individual missing lesion(s) would not change the assigned time point response. This would be most likely to happen in the case of PD. For example, if a patient had a baseline sum of 50 mm with three measured lesions and at follow-up only two lesions were assessed, but those gave a sum of 80 mm, the patient will have achieved PD status, regardless of the contribution of the missing lesion.

Best overall response: all time points

The best overall response is determined once all the data for the patient is known. Best response determination in trials where confirmation of complete or partial response IS required (as is the case in the MK-6482-004 study): Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (in the case of the MK-6482-004 study specifically, best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response is documented). In this circumstance, the best overall response can be interpreted as in Table 3.

Table 173 Best overall response when confirmation of CR and PR required.

Overall response	Overall response	BEST overall response
First time point	Subsequent time point	
CR	CR	CR

Overall response	Overall response	BEST overall response
First time point	**Subsequent time point**	
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.

[a]If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

Appendix O: MK-6482-004 study results from other data cut-off dates

Overall response rate

VHL RCC

01-JUN-2020

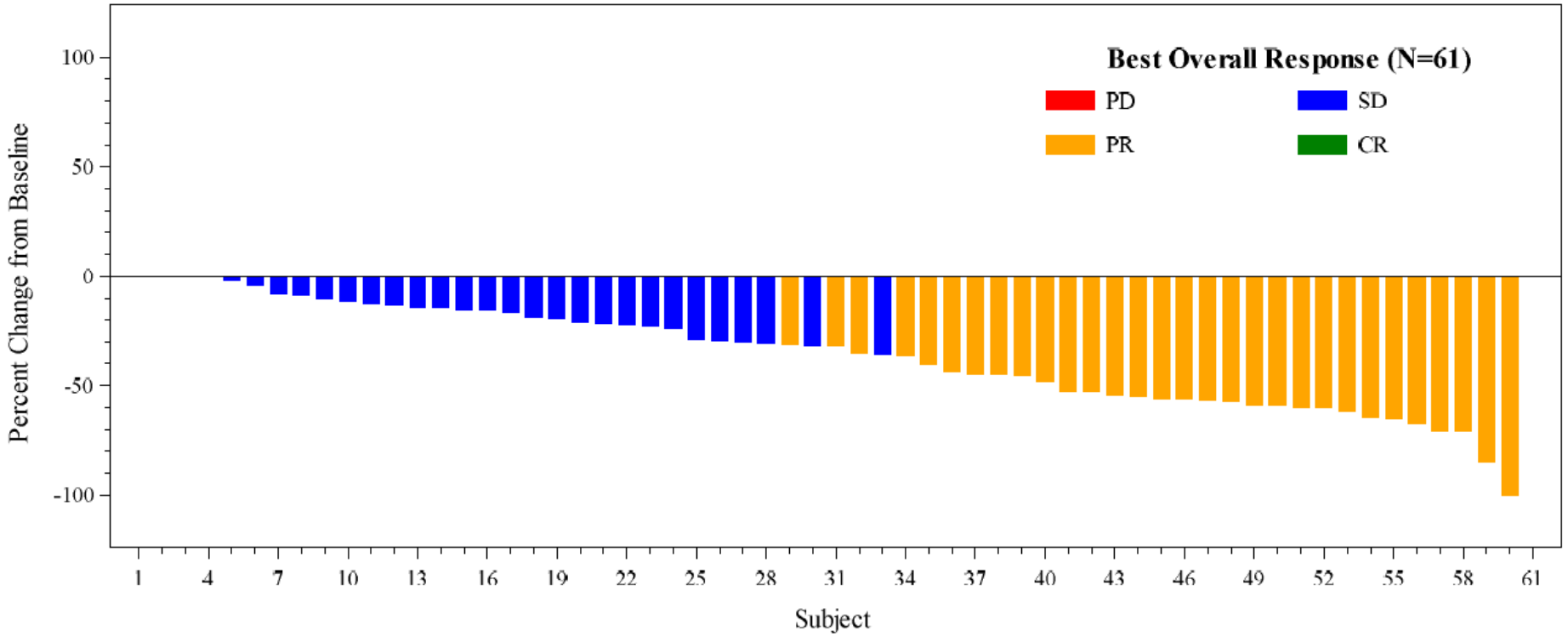
Figure 41 Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set): 01-JUN-2020

Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.

Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement.
Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8)

01-DEC-2020

Figure 42 Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set): 01-DEC-2020



Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement.
Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8)

15-JUL-2021

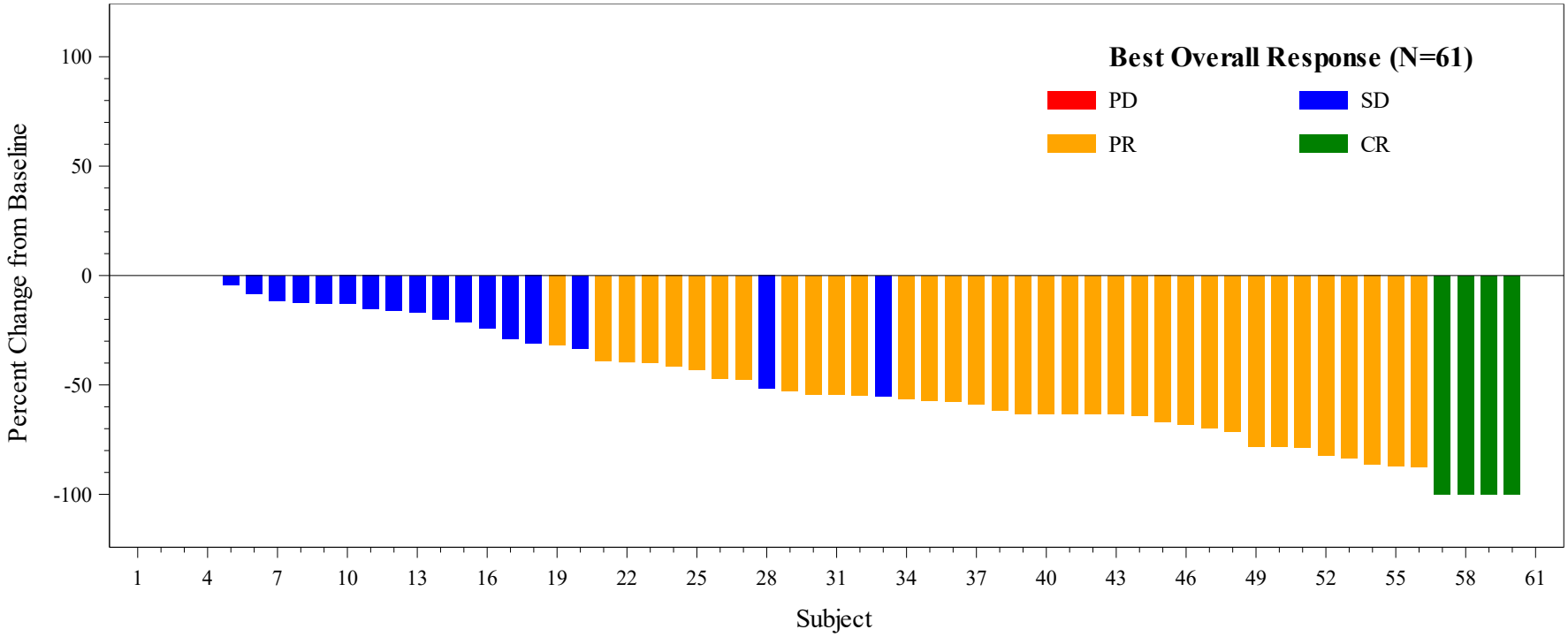
Figure 43 Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set); 15-JUL-2022

Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement.

Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8)
Date of Data Cut-off: 15JUL2022

01-APR-2022

Figure 44 Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set); 01-APR-2022



Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.

Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement.
Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8)
Date of Data Cut-off: 01APR2022

Appendix P: Additional MK-6482-004 study results

Additional baseline characteristics data

All patients

Table 174 MK-6482-004 study demographic and baseline characteristics (safety analysis set) - all patients - number of key target lesions

	120 mg QD	
	n	(%)
Participants in population	61	
Number of CNS Hemangioblastoma Target Lesions		
No tumour	31	(50.8)
1 tumour	12	(19.7)
2 tumours	10	(16.4)
>= 3 tumours	8	(13.1)
Number of CNS Hemangioblastoma Target Lesions (Solid Component Only)		
No tumour	36	(59.0)
1 tumour	15	(24.6)
2 tumours	8	(13.1)
>= 3 tumours	2	(3.3)
Number of Pancreatic Target Lesions		
No tumour	3	(4.9)
1 tumour	33	(54.1)
2 tumours	17	(27.9)
>= 3 tumours	8	(13.1)
Number of Pancreatic Neuroendocrine Target Lesions		
No tumour	39	(63.9)
1 tumour	18	(29.5)
2 tumours	3	(4.9)
>= 3 tumours	1	(1.6)
Number of RCC Target Lesions		
1 tumour	29	(47.5)
2 tumours	18	(29.5)
>= 3 tumours	14	(23.0)
Database Cutoff Date: 01APR2022		

Subgroup of patients with CNS hemangioblastoma

Table 175 MK-6482-004 study demographic and baseline characteristics (safety analysis set) – subgroup of patients with CNS hemangioblastoma

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	30		20		50	
Age (Years)						
Participants with data	30		20		50	
Mean	39.4		41.8		40.4	

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
SD	12.8		13.0		12.8	
Median	37.5		43.0		40.5	
Range	22.0 to 65.0		19.0 to 62.0		19.0 to 65.0	
Ethnicity						
Hispanic or Latino	3	(10.0)	2	(10.0)	5	(10.0)
Not Hispanic or Latino	26	(86.7)	18	(90.0)	44	(88.0)
Unknown	1	(3.3)	0	(0.0)	1	(2.0)
Race						
Asian	1	(3.3)	0	(0.0)	1	(2.0)
Black or African American	1	(3.3)	1	(5.0)	2	(4.0)
Native Hawaiian or Other Pacific Islander	1	(3.3)	0	(0.0)	1	(2.0)
White	26	(86.7)	18	(90.0)	44	(88.0)
Unknown	1	(3.3)	1	(5.0)	2	(4.0)
Weight (kg)						
Participants with data	30		20		50	
Mean	86.4		73.5		81.2	
SD	20.8		25.6		23.5	
Median	81.5		65.7		75.5	
Range	63.6 to 147.6		47.7 to 147.0		47.7 to 147.6	
Height (cm)						
Participants with data	30		18		48	
Mean	176.5		160.4		170.5	
SD	9.0		7.4		11.5	
Median	174.8		159.5		170.7	
Range	159.5 to 195.0		148.0 to 174.0		148.0 to 195.0	
BMI (kg/m2)						
Participants with data	30		18		48	
Mean	27.7		28.6		28.0	
SD	5.8		9.5		7.3	
Median	27.0		26.3		26.7	
Range	18.4 to 42.7		17.2 to 52.0		17.2 to 52.0	
ECOG Performance Status						
0	22	(73.3)	17	(85.0)	39	(78.0)
1	8	(26.7)	2	(10.0)	10	(20.0)
2	0	(0.0)	1	(5.0)	1	(2.0)

Database Cutoff Date: 01APR2022
Number of participants: Safety Population
Note: Baseline is defined as the last available measurement prior to the first dose administered.

Table 176 MK-6482-004 study demographic and baseline characteristics (safety analysis set) - subgroup of patients with CNS hemangioblastoma - number of key target lesions

	120 mg QD	
	n	(%)
Participants in population	50	
Number of CNS Hemangioblastoma Target Lesions		
No tumour*	20	(40.0)
1 tumour	12	(24.0)
2 tumours	10	(20.0)
>= 3 tumours	8	(16.0)
Number of Pancreatic Target Lesions		
No tumour	2	(4.0)
1 tumour	24	(48.0)
2 tumours	16	(32.0)
>= 3 tumours	8	(16.0)
Number of Pancreatic Neuroendocrine Target Lesions		
No tumour	33	(66.0)
1 tumour	13	(26.0)
2 tumours	3	(6.0)
>= 3 tumours	1	(2.0)
Number of RCC Target Lesions		
1 tumour	23	(46.0)
2 tumours	15	(30.0)
>= 3 tumours	12	(24.0)
Database Cutoff Date: 01APR2022		

For CNS hemangioblastomas, 20 of these tumours did not meet the criteria for designation at baseline due their measurable tumour plus associated cyst size <1 cm.

Subgroup of patients with pancreatic neuroendocrine tumour

Table 177 MK-6482-004 study demographic and baseline characteristics (safety analysis set) – subgroup of patients with pNET

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	11		11		22	
Age (Years)						
Participants with data	11		11		22	
Mean	42.4		42.5		42.4	
SD	13.6		16.1		14.6	
Median	41.0		42.0		42.0	
Range	24.0 to 64.0		19.0 to 66.0		19.0 to 66.0	
Ethnicity						
Hispanic or Latino	1	(9.1)	0	(0.0)	1	(4.5)
Not Hispanic or Latino	10	(90.9)	11	(100.0)	21	(95.5)
Race						
Asian	1	(9.1)	0	(0.0)	1	(4.5)

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Black or African American	1	(9.1)	1	(9.1)	2	(9.1)
White	9	(81.8)	9	(81.8)	18	(81.8)
Unknown	0	(0.0)	1	(9.1)	1	(4.5)
Weight (kg)						
Participants with data	11		11		22	
Mean	84.5		81.5		83.0	
SD	23.0		34.7		28.8	
Median	75.7		62.9		75.5	
Range	63.0 to 147.6		48.4 to 147.0		48.4 to 147.6	
Height (cm)						
Participants with data	11		11		22	
Mean	174.4		162.0		168.2	
SD	8.5		6.8		9.9	
Median	174.6		161.7		167.5	
Range	159.5 to 185.9		149.9 to 174.0		149.9 to 185.9	
BMI (kg/m2)						
Participants with data	11		11		22	
Mean	27.5		31.1		29.3	
SD	5.5		12.5		9.6	
Median	26.3		26.2		26.3	
Range	19.9 to 42.7		17.2 to 52.0		17.2 to 52.0	
ECOG Performance Status						
0	10	(90.9)	10	(90.9)	20	(90.9)
1	1	(9.1)	1	(9.1)	2	(9.1)
Database Cutoff Date: 01APR2022 Number of participants: Safety Population Note: Baseline is defined as the last available measurement prior to the first dose administered.						

Table 178 MK-6482-004 study demographic and baseline characteristics (safety analysis set) – subgroup of patients with pNET - number of key target lesions

	120 mg QD	
	n	(%)
Participants in population	22	
Number of CNS Hemangioblastoma Target Lesions		
No tumour	13	(59.1)
1 tumour	4	(18.2)
2 tumours	3	(13.6)
>= 3 tumours	2	(9.1)
Number of CNS Hemangioblastoma Target Lesions (Solid Component Only)		
No tumour	12	(54.5)
1 tumour	7	(31.8)
2 tumours	2	(9.1)
>= 3 tumours	1	(4.5)

	120 mg QD	
	n	(%)
Number of Pancreatic Target Lesions		
1 tumour	14	(63.6)
2 tumours	7	(31.8)
>= 3 tumours	1	(4.5)
Number of Pancreatic Neuroendocrine Target Lesions		
1 tumour	18	(81.8)
2 tumours	3	(13.6)
>= 3 tumours	1	(4.5)
Number of RCC Target Lesions		
1 tumour	13	(59.1)
2 tumours	3	(13.6)
>= 3 tumours	6	(27.3)
Database Cutoff Date: 01APR2022		

Locations of CNS hemangioblastoma tumours in patients with each best overall response type

Complete response

Table 179 Locations of the CNS hemangioblastomas, in participants with CNS hemangioblastoma (RECIST 1.1) by IRC at baseline with complete response, efficacy analysis set

Study: MK6482-004	**120 mg QD**
Characteristic	**N[a]=4**
Locations of the CNS hemangioblastomas	
SPINE	[redacted]
Missing	
Database Cutoff Date: 01APR2022	
a: Number of participants: Efficacy Analysis Population	

Partial response

Table 180 Locations of the CNS hemangioblastomas, in participants with CNS hemangioblastoma (RECIST 1.1) by IRC at baseline with partial response, efficacy analysis set

Study: MK6482-004	**120 mg QD**
Characteristic	**N[a]=18**
Locations of the CNS hemangioblastomas	
BRAIN STEM; CEREBELLUM; SPINE CEREBELLUM CEREBELLUM; FRONTAL LOBE CEREBELLUM; SPINE FRONTAL LOBE	[redacted]

SPINAL CORD SPINE OTHER	[redacted]
Database Cutoff Date: 01APR2022 a: Number of participants: Efficacy Analysis Population	

Stable disease

Table 181 Locations of the CNS hemangioblastomas, in participants with CNS hemangioblastoma (RECIST 1.1) by IRC at baseline with stable disease, efficacy analysis set

Study: MK6482-004	**120 mg QD**
Characteristic	**N[a]=23**
Locations of the CNS hemangioblastomas	
CEREBELLUM CEREBELLUM; SPINE LEPTOMENINGEAL LEPTOMENINGEAL; SPINAL CORD LEPTOMENINGEAL; TEMPORAL LOBE Missing	[redacted]
Database Cutoff Date: 01APR2022 a: Number of participants: Efficacy Analysis Population	

Progressive disease

Table 182 Locations of the CNS hemangioblastomas, in participants with CNS hemangioblastoma (RECIST 1.1) by IRC at baseline with progressive disease, efficacy analysis set

Study: MK6482-004	**120 mg QD**
Characteristic	**N[a]=3**
Locations of the CNS hemangioblastomas	
CEREBELLUM CEREBELLUM; SPINE Missing	[redacted]
Database Cutoff Date: 01APR2022 a: Number of participants: Efficacy Analysis Population	

Not evaluable

Table 183 Locations of the CNS hemangioblastomas, in participants with CNS hemangioblastoma (RECIST 1.1) by IRC at baseline with response not evaluable, efficacy analysis set

Study: MK6482-004	**120 mg QD**
Characteristic	**N[a]=2**
Locations of the CNS hemangioblastomas	
SPINAL CORD	[redacted]

Missing	[redacted]

Database Cutoff Date: 01APR2022
a: Number of participants: Efficacy Analysis Population

Appendix Q: VHL Natural History Study summary of key details

Rationale

A comprehensive understanding of the natural history and progression of renal solid tumours associated with VHL syndrome is critical for informing treatment strategies and for interpreting results from clinical trials. Currently, MSD is conducting the single-arm MK-6482-004 trial to investigate belzutifan therapy for treatment of VHL disease-associated RCC. To help demonstrate the clinical benefit of this investigational therapy, and to inform future product developments, additional work was needed to describe growth kinetics and assessment of change in tumour burden (using Response Evaluation Criteria in Solid Tumours [RECIST] V1.1 criteria), patterns of care (including surgical patterns), and renal function over time among this patient population. The assessment of growth kinetics and change in tumour burden can support the claim that tumour growth is likely to occur in the absence of pharmacologic intervention. The current study aims to use a large data source of patients with VHL syndrome managed and treated at the National Cancer Institute (NCI) to fill in gaps in our understanding on the natural history of the disease.

Objectives

Primary objective

The primary objective was to evaluate the linear growth rate (LGR) of renal solid tumours among patients with VHL syndrome who have at least one measurable renal solid tumour (defined using RECIST 1.1) and at least three measurements of renal tumour diameter (one initial measurement and at least two subsequent measurements) for unique tumour(s) during the assessment window.

Secondary objectives

Separately, among patients with VHL syndrome who have at least three measurements (one initial measurement and at least two subsequent measurements) of diameter of any target or measurable non-target renal solid tumors during the assessment window, the secondary objectives were:

- To report the real-world tumor response and disease progression metrics, as determined by RECIST 1.1 criteria:
    - Real-world objective response rate (rwORR), defined as the proportion of patients with a best confirmed response of complete response (CR) or partial response (PR);
    - The number of patients by the real-world best overall response (rwBOR);
    - Time from Patient-Level Index Date until first documentation of progressive disease (PD);
    - Time from Patient-Level Index Date until the first documentation of confirmed response, defined using best response of CR or PR;
    - Real-world progression-free survival (rwPFS);

Among patients with VHL syndrome who have at least one renal solid tumor measurement during the study period:

- To describe surgical and other renal procedure patterns, specifically the:
    - Proportion of patients with renal surgeries with and without complications (including estimated blood loss and perioperative mortality);
    - Frequency and type of renal procedures;a
    - Time from Patient-Level Index Date in the study period to first surgery with therapeutic intent;
    - Time between surgeries with therapeutic intent;
- To describe VHL disease-associated metastasis patterns, specifically the:
    - Proportion of patients who develop metastasis;
    - Time from Patient-Level Index Date to metastasis detection;
    - Time from metastasis detection to date of mortality;

- To describe mortality patterns, specifically the:
    - Time from Patient-Level Index Date to mortality (overall and cause-specific);
    - Proportion of patients with mortality as compared to age- and sex-standardized mortality rates in the United States (US) population;
- To describe baseline demographic and clinicopathologic characteristics (e.g., detection of other organ-specific tumor types [CNS, pancreatic, etc.]);
- To describe laboratory values at key time points (including index date, pre- and post- each renal procedure, and at the end of follow-up), and the occurrence of chronic kidney disease (CKD), anemia, and patients qualifying for dialysis;
- To describe the prevalence and incidence of hypertension, as well as the frequency of risk factors for hypertension; and
- To describe the association between renal solid tumor LGR and other demographic and clinicopathologic characteristics.

Study design

Overview

The study was a retrospective non-interventional study of existing medical records, with supplemental electronic medical record (EMR) data abstraction and central imaging review of abdominal imaging scans obtained during routine clinical care. Patients with VHL syndrome who had ≥1 measurable renal solid tumour measured during the study period and met other study inclusion and exclusion criteria (described later subsections), were identified and followed until the end of the assessment window. The primary objective was to describe the LGRs of renal solid tumours among patients with VHL. Additional information characterising the natural history of VHL syndrome in the study population were to be presented as secondary objectives, including real-world tumour response and progression assessments using RECIST 1.1 criteria, and assessments of surgical and other renal procedures, metastasis, mortality, laboratory values including renal function, hypertension, and

other clinicopathologic characteristics. For some outcomes, supplemental EMR abstraction of data were necessary.

Data source

The study used data collected by the NCI's Urologic Oncology Branch (UOB) of patients with VHL syndrome managed and treated at the National Institutes of Health (NIH) Clinical Center in Bethesda, Maryland. Data collected on VHL patient characteristics, treatment, and follow-up information was available in medical records for all patients treated at the NCI's UOB since approximately 1987. For the current study, the data source included data registered by the NCI using a hybrid of existing structured data fields in the UOB Hereditary Database, linked to other structured NCI and external database (e.g., NCI laboratory and prescription database, and US National Death Index [NDI]), supplemented with additional medical record abstraction of unstructured data fields and results of central imaging review of Digital Imaging and Communications in Medicine (DICOM) images for patients receiving care at the NIH Clinical Center in Maryland. The current EMR system (Clinical Record Interactive System [CRIS]) was introduced by the NCI on July 31, 2004.

The study leveraged centralised review of comprehensive DICOM image data to widen the capture of the number of solid renal tumours and serial measurements included in the growth rate analysis and the assessment of change in tumour burden using RECIST 1.1. The central imaging vendor for the study, Bioclinica, was the same as that used for the MK-6482-004 clinical trial, and the central imaging review for the natural history study was performed using a similar independent review (IR) charter. Dates of metastasis detection, start and stop dates of systemic oncologic therapies, and complication information for all renal surgeries were abstracted from the unstructured EMR data. Prescription, laboratory, and vital statistics data were obtained through linkage with a separate structured NCI EMR database. Finally, mortality data (including date of death and cause of death) were obtained through linkage with the NDI.

Patient-level study populations

Two primary patient-level study populations and subgroups were identified:

- The Primary Study Population; and
- The Trial Population Subgroup.

Patient-Level Index Date

The earliest date that a measurable renal solid tumour was detected during the study period (Section 6.5 [Study Period]) was classified as the Patient-level index date.

NOTE: All dates in the dataset provided by the NCI have been defined relative to Patient-level index date. Thus Patient-level index date for all patients is defined as a relative date variable equal to 0.

Primary Study Population (Patient-Level)

The Primary Study Population consisted of patients treated at the NCI with confirmed VHL syndrome, residents of the US or Canada, ≥1 renal solid tumour with available measurement(s), available follow-up, and no investigational therapy, oncologic therapy, or renal procedures proximate (within 30 days) to Patient-level index date.

<u>Inclusion Criteria</u>

The following are the study inclusion criteria for the Primary Study Population (Patient-Level):

- Patients with VHL syndrome who are residents of the US or Canada; and

  NOTE: These criteria have already been applied in the dataset. All included patients (N=776) have VHL syndrome and are residents of either the US or Canada.

- Patients with ≥1 renal solid tumor identified and measured during the study period; and
- Patients with measurable disease, as defined by IR, at first renal solid tumour detection date.

Exclusion Criteria

The following are the study exclusion criteria for the Primary Study Population (Patient-Level):

- Patients who received systemic oncologic or investigational therapy within 30 days on or prior to Patient-level index date;
- Patients with any renal procedure in the 30 days on or prior to Patient-level index date; and
- Patients whose last follow-up date was on or prior to Patient-level index date.

Trial Population Subgroup (Patient-Level)

To benchmark the results of the MK-6482-004 study, the study conducted subgroup analyses among a sub-population of patients meeting inclusion and exclusion criteria that are closely matched to that of the trial.

Inclusion Criteria

The following are the study inclusion criteria for the Trial Population Subgroup (Patient-Level):

- Patients in the Primary Study Population; and
- Patients with a diagnosis of VHL based on germline VHL alteration.

Exclusion Criteria

The following are the study exclusion criteria for the Trial Population Subgroup (Patient-Level):

- If the largest tumor at Patient-level index date is ≥30 millimeters (mm; Renal solid tumor measurement at patient-level index date ≥30), patients with a renal surgical procedure performed within 60 days on or after Patient-level index date
- Patients who received treatment with MK-6482 or another hypoxia inducible factor 2 alpha (HIF-2α) inhibitor any time prior to Patient-level index date

- Patients who received systemic oncologic or investigational therapy any time prior to Patient-level index date
- Patients with evidence of VHL disease-associated metastatic disease prior to Patient-level index date

Study period

The study period began on July 31, 2004 (the date in which the CRIS EMR was implemented) and ended on June 30, 2020.

Patient-Level Study Period

Patient-Level Study Period: All Follow-up Time

Patients were followed from Patient-level index date until the first of:

- Mortality date; or
- Last clinical encounter date.

Patient-Level Study Period: Maximum 2 Years

In the 2-year follow-up window, patients were followed from Patient-level index date until the first of either Last patient-level follow-up date (Overall) or 2 years (730 days; see derived variable: Last patient-level follow-up date (2 years)).

Patient-Level Study Period: Maximum 5 Years

In the 5-year follow-up window, patients were followed from Patient-level index date until the first of Last patient-level follow-up date (Overall) or 5 years (1,825 days; see derived variable: Last patient-level follow-up date (5 years)).

RECIST-based Assessments Study Period

For the RECIST-based analyses, patients were followed from Patient-level index date until the first of the following events:

- Mortality date or Last clinical encounter date;

- Date of receipt of an investigational therapy (first Investigational therapy initiation date on or after Patient-level index date) or date of receipt of an oncologic therapy (first Oncologic medication initiation date on or after Patient-level index date); or
- Date of a renal procedure impacting any tumour (Earliest First renal procedure (tumour-level)).

Statistical analysis

Overview

Results were provided as descriptive statistics. The study population characteristics and frequency of the outcomes of interest were described using frequency and percentage distributions (cross-tabulations) for categorical variables and descriptive statistics (mean, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum) for continuous and count variables. Continuous and count variables were categorised into ranges and described using frequency and percentage distributions. Time to event data were summarised by Kaplan Meier methodology. The 25th percentile, median, and 75th percentile of survival time were provided with 95% confidence intervals (CIs). Kaplan-Meier curves and survival probabilities at key time points (1, 2, 3, 5, and 7 years) were presented.

All analyses used SAS (SAS Institute Inc., Cary, North Carolina). Results were summarised in tables and figures in Microsoft® Excel format.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Belzutifan for treating tumours associated with von Hippel-Lindau disease [ID3932]

Summary of Information for Patients (SIP)

March 2023

File name	Version	Contains confidential information	Date
NICE ID3932 Summary of information for patients	1.0	No	13th March 2023

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Belzutifan (WELIREG®)

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

The patient population being appraised by NICE are for adult patients (18 or older) that have certain types of tumours that have been caused by von Hippel-Lindau (VHL) disease. Patients must have at least one of the following types of tumours to be eligible:

- A type of tumour located in the kidney, referred to as a "renal cell carcinoma" or RCC
- A type of tumour located in the brain or spinal cord, referred to as a "central nervous system hemangioblastoma" or CNShb
- A type of tumour located in the pancreas, referred to as a "pancreatic neuroendocrine tumour" or pNET

The exact wording of the patient population being appraised by NICE is below:

Adults with untreated renal cell carcinoma, central nervous system hemangioblastomas, or pancreatic neuroendocrine tumours caused by von Hippel-Lindau disease

It should also be noted that the marketing authorisation granted for belzutifan, given in section 1c, is more specific than the above NICE wording. It specifies that patients have any one of these tumours, and that a doctor has decided that localised procedures (like surgery) are unsuitable or undesirable. This means if a patient is suitable to have a localised procedure and the outcome is likely to be desirable, they should have one, and if not belzutifan could be an option.

MSD were disappointed and surprised that the NICE Topic Selection Oversight Panel (TSOP) made the decision not to route this indication into the Highly Specialised Technologies (HST)

programme. We disagree with the decision. However, in order to facilitate access for patients with VHL we are moving ahead the STA process.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Belzutifan has a GB marketing authorisation for the indication in this submission that was first granted on 31-MAY-2022 with Medicines & Healthcare products Regulatory Agency (MHRA) marketing authorisation number PL 53095/0087 (1).

The indication relevant to this appraisal is provided below:

WELIREG (belzutifan) is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable (1).

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

MSD have had no financial transactions with VHL patient groups. MSD has been in regular contact with VHL UK & Ireland through one-to-one meetings over the past year to seek their support in understanding VHL disease, and at their request, to keep them up to date with the belzutifan submission.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

Below we outline what VHL disease is, what the impact is on patient's quality of life and what the mortality impact is.

What is VHL disease?

VHL is a rare disorder caused by a faulty gene. It is named after the two doctors who first described the disease and affects about one in 70,000 people in the UK (2). Tumours develop in one or more parts of the body. Many of these tumours involve the abnormal growth of blood vessels in parts of the body which are particularly rich in blood vessels. The areas most frequently affected are the eyes, the back of the brain (cerebellum), the spinal cord, the kidneys, the adrenal glands and the pancreas.

Different VHL features tend to develop at different ages. The eye features (angiomas) often develop in childhood. Others, including tumours found in the cerebellum, spinal cord or adrenal glands (hemangioblastomas and phaeochromocytomas), can develop from late childhood onwards. The kidney tumours usually develop the latest, from the mid-twenties onwards.

The features of VHL can be divided into two groups. The first are those that if left untreated, would go on to cause some kind of problem. The second group of features hardly ever cause symptoms and can just be regarded as helpful in diagnosing VHL.

Most often the tumours seen in VHL are classified as benign because they are not cancerous and do not have the potential to spread to other parts of the body. However, these tumours can have serious effects on the organs where they develop (3). Some of the disabilities possible from the progression of this disease are lifetime dialysis (visiting the hospital several times a week to have your blood cleaned), blindness, motor impairment (loss of function of body parts), and cancer.

Impact on mortality

VHL disease also has a significant impact on a patient's life expectancy. Results from a study using data from the North West Regional Genetic Register Service and the North West Cancer Intelligence Service indicate that VHL disease considerably shortens life, with the average life expectancy reduced by nearly 19 years in males and by nearly 34 years in females.

Impact of VHL disease on life expectancy

	Median life expectancy in years	
	General population	**People with VHL disease**
Males	78	59.4
Females	82	48.4
Overall	80	52.5

Source: Wilding et al. 2012 (4)

What is the impact of VHL to quality of life?

Every patient's journey with VHL is different. At one end of the scale, VHL may have a limited impact on a patient's life where some benign tumours are found but no treatment is required. However, many patients have a disease experience at a more severe end of the scale, whose lives revolve around scans, invasive surgeries, long periods of recovery requiring significant caregiver input and living with the day-to-day impacts reduced function due to the tumours or surgical procedures (for example reduced vision from tumours or surgery of the eye or reduced mobility and balance from tumours or surgery of the brain).

The best way to understand the impact of the disease is to hear their stories. They are captured by VHL UK & Ireland and can be found here (5). We share one illustrative example from Barry Maloney who describes the psychological, physical and emotional pain of the disease and the burden it has placed on him and his family.

With me, the cancer was close to the major blood vessels in the kidney. No one can say that you're going to be ok and be completely sure. It's impossible. I'm not a child anymore. I've seen enough in my life to know things don't always go to plan. Whether it's a post operative infection or a post operative bleed. There's no way of knowing what's going to happen in the future. No matter how much you pray, what's going to happen is going to happen.

I spent hours every day for a week after my surgery praying to God to stop the pain. I cried. I'm not ashamed to say it. I'm a teenage boy and I cried in front of my Aunt and Uncle, in front of nurses, and even in front of one of the most beautiful doctors I have ever seen in my life. Men need to know there's nothing wrong with crying. People don't think any less of you for doing it. So many men I know hold back tears no matter how much pain they're in. I spent hours praying, but there was nothing. I even pleaded to God to "take me away". Being in so much pain that you are literally praying for death is not a place I ever want to be in again. I have so much in my life. I have so many people in my life that I love, who love me. However, at one point that all went out the window. There's something about pain. It's a strange thing. It can make you want to throw your whole life away just for relief. At that moment I was in so much pain and I just wanted it all to end. It was a pain I've never experienced, a pain I never even imagined existed. It was what I imagine a gunshot wound or a stab wound feels like. If it is they really don't convey it very well in the movies. I screamed and shouted, all I remember is seeing nurses and doctors running towards me. Everything was hazy. I couldn't hear anything, all I could focus on was the pain.

…

That was the worst pain I had ever felt. That's what I thought, until December 4th 2015. That was the day my incredible Mam, the woman who raised me and my two big brothers by herself while fighting this awful disease, passed away. She fought everyday for 11 years. She went through three major and extremely life-changing surgeries over the 11 years. She spent the last 4 years of her life on dialysis for 4 hours a day, 3 days a week. She was the strongest woman I have ever known. She lost her husband to cancer and still managed to successfully raise three lads all by herself. VHL changed her life completely. That day in the hospital was like a paper cut compared to this pain. I lost a part of me the day she died.

Months before her passing Marie vowed that she would climb the Queen Maeve Trail up Knocknarea Mountain, Strandhill. That dream of hers never came true as she became too sick.

…

VHL has taken so much from me, my childhood, my Mam and even part of my body. My life has changed dramatically since that winter morning when I was 11 years old. The story doesn't end there. I know that eventually VHL could take my life, but it hasn't taken it yet.

- Barry Moloney - VHL Patient - Sep 8, 2019 (5)

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?
VHL disease is a rare, inherited disorder caused by genetic mutations in the VHL gene. The children of an individual with VHL have a 50% (one-in-two) chance of inheriting the genetic alteration. Occasionally patient with VHL is the first to have the gene mutation in the family due to a new ('de novo') mutation. The VHL gene helps to stop tumour growth and is important in a variety of cell growth processes, including the development of new blood vessels. In most cases, it is possible to identify the precise gene mutation causing VHL, which enables doctors to test for VHL disease before any symptoms occur. Testing for VHL can help doctors closely monitor (or 'screen') patients who have the disease and help identify family members at risk.

Diagnosis is usually made through identification of a disease-causing genetic mutation in the VHL gene. A genetic mutation can be identified in around 90% of those with VHL, however, some families can have a diagnosis of VHL based on the patterns of tumours that occur. Genetic testing is recommended in families with two or more lesions suggestive of VHL (retinal angioma, hemangioblastomas, multiple renal or pancreatic cysts, renal cell carcinoma, phaeochromocytoma, and endolymphatic sac tumours).

Genetic testing of the VHL gene should be offered to people with suspected VHL to identify the specific genetic mutation in the family. Testing for the specific gene mutation can then be offered to at-risk family members to identify those who have VHL and require screening. Testing in childhood is recommended due to the early development of tumours (6).

UK doctors who have expertise in VHL have suggested that patients at-risk or with confirmed VHL undergo the following monitoring (or 'screening') for their disease (7):

- Annual eye exams, starting from early childhood
- Brain scans every 12-36 months starting in adolescence
- Scans of the abdomen every year, starting from age 16
- Annual blood pressure and urine monitoring, starting from age 8 if patient at high risk of phaeochromocytoma (tumour of the adrenal gland)

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

VHL patients are managed at hospitals that have a speciality in this disease. A 'multidisciplinary team' (MDT) will look after the patients, involving doctors and other staff who have expertise in treating all types of tumour that can result from the disease. Through the 'screening' guidance given in 2b, this team will look out for tumours.

There are no drug treatments indicated for VHL disease, and currently the main way to treat the disease is by removing the tumours via surgery. This can be an effective way of stopping cancer. However, there are always risks to performing surgery. It is painful and can require patients to take several weeks or months off school or work to recover (8). There are also risks of things going wrong during surgery, from issues that can pass over time such as infection to a small risk of dying (9).

Because of the risks of surgery often hospital staff may decide a patient is better off leaving the tumour untreated. Tumours may not cause much of an issue to a patient until it's reached a certain size. If tumours grow, they can cause issues with how the body functions. For example a tumour in the brain can cause sickness, behavioural changes and memory problems (10). Large

tumours may also lead to greater risks of "metastatic" disease, or cancer that spreads beyond one organ (11). The MDT will carefully weigh these risks against the benefits of surgery before making a recommendation to a patient.

In some instances, an MDT may decide that a 'radical' surgery is required, meaning that a surgeon may aim to remove all of an organ. This is only done as a 'last resort', to aim to prevent cancer or other consequences of leaving the tumour. Removing a whole organ has permanent consequences for a patient. Removing both kidneys means a patient will have to visit the hospital several times a week for the rest of their life (or until a kidney transplant) to have their blood cleaned through dialysis (12). Removing the pancreas can lead to lifelong pancreatogenic diabetes (13).

2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

There are two recent patient-based studies into the VHL disease that were conducted that help to understand the patient needs and disease experience. The first, named "The VHL patient survey", was a study of over 200 patients run in collaboration with the VHL Alliance in the US, Canada, the UK, France, and Germany (14). The second is a UK based survey of 39 patients and carers run by VHL UK & Ireland, named the "VHL UK/Ireland Patient/Carer Survey" (15). Below we summarise the key results from each study.

<u>The VHL patient survey</u>

To understand the patient experience of VHL disease a large study was undertaken in collaboration with the VHL Alliance, called "The VHL patient survey". 200 patients were asked to complete a survey, designed to be completed in one sitting. Participation in the survey was voluntary, and data was collected between December 2021 and June 2022.

The objective of this survey was to understand the patient experience and burden of disease of patients with certain types of kidney tumours, brain or spine tumours, or pancreas tumours. The survey specifically aimed:

- To assess the patients' health-related quality of life
- To assess the patients' work productivity loss and activity impairment

The data collected in this survey included the patient experiences of undergoing surgery. The survey also included patient-reported outcome tools to help describe the burden (direct and indirect) of the condition. These tools included the quality-of-life assessment "EQ-5D" and the work impairment score "WPAI".

The EQ-5D has five questions on mobility, self-care, pain, usual activities, and psychological status with three possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). Results from these questions can then be combined and scaled to produce a single

score with a maximum score of 1. Scores can vary from 0, which represents death, to 1 which represents the best possible health state.

Overall, patients with VHL-related tumours had a mean EQ-5D score of 0.699. Patients with metastatic disease (n=16) had a mean EQ-5D score of 0.550 and patients without metastatic disease (n=195) had a mean EQ-5D score of 0.714.

The WPAI is a patient-reported numerical score of the amount of absenteeism, presenteeism and daily activity impairment attributable to a specific health problem. Respondents are asked questions concerning impairment due to VHL. The outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, i.e., worse outcomes. Someone completely unable to work or unemployed will score 100%.

Patients with VHL-related tumours had a mean percent overall work impairment of 30.8%. Patients with metastatic disease had a mean percent overall work impairment of 40.6% and patients without metastatic disease had a mean percent overall work impairment of 29.3%.

From these results we can see a clear impact of VHL disease on both a patient's quality of life and their ability to work. There is an unmet need for a treatment that can help address both issues.

<u>VHL UK/Ireland Patient/Carer Survey</u>

This survey was commissioned by the Trustees of VHL UK/Ireland Charity to aid the appraisal of belzutifan for VHL, focusing specifically on RCC (as it was only later understood CNShb and pNET tumours would also be part of the NICE appraisal). These results are based on 39 responses across patients and their carers. 74% of respondents captured the patients view, 18% of respondents captured the carers view and 8% provided both perspectives. Below is a summary of the findings.

- VHL often affects many aspects of a patient's life, including their careers, travel, relationships, family planning, finances, and social activities and in most of these, more than a 10% feel these have been affected permanently.
- Carers lives can also be greatly impacted by VHL. The data suggested carers often worry even more than patients about scans and follow appointments.
- Kidney cancer affected 51% of respondents to this survey, with at least 60% of those patients having experienced at least 1 surgery but some up to 6. Recovery from these surgeries is long and can include a high occurrence of mobility issues, pain, fatigue and anxiety/depression.
- The impact of VHL on patients and carers is significant, and lifelong.

Below is the percentage of patients who say symptoms of VHL tumours has affected decisions on the following areas of their lives at some point (either rarely, occasionally, frequently or permanently):

Table of patients who say symptoms of VHL tumours has affected decisions on the following areas of their lives

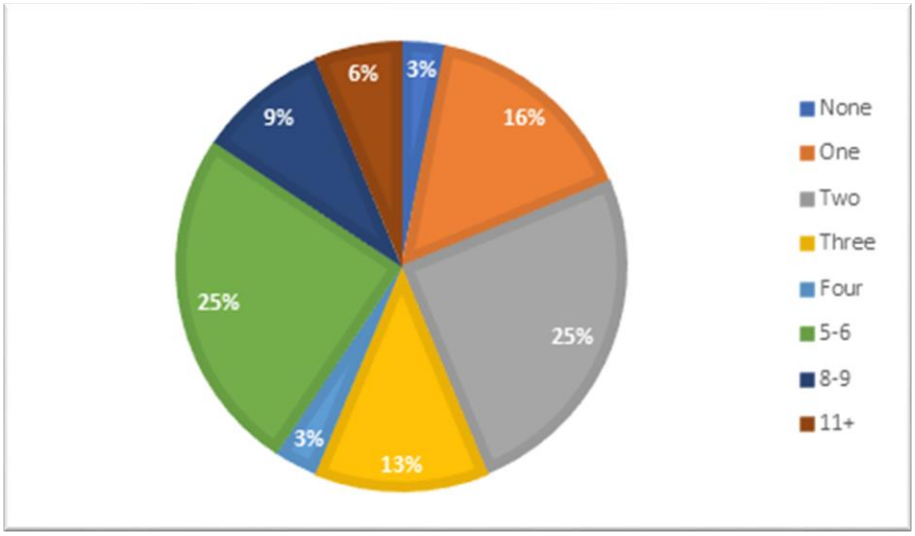
	Ever	Frequently	Permanently
Mild Mobility Issues	56%	13%	9%
Severe Mobility Issues	53%	6%	13%
Change in sensation in at least one area of my body	59%	6%	25%

Inability to work or study	63%	6%	9%
Inability to do housework/ daily chores	66%	9%	3%
Inability to do leisure activities	69%	16%	6%
Issues affecting family or personal relationships	63%	16%	6%
Pain	72%	13%	13%
Anxiety	78%	31%	3%
Low Mood	88%	25%	6%
Clinical depression	41%	9%	0%

(Source: VHL UK/Ireland Patient/Carer Survey. Number of respondents = 39)

The following graph shows how many VHL-related surgical procedures the respondents have had in their lifetime that required general anaesthetic.

Proportion of VHL-related surgical procedures the respondents have had in their lifetime that required general anaesthetic



(Source: VHL UK/Ireland Patient/Carer Survey. Number of respondents = 39)

97% have had a least one surgical procedure requiring a General Anaesthetic, 41% have had 5 or more.

This study also highlights the burden of VHL on everyday life, and that many major surgeries are common for a significant number of these patients. Again, this study shows there is a need for a treatment that can help to address these issues.

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

Belzutifan aims to stop or slow the growth of tumours in patients with VHL disease. Belzutifan aims to block hypoxia-inducible factor 2 alpha (HIF-2α). HIF-2α plays a role in oxygen sensing by regulating genes that help the body adaptation to low levels of oxygen. Under normal oxygen levels, a protein in the body named VHL will target HIF-2α for breakdown. If the VHL protein is not functioning properly (as in VHL disease) this results in the build-up of and stabilization HIF-2α. Upon stabilization, HIF-2α starts a messaging process that promotes tumour growth (16).

Belzutifan binds to HIF-2α, and when the VHL protein is not functioning properly, belzutifan blocks the interactions that can lead to tumour growth.

The discovery of the link between the VHL gene and the role of HIF2α was a significant scientific breakthrough, where William Kaelin, Jr., Sir Peter Ratcliffe, and Gregg Semenza were awarded The Nobel Prize in Physiology or Medicine for 2019 for its discovery (17).

Belzutifan was also awarded the first 'Innovation Passport' by the Medicines and Healthcare products Regulatory Agency, National Institute for Health and Care Excellence (NICE) and the Scottish Medicines Consortium (SMC). The Innovative Licensing and Access Pathway (ILAP) aims to accelerate the time to market, facilitating patient access to medicines (18).

When a medicine is approved in the UK two documents are published. One is aimed to inform healthcare staff about the medicine, called the Summary of Product Characteristics (SPC), and the other is aimed to inform patients about the medicine, called the patient information leaflet (PIL) (19). These documents can be found by clicking on the following link: https://products.mhra.gov.uk/product/?product=WELIREG%2040%20MG%20FILM-COATED%20TABLETS

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

Belzutifan is not intended to be used in combination with other medicines for this indication.

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

The recommended dose of Welireg is 120 mg (three 40 mg tablets) to be taken - once a day, at the same time each day. In the event of certain side effects it is recommended that belzutifan is stopped until the side effects improve.

Patients in the study MK-6482-004 continued to take belzutifan unless there were unacceptable side effects or evidence of disease progression.

3d)

Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

A search on clinicaltrials.gov for recruiting, enrolling by invitation, active but not recruiting, or completed studies on belzutifan returns 19 studies. (Search conducted on 6th February 2023). The results of study NCT03401788 provides the data for this appraisal. A further study into VHL disease is being conducted (highlighted in green) which focuses of certain types of tumour in the adrenal glands and pancreas. Further details of these studies can be found by searching for the study name on clinicaltrials.gov.

Study name	Phase	Locations	Study Title	Conditions	n	Interventions	Completion Date
NCT04924075	Phase 2	Global	Belzutifan/MK-6482 for the Treatment of Advanced Pheochromocytoma/Paraganglioma (PPGL), Pancreatic Neuroendocrine Tumor (pNET), or Von Hippel-Lindau (VHL) Disease-Associated Tumors (MK-6482-015)	Pheochromocytoma/Paraganglioma|Pancreatic Neuroendocrine Tumor|Von Hippel-Lindau Disease	232	Drug: Belzutifan	August 12, 2026
NCT04846920	Phase 1	United States	A Study of Belzutifan (MK-6482) in Participants With Advanced Clear Cell Renal Cell Carcinoma (MK-6482-018)	Carcinoma, Renal Cell	52	Drug: Belzutifan	July 17, 2025
NCT04994522	Phase 1	United States	A Study of Belzutifan (MK-6482) in Participants With Renal Impairment (MK-6482-021)	End Stage Renal Disease|Renal Impairment	12	Drug: Belzutifan	March 14, 2023
NCT03445169	Phase 1	United States	A Food Effect Study in Healthy Volunteers With Belzutifan (PT2977, MK-6482) Tablets	Healthy	16	Drug: Belzutifan	May 26, 2018
NCT04976634	Phase 2	Global	Pembrolizumab Plus Lenvatinib in Combination With Belzutifan in Solid Tumors (MK-6482-016)	Carcinoma, Hepatocellular|Colorectal Neoplasms|Pancreatic Ductal Adenocarcinoma|Biliary Tract Neoplasms|Endometrial Neoplasms|Esophageal Neoplasms	730	Drug: Pembrolizumab|Drug: Belzutifan|Drug: Lenvatinib	August 18, 2026
NCT04995484	Phase 1	United States	Belzutifan (MK-6482) Hepatic Impairment Study (MK-6482-020)	Moderate Hepatic Impairment	16	Drug: Belzutifan	March 15, 2023

NCT05239728	Phase 3	Global	A Study of Belzutifan (MK-6482) Plus Pembrolizumab (MK-3475) Versus Placebo Plus Pembrolizumab in Participants With Clear Cell Renal Cell Carcinoma Post Nephrectomy (MK-6482-022)	Clear Cell Renal Cell Carcinoma	1600	Drug: Belzutifan|Biological: Pembrolizumab|Drug: Placebo	January 25, 2030
NCT03634540	Phase 2	United States	A Trial of Belzutifan (PT2977, MK-6482) in Combination With Cabozantinib in Patients With Clear Cell Renal Cell Carcinoma (ccRCC) (MK-6482-003)	Renal Cell Carcinoma (RCC)|Clear Cell Renal Cell Carcinoma (ccRCC)|Kidney Cancer|Renal Cancer|Renal Cell Carcinoma|Renal Cell Cancer Metastatic|Renal Cell Carcinoma Recurrent|Renal Cell Cancer, Recurrent|Kidney	118	Drug: Belzutifan|Drug: Cabozantinib	August 31, 2025
NCT03401788	Phase 2	United States, France, United Kingdom	A Phase 2 Study of Belzutifan (PT2977, MK-6482) for the Treatment of Von Hippel Lindau (VHL) Disease-Associated Renal Cell Carcinoma (RCC) (MK-6482-004)	VHL - Von Hippel-Lindau Syndrome|VHL Gene Mutation|VHL Syndrome|VHL Gene Inactivation|VHL-Associated Renal Cell Carcinoma|VHL-Associated Clear Cell Renal Cell Carcinoma	50	Drug: Belzutifan	March 29, 2026
NCT04736706	Phase 3	Global	A Study of Pembrolizumab (MK-3475) in Combination With Belzutifan (MK-6482) and Lenvatinib (MK-7902), or Pembrolizumab/Quavonlimab (MK-1308A) in Combination With Lenvatinib, Versus Pembrolizumab and Lenvatinib, for Treatment of Advanced Clear Cell Renal Cell Carcinoma (MK-6482-012)	Carcinoma, Renal Cell	1653	Biological: Pembrolizumab|Drug: Belzutifan|Biological: Pembrolizumab/Quavonlimab|Drug: Lenvatinib	October 29, 2026
NCT05468697	Phase 1|Phase 2	United States, Israel, Australia	A Study of Belzutifan (MK-6482) in Combination With Palbociclib Versus Belzutifan Monotherapy in Participants With Advanced Renal Cell Carcinoma (MK-6482-024)	Renal Cell Carcinoma	180	Drug: Belzutifan|Drug: Palbociclib	March 16, 2027
NCT04489771	Phase 2	Global	A Study of Belzutifan (MK-6482) in Participants With Advanced Renal Cell Carcinoma (MK-6482-013)	Carcinoma, Renal Cell	150	Drug: Belzutifan	October 4, 2025

NCT05030506	Phase 1	China	A Study of Belzutifan (MK-6482) as Monotherapy and in Combination With Lenvatinib (E7080/MK-7902) With or Without Pembrolizumab (MK-3475) in China Participants With Advanced Renal Cell Carcinoma (MK-6482-010)	Renal Cell Carcinoma	45	Drug: Belzutifan|Biological: Pembrolizumab|Drug: Lenvatinib	October 21, 2026
NCT02974738	Phase 1		A Trial of Belzutifan (PT2977, MK-6482) Tablets In Patients With Advanced Solid Tumors (MK-6482-001)	Advanced Solid Tumors|Solid Tumor|Solid Carcinoma|Solid Tumor, Adult|ccRCC|RCC, Clear Cell Adenocarcinoma|RCC|Kidney Cancer|Clear Cell Renal Cell Carcinoma|Renal Cell Carcinoma, Metastatic|Renal Cell Carcinoma Recurrent|Renal Cell Carcinoma, Clear Cell Adenocarcinoma|Glioblastoma|Glioblastoma, Adult|GBM|Glioblastoma Multiforme	120	Drug: Belzutifan	April 14, 2025
NCT04195750	Phase 3	Global	A Study of Belzutifan (MK-6482) Versus Everolimus in Participants With Advanced Renal Cell Carcinoma (MK-6482-005)	Carcinoma, Renal Cell	736	Drug: Belzutifan|Drug: Everolimus	September 17, 2025
NCT04586231	Phase 3	Global	A Study of Belzutifan (MK-6482) in Combination With Lenvatinib Versus Cabozantinib for Treatment of Renal Cell Carcinoma (MK-6482-011)	Carcinoma, Renal Cell	708	Drug: Belzutifan|Drug: Lenvatinib|Drug: Cabozantinib	December 23, 2024
NCT04626479	Phase 1|Phase 2	Global	Substudy 03A: A Study of Immune and Targeted Combination Therapies in Participants With First Line (1L) Renal Cell Carcinoma (MK-3475-03A)	Carcinoma, Renal Cell	400	Biological: Pembrolizumab|Biological: Favezelimab/Pembrolizumab|Drug: Belzutifan|Drug: Lenvatinib|Biological: Pembrolizumab/Quavonlimab|Drug: Vibostolimab/Pembrolizumab	March 16, 2026
NCT04989959	Phase 1	United States, France, United Kingdom	[18F]PT2385 PET/CT in Patients With Renal Cell Carcinoma	Renal Cell Carcinoma|Clear Cell Renal Cell Carcinoma	50	Drug: [18F]PT2385|Procedure: Positron Emission Tomography/Computed Tomography|Procedure: Biopsy	August 18, 2026
NCT04626518	Phase 1|Phase 2	Global	Substudy 03B: A Study of Immune and Targeted Combination Therapies in Participants With Second Line Plus (2L+) Renal Cell Carcinoma (MK-3475-03B)	Carcinoma, Renal Cell	370	Biological: Pembrolizumab|Biological: MK-4830|Drug: Belzutifan|Drug: Lenvatinib|Biological: Pembrolizumab/Quavonlimab|Biological: Favezelimab/Pembrolizumab	May 29, 2025

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

There is one study that provides the efficacy data for this appraisal, MK-6482-004. The MK-6482-004 study was a single-arm open-label Phase 2 study that evaluated the efficacy and safety of belzutifan in patients with VHL disease who have at least 1 measurable renal cell carcinoma (RCC), a type of kidney tumour. This means that every patient on the study received belzutifan, and that there wasn't another group (or arm) on the current treatment to compare against (as you would expect in a phase 3 study).

This study set out to see what effect belzutifan had on the size of patient's tumours. A total of 61 patients received belzutifan in the study. The main measures that were taken were the following:

1. **ORR** - measured as a percentage, objective response rate, or ORR, is the proportion of patients in a trial whose tumour is destroyed or significantly reduced by a drug. ORR is generally defined as the sum of complete responses (CRs) – patients with no detectable evidence of a tumour over a specified time period – and partial responses (PRs) – patients with a decrease in tumour size over a specified time period. This is a useful measure for seeing how effective a drug is in shrinking a tumour.
2. **DCR** - measured as a percentage, disease control rate, or DCR, is the proportion of patients in a trial whose tumour is either destroyed, significantly reduced or does not significantly increase in size due to the drug. DCR is generally defined as the sum ORR and the patients who achieve stable disease (SD), where the tumour has neither shrunk or grown significantly. This is a useful measure for seeing how effective a drug is in stopping a tumour growing.
3. **DOR** - measured in months, duration of response, or DOR, is length of time that a tumour continues to respond to treatment without the cancer growing or spreading. Cancer drugs that demonstrate improved DOR can produce a durable, meaningful delay in disease progression, as opposed to a temporary response without any lasting benefit.

Patients who were eligible for the study had at least 1 measurable RCC. The study was designed to see if belzutifan could produce an ORR of 15% in the RCC tumours being measured. This is known as the "null hypothesis" and is considered a successful study if the ORR was found to be significantly greater than 15%. This measurement is known as the study's primary outcome. Other measures were taken, such as the DCR and DOR for RCC tumours. Also, some patients had other tumours such as CNShbs and pNETs, and measurements were also taken to determine the ORR, DCR and DOR of these tumours. All of these measurements were "secondary outcomes" and had no bearing on whether this study was successful or not.

The table below provides the summary results for ORR, DCR and DOR. Please note that in addition to the values given, a range is also provided in brackets. This range refers to an upper and lower estimate between which you can be 95% certain the true value lies, (named 95% confidence interval, CI). The table also mentions "not reached" meaning that the study has not yet been running for long enough for us to make a measurement. Where an "n" is given in the table below this refers to the number of patients.

Summary of MK-6482-004 study efficacy results

Outcome		Summary of results
RCC (all patients, n=61)		
Primary outcome	Overall response rate (ORR)	63.9% (95% CI: 50.6%, 75.8%)
Secondary outcomes	Disease control rate (DCR)	98.4% (95% CI: 91.2%, 100.0%)
	Duration of response (DOR)	Median DOR not reached (range: 5.4+ to 35.8+ months)
Subgroup of patients with CNS hemangioblastoma (n=50)		
Secondary outcomes	Overall response rate (ORR)	44.0% (95% CI: 30.0%, 58.7%)
	Disease control rate (DCR)	90.0% (95% CI: 78.2%, 96.7%)
	Duration of response (DOR)	Median DOR not reached (range: 3.7+ to 38.7+ months)
Subgroup of patients with pNET (n=22)		
Secondary outcomes	Overall response rate (ORR)	90.9% (95% CI: 70.8%, 98.9%)
	Disease control rate (DCR)	100% (95% CI: 84.6%, 100.0%)
	Duration of response (DOR)	Median DOR not reached (range: 11.0+ to 37.3+ months)

Date of Data Cut-off: 01APR2022

Based on the data above we can see that the majority of patients have seen their kidney cancers and their pNETs shrink (63.9% and 90.9% respectively). 44% of patients with a CNShb also saw those tumours shrink. Though the average patient had been on belzutifan for just over for 3 years at the point these measures were taken we don't yet know how long the average patient will continue to see their tumours shrink, though it's likely to be at least 3 years. The disease control rates for all the tumours studied were 90% or more, meaning that 9 in 10 patients had at least their tumours stable during this trial.

Though it wasn't a primary goal of the study, there was also some data collected about the number of surgeries patients had before and after they started belzutifan. More data will be collected on this over time, but it does appear that patients had far fewer surgeries in the years after taking belzutifan compared to the years before. This data is detailed in section B2.6 of document B of the submission.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

No patient reported outcomes were collected in the MK-6482-004 study. However, there are some key insights from the studies mentioned in 2d that can help us estimate the quality of life

benefit of belzutifan. Generally, it can be seen that the more VHL disease has progressed, the worse the quality of life for patients. Given that belzutifan may halt the progression of disease or even shrink tumours, we estimate that patients, for a period of time, may return to a state of lesser disease progression and therefore a better quality of life.

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

The safety of belzutifan was evaluated in the study mentioned in 3e) in 61 patients with VHL disease-associated RCC. Patients were treated with 120 mg belzutifan once daily. The most common side effects with belzutifan were anaemia (90%), fatigue (71%), dizziness (44%) and nausea (36%). The most common side effects which required treatment at hospital were anaemia (10%), and fatigue (5%) (see "further information" for how side effects are graded). Serious side effects occurred in 5% of patients who received belzutifan, including anaemia, dyspnoea and hypoxia (1 patient each). A definition of these side effects is given in the glossary of terms.

The summary of product characteristics (SPC) provides doctors and other hospital staff with information on how to deal with some of the side effects common to belzutifan. This guidance is given below:

Patients should be monitored for anaemia before initiation of and periodically throughout treatment with belzutifan with more frequent monitoring within the first 6 months of treatment . For patients who develop Grade 3 anaemia, belzutifan should be withheld and patients should be treated according to standard medical practice, including ESA administration (see glossary of terms) *until resolved to ≤ Grade 2. For recurrent Grade 3 anaemia, belzutifan should be discontinued. For patients who develop Grade 4 anaemia, the dose of belzutifan should be reduced or permanently discontinued.*

Belzutifan can cause severe hypoxia that may require discontinuation, supplemental oxygen, or hospitalisation. In light of the risk of hypoxia, smoking cessation (stopping smoking) *is recommended. For Grade 2 hypoxia, providing supplemental oxygen and continuing or withholding treatment should be considered. If withheld, belzutifan should be resumed at a reduced dose. For patients who have Grade 3 hypoxia, belzutifan should be withheld, hypoxia treated, and dose reduction should be considered. If Grade 3 hypoxia continues to recur, treatment should be discontinued. For Grade 4 hypoxia, treatment should be permanently discontinued.*

For other side effects no changes of dose are recommended until Grade 3, where it is recommended to stop taking belzutifan until the side effects resolve to a Grade 2 or less. For Grade 4 the advice is to permanently discontinue belzutifan.

It should be noted that the other common side effects (fatigue, dizziness and nausea) are similar to common side effects for patients receiving cancer treatments such as chemotherapy (20). As such the hospitals that will see patients eligible for belzutifan have experience in managing these side effects.

In study MK-6482-004 belzutifan the proportion of participants who stopped taking the drug due to an AE was low (4 or 6.6% of participants).

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration

Below we list the key benefits of belzutifan for patients, caregivers and their communities.

- Belzutifan is the first medicine that has been licenced by the MHRA for the treatment of some VHL associated tumours. Should this medicine be approved by NICE it provides more options for doctors and patients to treat this disease. For patients in the most severe condition, it may provide an alternative option to undesirable surgery with difficult consequences.
- The data from this trial shows that belzutifan has shrunk some types of kidney, pancreatic and brain and spine tumours. This effect appears to last on average at least three years.
- In addition, the data from the trial suggests that the number of surgeries a patient may need after taking belzutifan may be reduced. Patients often experience anxiety, pain and need to take time to recover from surgery. Taking less time off due to fewer surgeries may also provide financial benefits to patients and be less burdensome for caregivers. This may help patients sustain a better quality of life for longer, although it's important to note that this wasn't measured directly in the trial.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

The key disadvantages to patients, caregivers and communities may include the following:

- As there is no medical alternative to belzutifan, patients who were otherwise not taking any medicine for their VHL disease will now have to take three tablets every day. This may have to continue for several years.
- Some patients on belzutifan will have side effects. A small proportion of patients may also experience side effects that require a hospital admission. There may be a chance this small proportion may have to visit the hospital more regularly than otherwise.
- The trial was only in 61 patients. Though the results on tumour shrinkage look very promising for nearly all patients on the trial, it's unlikely that belzutifan will work for everyone.

3i) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?)
- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?
- How the condition, taking the new treatment compared with current treatments affects your quality of life.

Cost-effectiveness relates to how much new health (or quality-adjusted life years, QALYs) the new medicine produces compared to its additional cost (vs. current care), for a typical/average patient and whether the new health is worth the extra cost.

The cost-effectiveness of belzutifan in this indication (vs. current standard of care that patients would otherwise receive which in this case would be surgery resulting in loss of organ function and/or other significant consequences for the majority of patients) is evaluated for the typical/average patient via modelling that uses short-term trial data to extrapolate efficacy and costs over a patient's lifetime.

The challenges of modelling average lifetime outcomes (overall survival, progression of disease and quality of life) from clinical trial data arise from the short-term nature of the trial (MK-6482-004 has around 3 years of patient follow-up data), the limited sample size for each cohort and the trial being single arm so that patients only receive belzutifan (i.e. the trial does not have a comparator arm).

The cost-effectiveness model is used to track a typical/average patient cohort as they move through the patient pathway and produces lifetime outcomes. The typical/average patient here is in line with severe patient population highlighted in the MHRA license and considered to be patients that have exhausted all alternative options for VHL tumour manifestations and are at the "end of the road". As there are no previously published cost-effectiveness models evaluating belzutifan (since it is a new medicine) nor of any other treatments of VHL, a 'de novo' (new) model was developed where patients moved through health states – pre-surgery, surgery, event-free after surgery, metastatic disease and death. These health states represent specific events that have been found to be important differentiators of subsequent outcomes, health-related quality of life and overall survival. The model allowed for one surgery relating to the patient's primary tumour type given the severity of disease. To produce results for the typical/average patient receiving belzutifan (or the current standard of care) for VHL-associated RCC, CNS HB or pNET, all the outcomes are averaged at the end.

How long patients stay in each health state depends on data from the trial. For the period beyond the trial, data extrapolation methods are used (known as "parametric survival models") and there is always uncertainty about which extrapolated curve not only fits the trial data the best, but also

which curve estimates more plausible outcomes in the long-term. This trial did not contain a comparator arm; however, based on the wording of the UK label indication, it is assumed that the majority of patients who require treatment for VHL associated tumours would receive surgery in the absence of belzutifan. As these patients are considered "unsuitable or undesirable" for localised procedures and also "require therapy", these surgeries would have significant impact on organ function or lead to problematic adverse effects. The rate at which events (progression to metastatic disease or death) can occur following surgery were estimated from a real-world study in VHL.

There will also be debates about whether additional adjustments should be made to extrapolations of data observed for belzutifan that make the risks of surgery, progression or death closer to the comparator treatments, this is known as "treatment effect waning". In this appraisal, belzutifan is assumed to have a treatment effect waning period based around its mechanism of action in reducing tumours; hence, when a patient stops taking belzutifan it is assumed that, over time, tumours will eventually grow back to their original size and therefore the risks of these events revert to those expected without belzutifan. Once tumours have grown back to their original size, treatment effect waning is expected to have taken full effect.

Belzutifan reduces both the risk of surgeries and the risk of metastatic disease, by decreasing the size of, or halting the growth, of tumours. Clinical experts assert that this in turn would reduce symptoms associated with VHL tumours, prevent surgery-related complications and keep patients alive for longer than if they were receiving current standard of care.

Quality of life tends to be better for patients in the pre-surgery and event-free after surgery states compared with the metastatic disease state. Given the improved progression-free survival, the typical belzutifan patient will tend to have a better quality of life than a patient receiving current standard of care. The model applies fixed quality-of-life "weights" to time spent in each state. Surgical complications, time spent in each health state and the time spent alive all impact overall quality of life. Side effects of treatment also have an effect, but this is not a big driver of results.

Under this STA process, the NICE committee may apply a greater weight to QALY gains if the product is indicated for a condition with a high degree of severity when compared to the general population of a similar age, this is known as a severity modifier. For belzutifan, the highest severity modifier was met and therefore the highest cost-effectiveness threshold should apply for this condition if considered under the STA process. Cost-effectiveness models are a simplification of disease pathways and in the case of belzutifan it does not capture the full benefit of belzutifan resulting in the company's base case estimate being marginally above the decision-making threshold to be deemed "cost-effective". Aspects specifically not captured are the societal benefits of increased work productivity for both patients and carers, reduced out-of-pocket costs for patients, reduced need for government assistance and the additive clinical benefit of belzutifan when acting to reduce tumour burden across multiple sites simultaneously. These factors result in an underestimation of the value of belzutifan in VHL disease.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

As mentioned in previous sections, belzutifan is the first medicine to be licenced for the treatment of VHL disease. Should the medicine be recommended by NICE then doctors and other hospital staff in England and Wales will be able to provide an alternative to "undesirable" surgery.

The development of this medicine came from Nobel prize winning science (17). NICE and the MHRA also awarded belzutifan the first "innovation passport" which aims to accelerate access to innovative medicines.

Detailed in 3f, the patient benefits of a reduced number of surgeries (including fewer absences from school or work) are not currently captured in the economic model but are likely to have a large positive impact for patients.

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.
Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics

More information on how NICE deals with equalities issues can be found in the NICE equality scheme
Find more general information about the Equality Act and equalities issues here

No equality issues are anticipated.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc.
Where possible, please provide open access materials or provide copies that patients can access.

CTCAE grading

In oncology clinical trials, the severity of adverse events are usually graded according to US National Cancer Institute's AE Severity Grading Scale - Common Terminology Criteria for Adverse Events (CTCAE) (21). CTCAE can also be used to grade the AE for non-oncology studies, but generally not appropriate for studies using healthy volunteers.

- Grade 1 Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; no intervention indicated
- Grade 2 Moderate; minimal, local or non-invasive intervention indicated; limiting age-appropriate instrumental ADL
- Grade 3 Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living (e.g. bathing, dressing or feeding).
- Grade 4 Life-threatening consequences; urgent intervention indicated.
- Grade 5 Death related to AE.

Further information on NICE and the role of patients**:**

- Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE
- NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE
- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

Anaemia - A low red-blood count. Your blood does not have enough of the cells that carry oxygen (haemoglobin) to your body. Also called "tired blood" or "low iron".
Dyspnoea - When you have trouble breathing.
ESA, erythropoiesis-stimulating agent - A medicine that stimulates the bone marrow to make more red blood cells
Fatigue - tired, weak feeling of the whole body, feeling tired all over.
Hypoxia - is low levels of oxygen in your body tissues. It causes symptoms like confusion, restlessness, difficulty breathing, rapid heart rate, and bluish skin.
Nausea - When you have an upset stomach or feel like throwing up.

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

Response:

1. MHRA. WELIREG 40 MG FILM-COATED TABLETS. Summary of Product Characteristics [Internet]. 2022. Available from: https://products.mhra.gov.uk/search/?search=welireg&page=1&doc=Spc&rerouteType=0.
2. Maher ER, Adlard J, Barwell J, Brady AF, Brennan P, Cook J, et al. Evaluation of tumour surveillance protocols and outcomes in von Hippel-Lindau disease in a national health service. Br J Cancer. 2022.
3. NHS. Von Hippel-Lindau disease (VHL) - Information for patients, families and carers 2021. Available from: https://www.uhs.nhs.uk/Media/UHS-website-2019/Patientinformation/Genetics/Von-Hippel-Lindau-disease-VHL-1438-PIL.pdf.

4. Wilding A, Ingham SL, Lalloo F, Clancy T, Huson SM, Moran A, et al. Life expectancy in hereditary cancer predisposing diseases: an observational study. J Med Genet. 2012;49(4):264-9.

5. VHL UK / Ireland. Stories: VHL UK / Ireland; 2022 [Available from: https://vhl-uk-ireland.org/stories/.

6. NHS. Von Hippel-Lindau disease2020. Available from: https://www.genomicseducation.hee.nhs.uk/wp-content/uploads/2019/05/Von-Hippel-Lindau-disease.pdf.

7. Maher ER, Neumann HP, Richard S. von Hippel-Lindau disease: a clinical and scientific review. Eur J Hum Genet. 2011;19(6):617-23.

8. NHS. Radical Nephrectomy. Patient Information [Internet]. 2019. Available from: https://www.uhcw.nhs.uk/download/clientfiles/files/Patient%20Information%20Leaflets/Surgical%20Services/Urology/119440_Radical_Nephrectomy_(741).pdf.

9. Fernando A, Fowler S, Van Hemelrijck M, O'Brien T. Who is at risk of death from nephrectomy? An analysis of thirty-day mortality after 21 380 nephrectomies in 3 years of the British Association of Urological Surgeons (BAUS) National Nephrectomy Audit. BJU Int. 2017;120(3):358-64.

10. NHS. Brain tumours: National Health Service; 2020 [updated 03-FEB-2020. Available from: https://www.nhs.uk/conditions/brain-tumours/.

11. Thompson RH, Hill JR, Babayev Y, Cronin A, Kaag M, Kundu S, et al. Metastatic renal cell carcinoma risk according to tumor size. J Urol. 2009;182(1):41-5.

12. CRUK. Daily life with kidney cancer2020. Available from: https://www.cancerresearchuk.org/about-cancer/kidney-cancer/living-with/daily-life.

13. Cui Y, Andersen DK. Pancreatogenic diabetes: special considerations for management. Pancreatology. 2011;11(3):279-94.

14. MSD. VHL-related Cancer Patient Survey. [MSD Data On File] [Internet]. 2022.

15. VHL UK / Ireland. VHL UK/Ireland Patient/Carer Survey2022. Available from: teamnice@vhl-uk-ireland.org.

16. MSD. Mechanism of Action - Welireg: Merck Sharp & Dohme; 2022 [Available from: https://www.welireghcp.com/mechanism-of-action/.

17. Nobel Foundation. 2019 Nobel Prize in Physiology or Medicine2019. Available from: https://www.nobelprize.org/uploads/2019/10/press-medicine2019.pdf.

18. MHRA. First Innovation Passport awarded to help support development and access to cutting-edge medicines: Medicines and Healthcare products Regulatory Agency; 2021 [updated 26-FEB-2021. Available from: https://www.gov.uk/government/news/first-innovation-passport-awarded-to-help-support-development-and-access-to-cutting-edge-medicines.

19. emc. WELIREG 40 mg film-coated tablets - Package leaflet: Information for the patient 2022 [updated 18-OCT-2022. Available from: https://www.medicines.org.uk/emc/product/14126/pil.

20. NHS. Side effects - Chemotherapy: National Health Service; 2020 [updated 29-JAN-2020. Available from: https://www.nhs.uk/conditions/chemotherapy/side-effects/.

21. CTEP. Common Terminology Criteria for Adverse Events (CTCAE) Version 5.02017. Available from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Belzutifan for treating tumours associated with von Hippel-Lindau disease

[ID3932]

Clarification questions

28 March 2023

Response submitted: 17 April 2023

File name	Version	Contains confidential information	Date
NICE ID3932 Clarification letter – MSD response [ACIC] v1.0	1.0	Yes	17-APR-2023

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Literature searches (clinical effectiveness)

A 1. Please provide the strategies used for the ClinicalTrials.gov search in Appendix D and the conference proceedings searches in Appendices D, G and H.

MSD response:

With regard to the ClinicalTrials.gov search in Appendix D, a search for all studies conducted on patients with VHL disease that reported results was performed.

For the search of conference proceedings used for the SLR of clinical effectiveness evidence documented in Appendix D of the company submission, the search strategy shown in Table 1 which was run in the Northern Light Life Sciences Conference Abstracts database was used.

Table 1 Search strategy used for the search for conference proceedings for the SLR of clinical effectiveness evidence in Appendix D of the company submission

#	Searches
1	renal cell carcinoma.mp.
2	((renal or kidney) adj2 cell adj2 (carcinoma or cancer* or cancer* or malignan* or tumor* or tumour* or neoplas* or adenocarcinoma*)).ti,ab.

#	Searches
3	(renal cell cancer or RCC or renal cell carcinoma or kidney cancer or kidney carcinoma).ti,ab.
4	or/1-3
5	(von Hippel Lindau Disease or Familial Cerebello-Retinal Angiomatosis or Angiomatoses, Familial Cerebello-Retinal or Angiomatosis, Familial Cerebello-Retinal or Cerebello-Retinal Angiomatoses, Familial or Cerebello-Retinal Angiomatosis, Familial or Familial Cerebello Retinal Angiomatosis or Familial Cerebello-Retinal Angiomatoses or Hippel-Lindau Disease or Hippel Lindau Disease or VHL Syndrome or VHL Syndromes or Lindau's Disease or Lindau's Diseases or Lindaus Disease or von Hippel-Lindau Syndrome or von Hippel Lindau Syndrome or Angiomatosis Retinae or Cerebelloretinal Angiomatosis, Familial or Angiomatoses, Familial Cerebelloretinal or Angiomatosis, Familial Cerebelloretinal or Cerebelloretinal Angiomatoses, Familial or Familial Cerebelloretinal Angiomatoses or Familial Cerebelloretinal Angiomatosis or Lindau Disease).mp.
6	4 and 5
7	American Society of Clinical Oncology.cs.
8	European Society for Medical Oncology.cs.
9	7 or 8
10	6 and 9

For the search of conference proceedings used for the SLR of cost-effectiveness and HRQoL evidence documented in Appendix G and H respectively of the company submission, all the conferences except for NCCN and AAO were indexed in Embase hence, were not manually searched. For NCCN and AAO PDFs of abstract proceedings were downloaded and screened manually, and no keyword searching was performed for these conferences. Table 2 below provides results of these searches.

Table 2 Results of searches from conference proceedings

Conference name	Year	Keywords	Hits	Number of relevant unique abstracts found
NCCN (a pdf document containing abstract of conference proceedings was screened hence keyword searching was not performed)	2022	-	-	0
	2021	-	-	0
	2020	-	-	0
AAO (a pdf document containing abstract of conference proceedings was screened hence keyword searching was not performed)	2022	-	-	0
	2021	-	-	0
	2020	-	-	0

A 2. No search terms relating to the VHL Natural History Study are included in the study design filters for clinical effectiveness.

a) Please explain why terms for the VHL Natural History Study are not included in the study design filters for clinical effectiveness in Appendix D, given that the main comparator study is a natural history study.

b) Given the above, please explain how the VHL Natural History Study was identified.

MSD response:

a) Search terms specifically for the VHL Natural History Study were not included in the study design filters for clinical effectiveness data, as the VHL Natural History Study is a (currently) unpublished study specifically commissioned by MSD to address the lack of available relevant comparator data in the published literature for this indication.

b) The VHL Natural History Study, not (yet) being a published study, was not (and could not be) identified via a systematic literature review. The VHL Natural History Study is a retrospective non-interventional study commissioned by MSD.

Literature searches (cost-effectiveness)

A 3. The search methods for all cost effectiveness and HRQoL searches (Appendices G and H) report a search of MEDLINE and Embase via Embase.com. Please confirm whether this refers to a search of Embase only, conducted on the understanding that it contains all records from MEDLINE.

MSD response:

We can confirm a search of Embase only was conducted on the understanding it contains all records from MEDLINE. All Embase searches reported in these appendices include search results from MEDLINE and Embase.

A 4. In the cost-effectiveness SLR, Table 125 appears to be a search of all PubMed records, rather than just MEDLINE In Process, as no limit (such as 'inprocess[sb]') appears to have been applied to only identify 'in process' records. Please confirm whether this is the case.

MSD response:

This search was not restricted to *in process/ahead of print* results hence, no “AND (inprocess[sb] OR pubstatusaheadofprint)” string was used however, please note the search results of this search strategy still included in process and ahead of print records. To clarify, PubMed was searched in addition to Embase to additionally identify ‘in process’ records; however, the search in PubMed was not restricted.

A 5. In Appendices G and H the update searches are for a narrower population than the original searches. The original searches are for all patients with VHL, whereas the update searches only identify records where VHL terms occur in conjunction with CNS hemangioblastomas, pNETs or RCC terms. Please explain the discrepancy between these two approaches, and what effect this may have had on the update search results.

MSD response:

The original searches were conducted in 2020 before the MHRA GB marketing authorisation was granted on 31-MAY-2022 with wording specifically for patients with VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET). Therefore, the inclusion criteria for the population was updated to include these manifestations; however, as stated in Table 129 of the CS “*In absence of a clear reporting about the diagnosis method, studies that mention ‘VHL disease’ will also be included*”. Hence, the updated search in July 2022 has not restricted the population but rather more closely align with the GB marketing authorisation and NICE scope for this appraisal.

A 6. In Appendix G, the title of the update search for Embase.com states the search dates as being from July 1, 2020 to July 26, 2022. However, line #35 of the strategy appears to limit the results to records added since 01-06-2020. Please confirm which starting date is correct.

MSD response:

We apologise for the mistake here; this is a typographical error. The correct limit is [01-07-2020]/sd NOT [26-07-2022]/sd.

A 7. Please confirm the exact dates on which all the searches in Appendices G and H were conducted.

MSD response:

Original searches were performed from database inception to July 1, 2020, and the updated searches was performed from July 1, 2020, to July 26, 2022.

A 8. The study design filter used in Appendix H for the original PubMed search appears to be much narrower in scope than the filter used on other databases. For example, the filters used in line #2 of Table 142 are much narrower than the filters in line #2 of Table 140. Please explain this discrepancy, and what effect this may have had on the results found.

MSD response:

We used only the search terms available in Pubmed, several of the terms such as the following were not found in Pubmed: "health year equivalent, "disutility", "disfigure", "european organization for research and treatment of cancer questionnaire", "routine electronic monitoring of hrqol", "fact-ksi". Therefore, the search appears narrower, however given the fact that we used indexed terms and free-text searches, this search is not likely to have lesser sensitivity than the 'broader' one.

Decision problem

A 9. Priority. Table 2 states: "No treatments for advanced or metastatic disease are relevant as comparators as these would be used after treatment with belzutifan. The purpose of belzutifan is to prevent tumours reaching the advanced or metastatic stage."

a) Please confirm that the population in the decision problem should be re-expressed as excluding advanced or metastatic stage.

b) **If this is not the case then please include evidence for advanced or metastatic stage with comparators appropriate for this stage, including monotherapy or combination therapy with immunotherapies or kinase inhibitors, as stated in the NICE scope.**

MSD response:

MSD confirms that the population of the decision problem excludes advanced or metastatic stage disease, as explicitly stated in the company submission in Document A Section A.2, Section A.5, Document B Section B.1.1 Table 2 and Section B.1.3, and in MSD's consultee comments on the draft scope to this appraisal as documented in *NICE's response to comments on the draft scope and provisional stakeholder list* (published in: https://www.nice.org.uk/guidance/gid-ta10817/documents/scope-consultation-comments-and-responses, specifically on page 18).

A 10. Priority. The company stated: "While some patients in the study also had VHL-associated CNS hemangioblastomas and/or VHL-associated pNETs, all patients had VHL-associated RCC. This therefore means that the population of the MK-6482-004 study does not align with (i.e. is narrower than, in this respect) the marketing authorisation for belzutifan as described previously in section B.1), the population under consideration in this assessment." It appears that the company are asserting that the population is narrower because not all patients had all types of tumour (CNS, pNET and RCC). However, the marketing authorisation appears to not require this, but instead that patients have to have at least one of those three tumour types, given that 'or' is used in the list (see Tables 1 and 3 in the CS). Also, a separate CEA is conducted for each of three populations, VHL-associated CNS hemangioblastomas, VHL-associated pNETs, and VHL-associated RCC, referred to as "primary tumour site" (p. 132). In each model it appears that there is the possibility of multiple tumour types and surgery for multiple tumour types e.g., RCC and CNS or RCC and pNET. Additional tumours are described as "non-primary tumours" (p.132). However, the clinical effectiveness section, in counting the number of each type of tumour in the MK-6482-004 study, does not differentiate by whether primary

or not: all 61 have tumours at baseline of type RCC, 50 have CNS+RCC (referred to as CNS subgroup), 22 have pNET+RCC (referred to as pNET subgroup) and 17 (see Figure 9) have all three.

a) **Please clarify what the intended population in the decision problem is i.e., all three tumour types or at least one of them.**

b) **Please explain the relevance of whether a tumour of is primary or not if no distinction is made in the MK-6482-004 study and the same treatment i.e., surgery is possible for both the type that defines the subgroup and any additional type in the CEA.**

c) **Given that all patients in the MK-6482-004 study have an RCC tumour, please clarify whether the intended population of the decision problem must include an RCC tumour and if that has to be the primary tumour.**

d) **Please specify the nature of the population in the decision problem and UK clinical practice in terms of the proportions of patients in each of the main tumour type combination subgroup: RCC only, CNS only, pNET only, CNS+RCC, pNET+RCC, CNS+pNET, all three.**

MSD Response:

a) The intended population in the decision problem is patients with *at least one* of the tumour types specified in the GB marketing authorisation.

We comment the study is narrower than the MA for the following reason: we consider the MA to permit patients who have a CNS hemangioblastoma tumour but *no* RCC tumour to receive belzutifan. MK-6482-004 does not have any CNS hemangioblastoma participants who do not also have RCC (as an example).

b) In cancer treatment, primary tumour refers to the original tumour locus, e.g., whether the original tumour was located in the lungs, is then lung cancer that then may become advanced or metastasise, for example. This is *not* how the term is used in this submission, and we apologise for any confusion. The primary tumour is the tumour that is driving treatment decisions. For example, a patient might have an RCC tumour and an intrusive CNS hemangioblastoma tumour. In this hypothetical

example the CNS hemangioblastoma would be considered the primary tumour as it is driving treatment decisions.

It is critical to know which is the primary tumour, as the treatments / *localised procedures* are not homogenous interventions. There is a substantive difference in the possible interventions and sequalae for an RCC treatment/localised procedure compared with treatment options/localised procedures that might be considered for an intrusive CNS hemangioblastoma. As described in section B.1.3 of the company submission, the relevant localised procedures for these primary tumours are:

- For RCC tumours, the localised procedures relevant for this patient population includes radical (i.e. full, bilateral) nephrectomies which would lead to dialysis dependency.

- For pNETs, localised procedures relevant for this patient population includes Whipple procedures/ pancreatectomies and splenectomies (A Whipple surgery involves removal of the head of the pancreas, a part of the duodenum, some of the bile duct and the gall bladder. The stomach, bile duct and remainder of the pancreas will then be rejoined to the small bowel. The operation usually takes 4-6 hours. Long-term effects of this procedure include but are not limited to diabetes, pancreatic insufficiency, and change in bowel habit).

- For CNS hemangioblastomas, treatment might be radiotherapy, stereotactic radiosurgery and neurosurgery including procedures on those close to the brain stem and in the spinal cord which have a significantly increased risk of morbidity/mortality.

This difference in the possible interventions and the possible treatment sequalae by “primary tumour” needs to be explored in the economic model, hence three “primary models” with an analogous model structure, because of different costs and different sequelae.

We note that all patients in the MK-6482-004 study have RCC. However, in order to emulate what we think the marketing authorisation was seeking to achieve, we have assumed for the pNET cohort in the economic model that the primary tumour of concern is a pNET. Similarly for the CNS hemangioblastoma cohort, we have

assumed the primary tumour of concern of a CNS hemangioblastoma. This is a simplifying assumption required due to the limitations of the model structure and the dataset.

c) We agree with the EAG's previous comment here, '*the marketing authorisation appears to not require this [that a patient must have all 3 tumours], but instead that patients have to have at least one of those three tumour types, given that 'or' is used in the list'.* We do not believe it is the intention of the marketing authorisation to require patients to have an RCC to receive treatment with belzutifan.

The population of the decision problem is that which is specified in the MHRA GB marketing authorisation ("Welireg is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable"). The discrepancy between the population recruited in the MK-6482-004 and the population described in the GB marketing authorisation and how this discrepancy came to be is explained in the company submission: In Document A Section A.9, Document B at the end of the "Definitions and descriptions of key terms used in the submission" section, section B.2.3, and section B.2.12.

d) There is a lack of published data on the proportion of eligible patients with VHL disease in UK clinical practice specifically with RCC only, CNS only, pNET only, CNS+RCC, pNET+RCC, CNS+pNET, and all three.

MSD agrees it would be helpful to have better visibility on this. In the MK-6482-004 study where all participants necessarily had VHL disease *and* RCC, 50 (82%) of these participants had RCC + CNS hemangioblastomas, 22 (36%) had pNETs, and 17 (28%) had RCC + CNS hemangioblastomas + pNETs at baseline as confirmed by independent review committee (IRC) assessment (see section B.2.7 of the company submission). As the objective of the MK-6482-004 study was to enrol patients with RCC, it is possible that this does not reflect the distribution of these tumours in a non "RCC-selected" population expected to receive belzutifan in UK clinical practice. However, due to the rarity of the disease, MSD has not been able to validate

whether these proportions in the MK-6482-004 study reflect those seen in UK clinical practice with any external data sources which report this data.

Without being able to identify the primary tumour in patients who have more than one tumour manifestation in the MK-6482-004 trial, which was designed prior to the GB marketing authorisation, each modelled VHL cohort was defined for the purposes of subgroup analysis as the subset of patients who all had a specific VHL-related tumour manifestation (i.e. 100% of patients in the VHL-RCC cohort had an RCC).

Patients with more than one tumour manifestation are therefore included in each cohort respective to their manifestation. For example, a patient with both RCC and pNET tumour manifestations provides data in two analyses: the VHL-RCC cohort in which RCC is the primary tumour (in the VHL-RCC cohort making pNET a non-primary tumour) and in the pNET cohort in which pNET is considered the primary tumour (in the VHL-pNET cohort making RCC a non-primary tumour). This is a simplifying assumption as no data is available on which is the primary tumour in patients with more than one tumour manifestation. All 61 patients are modelled through the VHL-RCC cohort. Patients in the VHL-CNS Hb and VHL-pNET cohorts would also be included in the VHL-RCC cohort as, by definition, patients had to have RCC to be included in the MK-6482-004 trial. Non-primary tumour (including, for example, a pNET in the VHL-RCC cohort) surgery incidence rates were incorporated into the economic model by including the associated costs and QALY decrements (see response to B 2. (d) below for more detail).

This represented a pragmatic approach to conducting analyses on patient subgroups who are not perfectly distinct and for whom data available in the clinical trial and from UK clinical practice is limited.

A 11. The decision problem (Table 2) does not mention tumour types other than RCC, CNS haemangioblastomas or pNETs. However, Sections B.2.4 (Table 12) and B.2.7 (“Other tumours”) mention participants recruited to the MK-6482-004 study having: non pNET pancreatic lesions; retinal haemangioblastomas; adrenal lesions; endolymphatic sac tumours; and epididymal cystadenomas. Please clarify whether participants who had these tumour types in addition to RCC, CNS haemangioblastomas and/or pNETs were relevant to the decision problem.

MSD Response:

These patients *are* relevant to the decision problem, however, the presence of these other tumours in the absence one of the three primary tumour types (RCC, pNETs, or CNS hemangioblastoma) specified in the MA, we believe, would make a patient ineligible for treatment with belzutifan. For example, a patient with only a retinal hemangioblastoma and no RCC, pNET or CNS hemangioblastoma tumour would *not* be eligible for treatment with belzutifan. These other "non-primary" tumours are relevant to the decision problem because they may have a significant impact on the patients' quality of life. For example, we believe there is a substantive difference between a patient that has only an RCC tumour compared with a patient that has an RCC tumour, a CNS hemangioblastoma, a retinal hemangioblastoma and a non-pNET pancreatic lesion. We have referred to this latter type of patient as having "multi-systemic disease".

There is evidence, although in very small numbers of patients, that belzutifan provides a clinical benefit in these non-primary tumours that we believe should be modelled in the CEA in order to fully evaluate the burden of disease in VHL and the potential for belzutifan to alleviate this burden.

We do not think it is a requirement for these "non-primary tumours" to be explicitly referenced in the decision problem statement as they are not directly referenced in the marketing authorisation but are a reality that patients with VHL must contend with.

A 12. Priority. In Table 2 of the decision problem, regarding population, the company states "Adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable." However, the company recognises that there is 'misalignment' between 'require immediate therapy' and 'for whom localised procedures are unsuitable or undesirable', and the MK-6482-004 study where patients who had an immediate need for surgical intervention for tumour treatment were excluded and there was no requirement for the latter criterion. It is also stated that: *"Patients must have sufficient organ*

function (as described in the MK-6482-004 study participant eligibility criteria…to be eligible to receive belzutifan." **(p. 38)**

a) **Please clarify that not requiring immediate surgery is the same as not requiring therapy. Please clarify that the implication of this particular form of misalignment is that the patients in the MK-6482-004 study are at an earlier and less severe stage of disease than those in the decision problem.**

b) **Please clarify that the implication of the misalignment in 'for whom localised procedures are unsuitable or undesirable' is also that the patients in the MK-6482-004 study are at an earlier and less severe stage of disease than those in the decision problem.**

c) **Please clarify that there is a misalignment in having sufficient organ function with the decision problem population.**

d) **Please explain how the eligibility criteria applied to the VHL Natural history study improve alignment with the MK-6482-004 study in terms of these three criteria i.e. requiring therapy, localised procedures being unsuitable or undesirable and having sufficient organ function.**

e) **One of the exclusion criteria applied to the MK-6482-004 study is: *"RCC tumour greater than 3.0 cm that requires immediate surgical intervention*" and the VHL Natural history study is: "If the largest renal solid tumor at patient-level index date is ≥30 millimetres (mm), patients with a renal surgical procedure with therapeutic intent performed within 60 days on or after patient-level index date" (p. 107). However, it is unclear the degree to which these criteria align. In particular, immediacy might imply greater severity and thus poorer prognosis and therapeutic intent might imply that the procedure was regarded as suitable or desirable and therefore associated with better prognosis. Please clarify what is meant by 'therapeutic intent' and the implications of these exclusion criteria are and the degree to which alignment is improved or worsened.**

f) Please clarify that there is no inclusion criterion for sufficient organ function applied to the VHL Natural history study. If this is the case, then please discuss the implications of this mismatch with the MK-6482-004 study, and, if possible, explore the impact on all ITC analyses of applying such a criterion.

MSD response:

a) It is the company's position that not requiring immediate surgery is *not* the same as not requiring therapy. There is potentially a typo in the question above ("'require immediate therapy'"), as the indication wording is "requiring therapy".

The initial "not requiring immediate surgery" is a requirement of the trial protocol in a novel product. We assume this criterion was included on the basis that giving a patient a treatment that, at the time of trial protocol development, had unproven efficacy, was not ethical. In line with standard clinical practice, the requirement of immediate surgery would occur when patients had an RCC tumour in the range of 3cm and would under usual conditions be removed surgically to minimise the likelihood of metastasis.

It is the company's understanding that the MHRA, having reviewed data demonstrating meaningful efficacy, did not intend belzutifan to be used in all VHL patients, only those who *did* require an intervention.

The implication of this is that it is possible *some* of the patients in the MK-6482-004 study *may* have less severe disease manifestations than those covered by the marketing authorisation (and therefore the decision problem). We caution against using terms such as "earlier and less severe stage of disease" as this indicates "cancer" thinking. VHL has far more varied presentation than typical cancers. Therefore, it is probably more helpful to think in terms of severe or less severe manifestations of the disease.

b) We clarify that the implication of "for whom localised procedures are unsuitable or undesirable" is that *some* patients in the MK-6482-004 study *may* have a less severe presentation of VHL than the population intended by the marketing authorisation and therefore the decision problem. As previously, we suggest avoiding the phrase

‘earlier and less severe’ and instead refer to more or less severe manifestations of VHL disease.

c) There is no misalignment here. The marketing authorisation and clinical experts specify belzutifan will be used at a “fork in the road” in terms of treatment decision making. The trial inclusion criteria required patients to have sufficient organ function, therefore, this is included only to highlight that there are currently no data on the efficacy of belzutifan in patients whose organ function is worse than that specified by the trial inclusion criteria.

d) The VHL Natural History Study, similar to the MK-6482-004 study, was initiated prior the final indication wording received from the MHRA. Therefore, the VHL Natural History Study aligns very well with the trial protocol. The same discrepancies exist between the MHRA indicated population and the VHL Natural History Study as exists between the MHRA indicated population and the trial population. The fact such discrepancies exist is not an ideal situation, however this is the inevitable reality with sourcing data for treatments for very rare indications with highly heterogeneous presentations.

e) The difference in the nature and wording of these exclusion criteria is due to the MK-6482-004 study being a prospective study while the VHL Natural History Study is a retrospective one, and not to do with any meaningful differences between “immediacy” and “therapeutic intent”.

In order for the investigators of the prospective MK-6482-004 study to exclude patients who at baseline would receive a surgical intervention (with therapeutic intent, the only type of surgical intervention of relevance in this context) at that point in time under normal clinical practice, they necessarily needed to define/exclude patients who they determined to require immediate surgical intervention. Whereas for the VHL Natural History Study, being retrospective in nature, it was possible to confidently determine/exclude which patients would have received a surgical intervention under normal clinical practice by defining/excluding those patients who did actually undergo such procedures.

f) The requirement of sufficient organ function in the VHL Natural History Study is largely irrelevant as patients in the study were not on active systemic treatments (the

sufficient organ function eligibility criteria was implemented in the MK-6482-004 to ensure participants had sufficient kidney and liver function to safely process a pharmacologic intervention, such as belzutifan, as is standard for such studies). In terms of whether this means the patients included in the VHL Natural History Study have a more or less severe presentation of the disease, we would suggest that *some* patients in the VHL Natural History Study *may* have a less severe presentation, in the same way as the study population. Given the rarity of this disease and its highly heterogenous presentation we consider any potential impact of this requirement on clinical efficacy estimates or in terms of creating uncertainty within the economic analysis to be largely irrelevant, given the complexity of this decision problem.

A 13. Priority. The nature of comparators is not clear for all subgroups based on tumour type/combination of different tumour types (see Question A10.). On page 131 of the CS, it is stated that *"In routine clinical practice, the decision point for a patient meeting the criteria of belzutifan eligibility would have three options: 1) surgery that is unsuitable or undesirable because it results in loss of organ function, 2) active surveillance to monitor a tumour that is above 3cm (RCC) or 2cm (pNETs) and therefore there is an increased risk of metastatic disease and/or other symptoms of tumour burden (particularly in CNS Hb tumours), or 3) belzutifan".*

a) **Please clarify whether options 1) and 2) above in some proportions define the comparator generally.**

b) **Please provide information on the nature of relevant comparators for patients in subgroups based on tumour type/combination of different tumour types, e.g., if a patient has more than one tumour type then the comparator might best be described by the combination of treatments for each tumour type, such as the two types of surgery.**

c) **Related to the above, please explain the rationale for selection of comparators for patients in subgroups based on tumour type/combination of different tumour types.**

d) **Please provide estimates for the percentage of patients receiving surgery as opposed to active surveillance in subgroups based on**

tumour type/combination of different tumour types, as would be expected to be SoC in the UK.

MSD response:

a) We can confirm that the understanding is correct; surgery or active surveillance does comprise our comparator arm (please see response to part (d) for a full explanation of what active surveillance in this appraisal refers to). Patients at the treatment decision point are facing a choice of having surgery that will have significant and severe consequences or *not* having surgery and increasing the risk of the cancer becoming advanced/metastatic (RCC/pNET) or increasingly symptomatically burdensome (CNS Hb).

b) Based on our discussions with clinicians it is difficult to apply "rules" to the scenario in this question. It is dependent upon the individual's presentation and their personal circumstances. The treatment rationale we model is that there is a dominant/primary tumour that is driving treatment decisions. Specific surgical procedures (e.g., full nephrectomy, Whipple's procedure) are not used in the model due to complexity and lack of available data on this aspect, instead a simplifying assumption is made where each surgical procedure is considered for the specific tumour type (i.e. RCC, CNS hemangioblastoma, or PNET). The associated costs are sourced using HRG codes that are not for a specific surgical procedure but reflect the complexity of these procedures in the target population.

Nothing in the clinical trial nor in our clinician engagements suggests that patients routinely have surgeries that deal with two tumour sites at one time (e.g., CNS and renal resections in one surgery). It is plausible that a patient may have multiple tumours in one organ resected in a single surgical procedure. Therefore, while clinicians are dealing with the dominant tumour, for some patients, the other tumours continue to grow with currently available SoC procedures. This negative element of current SoC is not something we have been able to model.

Furthermore, patients with multiple tumour manifestations are modelled in the CEA in each relevant cohort as described in response to A 10. d) above; they are modelled as though either RCC, CNS Hb or pNET is the primary tumour depending on which tumour manifestations they present.

c) The comparator arm is broadly described as surgery with poor outcomes for the majority of patients with a small proportion receiving active surveillance in the RCC/pNET cohorts. As described above, it is very difficult to apply “rules” to the scenario in prior question. Therefore, patients with multiple tumour manifestations are modelled in multiple cohorts as though each manifestation is the primary tumour.

In the model, surgical procedures are distinguished by tumour manifestation (i.e. RCC, CNS Hb, pNET) but not by specific type of surgical procedure (i.e. Whipple’s procedure). Costs are therefore applied to that effect, using HRG codes applying to surgical procedures relating to the tumour type. Disutilities associated with complications expected from these procedures (as identified in the Optum study) are also then applied, with a different set of complications for each primary tumour type surgery.

d) We do not think patients who are well-managed with active surveillance meet the MHRA’s eligibility requirement for belzutifan: patients “require therapy”. As such patients that are well managed with active surveillance would not be typical belzutifan patients and therefore are not SoC in this appraisal. However, there are some patients who do not ‘take the risk’ with surgery though they know it will increase their risk of tumours becoming advanced. These patients are still scanned (receive active surveillance) but they are not typical ‘active surveillance’ patients.

For patients with a less severe manifestation and for the VHL population in general active surveillance is the foundation/backbone of care. Belzutifan is for when surveillance has identified something that requires an intervention (the patient “requires therapy”).

For the patient population relevant to this appraisal, we would like to distinguish between patients for whom active surveillance is the most appropriate care and patients who continue to be monitored/scanned but for whom there are no (good) treatment options. For some VHL patients there are no reasonable treatment options even after active surveillance has revealed the presence of a tumour that should be resected. Those that do not have surgery progress to advanced and metastatic cancer (which is not the objective of general active surveillance). General active surveillance is a precursor to a treatment intervention however, the active

surveillance termed in the statement highlighted in the initial question (page 131 of the CS) describes surveillance for patients who require therapy but have no reasonable treatment options. We apologise for the confusion in the term “active surveillance” in our CS and would like to clarify that here this refers to patients who have run out of treatment options, *not* those who are well-managed with regular monitoring.

A 14. A general definition of the term “*localised procedures*” is provided in Table 1 of Document B.

a) Please provide specific definitions of the “*localised procedures*” that are relevant for each patient subgroup based on tumour type/combination of tumour types.

MSD response

As detailed in the response to question A 13. b), it is very difficult to apply “rules” for comparator localised procedures based on each specific tumour type and combination of tumour types; it is entirely dependent upon the individual’s presentation and their personal circumstances. The treatment rationale we model is that there is usually a dominant or primary tumour that is driving treatment decisions and that patients do not have surgeries that deal with two or more tumour sites at one time.

Document B Table 1 of the company submission indeed provides a general definition of the term “localised procedures” though specific examples are provided as well: “This includes radiotherapy, radiofrequency ablation, thermo-ablation, cryoablation, microwave ablation, irreversible electroporation, and any other image-guided ablation targeted at these tumour(s), and all surgical procedures with the objective of removing or reducing the size of the tumour”. With regard to localised procedures used in clinical practice for the relevant VHL disease associated tumours:

- For RCC these would be radiofrequency ablation, thermo-ablation, cryoablation, microwave ablation, irreversible electroporation, partial nephrectomy.

- For CNS hemangioblastomas these would be radiotherapy, stereotactic radiosurgery and neurosurgery.

- For pNETs these would be endoscopic ultrasound-guided radiofrequency ablation, enucleation, radiotherapy, Whipple's procedure, partial pancreatectomy, and radical pancreatectomy, pancreatectomy with full splenectomy.

In the context of the company CEA, as described in the response to question A 13 part b), the treatment rationale we model is that there is usually a dominant or primary tumour that is driving treatment decisions. Specific surgical procedures are not used in the model, instead each surgical procedure is considered as one for the specific tumour type (i.e. RCC, CNS hemangioblastoma, or PNET) and the associated costs are sourced appropriately, but these are not for a specific surgical procedure (e.g. full nephrectomy, Whipple's procedure, etc.).

Systematic review

A 15. Priority. Table 109, Appendix D, states that any intervention might be included, but does not mention best supportive care in the Interventions criterion for the SLR. Also, the comparator evidence is from a natural history study, where natural history might not be regarded as an intervention. Table 109 also excludes case series, but this is effectively the design of the VHL Natural History study.

a) Please clarify whether the VHL Natural History study was retrieved as part of the SLR.

b) Please clarify whether the SLR was designed in such a way that all natural history (or non-intervention) studies could have been found (see also Question A2).

c) If the SLR was not designed in this way then please conduct another SLR to ensure that all studies in the population in the scope, interventional and non-interventional, of any treatment or no treatment,

BSC or natural history are found and fully analysed and reported in the clinical evidence.

MSD response:

a) The VHL Natural History Study was not retrieved as part of the SLR.

b) As described in the response to question A 2, the VHL Natural History Study is a (currently) unpublished study specifically commissioned by MSD to address the lack of available relevant comparator data in the published literature. The VHL Natural History Study, not (yet) being a published study, was not (and could not be) identified via a systematic literature review.

The inclusion criteria of the SLR of clinical evidence is shown in Appendix D Table 109 of the company submission and explicitly states that relevant prospective and retrospective cohort studies would be included (the search strategies used that explicitly include search terms for such studies are shown in Tables 106-108), therefore all relevant natural history (or non-intervention) studies could have been found if they are published. If the study is not published (as is the case for the VHL Natural History study) the study will not be found by this SLR.

c) Not applicable. As described in the response to part b) of the question above, the SLR of clinical effectiveness evidence was designed to include such studies.

A 16. Priority. Only the VHL Natural History study was included in the clinical effectiveness section. However, in the cost effectiveness model other sources of data for the comparator were used: retrospective analysis of the pre-treatment phase of MK-6482-004 and the Optum Clinformatics Data Mart claims study. Please provide a full description of the VHL Natural History study, the pre-treatment phase of MK-6482-004 and the Optum Clinformatics Data Mart claims study, including:

a) **Study design**

b) **Baseline characteristics, including proportion of patients in each of the tumour type combination subgroups (see question A10).**

c) **Treatment description, including type of surgery**

d) **All outcomes, including incidence of VHL-related surgeries, including by tumour type combination subgroups (see questions A10)**

e) **A comparison of these studies and their outcomes with reference to applicability to the UK and comparability to the treatment phase of MK-682-004**

MSD response:

a) to d) A full description of the VHL Natural History Study, including all information relevant to its applicability to the cost-effectiveness model, is already reported in detail in the company submission in section B.2.9 and Appendix Q. The relevant details of the analysis of the pre-treatment phase of MK-6482-004 (named Appendix R) is provided along with this response document. The publication of the Optum Clinformatics Data Mart claims study is reference #34 in the company submission and the full-text has been provided in the reference pack ["Jonasch 2022 (Clinical Genitourinary Cancer)"].

e) With regard to the applicability of the these studies to the UK, it should be noted that due to the rarity of this disease in the UK (the prevalence of VHL disease is between 1 in 77,340 and 1 in 68,493, with between 55 and 120 patients in England likely to be eligible for treatment with belzutifan, as described in section B.1.3 of the company submission), there is a lack of published data on the population characteristics of patients with VHL disease in the UK, much less data stratified specifically for the subgroups of patients with different tumour manifestations. The most up-to-date published information on the UK VHL disease patient population is that presented in the publication of the national audit of VHL disease in the UK by Maher et al. (https://www.nature.com/articles/s41416-022-01724-7), which does not report detailed/tabulated population characteristics. Therefore, it is not feasible to provide a meaningful quantitative comparison between the relevant patient population in UK clinical practice and the patient populations of these studies.

With regard to comparability to the treatment phase of the MK-682-004 study, the patient population of the retrospective analysis of the pre-treatment phase of the MK-

6482-004 is necessarily fully comparable due the fact that these would be the same patients.

The Optum Clinformatics Data Mart claims study in general was not designed for comparability with the MK-6482-004 study, the data from this study were used to inform adjustment factors in the cost-effectiveness model along with short-term and long-term surgical complication rates in the model; they were not used to inform the comparative effectiveness/transition probabilities (before real-world adjustment). Like many other sources (e.g., published literature) used in the model for non-efficacy inputs, population matching for all input sources is not required.

A 17. In Appendix D the company stated that the systematic review was conducted "to identify relevant studies that investigated Belzutifan and any relevant comparator treatments for the indication of interest for this appraisal as described in Table 1 of Document B section B1.1". However, in the findings of the systematic literature review section, the company states a total of 26 citations representing 26 unique studies were initially included in this review. However, "only one of the 26 studies identified investigated the efficacy of belzutifan, specifically the MK-6482-004 study as reported in the Jonasch et al. 2021 publication (1), and so only that one study has been included for the purposes of this submission."

a) Please explain and clarify why only one citation out of 26 was included. How and why did the other 25 citations (Table 110) meet the inclusion criteria for the systematic literature review, but were then excluded from the results and findings?

b) It appears that some of the studies listed in Table 110 may have provided relevant data on natural history or comparator interventions e.g., Chan et al. (2022) and Ploussard et al. (2007). Please confirm whether all 25 studies were checked for having relevance for providing relevant natural history/comparator data for the submission.

c) Table 111 of Appendix D lists "*…studies initially excluded after full-text screening*". As above, it is possible that some of these studies could have provided data on natural history or comparators to Belzutifan. E.g., it is not clear why Joly et al. (2011) was excluded based on population when the recruited patients had VHL RCC (also, survival data were mentioned in the abstract). Other studies may also have provided relevant outcome data for natural history/comparators to Belzutifan e.g., Arnon et al. (2021) and Iwamoto et al. (2011). Please confirm whether all the studies in this table were considered for availability of relevant natural history/comparator data.

MSD Response:

a) The SLR for clinical effectiveness evidence was performed primarily in order to identify all relevant published and unpublished randomised controlled trials (RCTs) and non-randomised clinical trials (non-RCTs) relating to belzutifan as per the final scope. Consequently, for this purpose (as stated in section B.2.2 of the company submission), only studies that reported relevant data on belzutifan are of interest and any studies that did not do so were excluded at the true final stage.

b) The SLR of clinical effectiveness evidence was originally designed to include single-arm studies that reported data on potential comparators. However, it soon became quite clear that not only would RCT-based network meta-analysis/indirect treatment comparison be unfeasible (as the one study in belzutifan in this indication is single-arm), but that any single-arm (or otherwise) studies reporting aggregate data on a potential comparator would also not be useful for a MAIC. This is due to the small size of the MK-6482-004 study, such that any analytical approach that involved matching/adjusting the population of the MK-6482-004 (which we have individual patient data for) to the population characteristics of a potentially relevant published comparator study (which would only be reported in aggregate form without individual patient data) would yield MAIC results of such great uncertainty they would not be of any value for/would be inappropriate for decision-making purposes.

The only form of MAIC that could feasibility yield results of sufficient certainty/value would be one where individual patient data from a large comparator study was available which would be large enough to match/adjust to the characteristics of the

MK-6482-004 study population, which was why the VHL Natural History Study was commissioned by MSD and used as part of this appraisal.

Consequently, the 25 comparator studies initially included the SLR of clinical effectiveness evidence would not have contained useful relevant information that could be used to meaningfully derive a comparison between belzutifan and comparators.

c) Studies were included based on the PICOS criteria provided in Table 109 of the company submission. Therefore, any natural history/comparator studies that did not report any outcomes of interest were not included. Joly et al. (2011) was excluded based on the population was not clearly defined as non-metastatic and no outcomes of interest were reported. Arnon et al. (2021) was excluded because the abstract did not provide any outcomes of interest and the included population was not clearly defined as non-metastatic. Iwamoto et al. (2011) was excluded as the study did not report any outcomes of interest.

A 18. Appendix H reports methods and results for an SLR of HRQoL studies. The review was initially conducted during July 2020 and updated July 2022. Please explain why the study selection criteria (Tables 147 and 148 of Appendix H) differed between the two dates and the implications for differential study identification. Please also see questions A5 and A8 that cover similar queries regarding the search strategy.

MSD response:

Please see response to A 5. for the explanation of the difference of inclusion criteria for the population in the two searches. The updated search in July 2022 included specific utility outcomes of interest in VHL disease and as required for cost-effectiveness analyses as stipulated by the NICE methods guide. Although this update may have excluded certain studies with outcome types that are not listed in the inclusion criteria, this is unlikely as the inclusion criteria includes a comprehensive list of outcomes, Furthermore, this produces for specific results relevant for the cost-effectiveness analyses in this appraisal.

A 19. Table 109 ("*Other*" inclusion criteria) indicates that only studies published in English language were included in the SLR.

a) Please provide the number of relevant studies omitted from the review because of being published in non-English languages.

b) Please consider the impact of exclusion of studies published in non-English languages on the estimates in the submission.

MSD response:

a) A total of two non-English studies were identified during full text review and were excluded.

b) Due to the very limited number of studies that were excluded for non-English language (i.e. two), we anticipate that the impact of this is minimal.

A 20. Please describe the process used for data extraction and risk of bias assessment of the included studies (for intervention and comparator data). Please state:

a) For each process, how many reviewers were involved and the methods for resolving disagreements.

b) Please provide the *a priori* plan for data extraction (i.e., what types of data were extracted?).

c) The Cochrane risk of bias tool is not suitable for non-comparative studies. Please provide a risk of bias assessment for all included intervention and comparator studies, using checklists suitable for the respective study designs.

MSD response:

a) Two reviewers working independently reviewed eligible studies for the final list of selected eligible studies. Any discrepancies observed between the two reviewers were resolved by involving a third reviewer and coming to a consensus.

b) The following data extraction variables were captured *a priori*:

The following information was extracted regarding study characteristics: study title, first author, study identifiers (e.g. cohort name, NCT number), study design, study duration (year of initiation/completion), phase, masking, number of patients/subjects

enrolled and number completed, study duration, initiation, and completion dates, follow-up duration, inclusion/exclusion criteria, outcomes reported, study quality items, data source, eligibility period for observational data, and date of data cut-off.

The following information was extracted regarding interventions: treatment name, treatment dose, method of administration, frequency of administration, planned treatment duration, observed treatment duration, and concomitant/background therapies.

The following information was extracted regarding baseline/study inception patient/subject characteristics: age, age of VHL diagnosis, age of VHL associated RCC diagnosis, any other VHL associated tumours, age at start of treatment, gender, race and ethnicity, region/country, method of VHL status diagnosis, performance status (ECOG, KPS), and VHL type (1, 2A, 2B, 2C).

The following information was extracted regarding outcomes: tumour response proportion (objective, complete, partial, stable disease, and progressive disease); including response criteria used (e.g. RECIST 1.1, iRECIST or mRECIST), duration of response, time to response, time to surgery, progression-free survival, overall survival, drug-related adverse events, grade 3-5 AEs (all, drug-related), discontinuation due to AEs, serious AEs, and deaths.

c) Risk of bias assessment has now also been conducted using the ROBINS-I risk of bias tool for the MK-6482-004 study included in the SLR of clinical effectiveness evidence (see Table 3 below). Why only the MK-6482-004 study was included in this SLR (and consequently assessed for risk of bias below) is described in the response to question A 17 previously.

Table 3 Risk of bias assessment via the ROBINS-I risk of bias tool for the MK-6482-004 study.

Study	MK-6482-004
Bias due to confounding	Low
Bias in selection of participants into the study	Low
Bias in classification of interventions	Low
Bias due to deviations from intended interventions	Low
Bias due to missing data	Low
Bias in measurement of outcomes	Low
Bias in selection of the reported result	Low
Overall risk of bias	Low

Clinical effectiveness evidence

A 21. Priority. Please clarify whether the MK-6482-004 trial population is representative for the UK patient population. Please compare trial and UK patient characteristics for all three subgroups.

MSD response:

As noted in the company submission in section B.2.3 in the "Baseline characteristics of trial participants" subsection, the baseline characteristics of this study have been presented to UK clinical experts with experience treating patients with VHL disease. Experts broadly agreed that these were representative of/applicable to the patients in the UK who would be treated with belzutifan in accordance with the marketing authorisation.

Due to the rarity of this disease in the UK (the prevalence of VHL disease is between 1 in 77,340 and 1 in 68,493, with between 55 and 120 patients in England likely to be eligible for treatment with belzutifan, as described in section B.1.3 of the company submission), there is a lack of published data on the population characteristics of patients with VHL disease in the UK, much less data stratified specifically for the subgroups of patients with different tumour manifestations. The most up-to-date published information on the UK VHL disease patient population is that presented in the publication of the national audit of VHL disease in the UK by Maher et al. (https://www.nature.com/articles/s41416-022-01724-7), which does not report detailed/tabulated population characteristics. Therefore, it is not feasible to provide such a comparison.

It should also be noted that the MK-6482-004 study focused on/included patients who had to have VHL disease-associated RCC, and so the patient population included in the study may not be fully representative of the general VHL disease population in the UK.

A 22. Priority. In Table 16, please add the number of patients in each subgroup used to get the estimated efficacy results. Please also clarify if the efficacy measures (e.g., ORR) are based on subgroup-specific tumours: for example, in the subgroup of patients with CNS hemangioblastoma the ORR = 44% refers to CNS Hb tumours only or refers to any tumour that a patient in the CNS subgroup might have (thus CNS Hb, RCC or pNET-associated tumours)?

MSD response:

The data on ORR, DCR, DOR, TTR, PFS, and TTS shown in the Table 16 summary table are specific to the tumours for which they are reported i.e. the results reported under the "RCC (all patients)" subheading are specific to RCC tumours, the results reported under the "Subgroup of patients with CNS hemangioblastoma" subheading are specific to CNS hemangioblastomas, and the results reported under the "Subgroup of patients with pNET" subheading are specific to pNETs. For example, in the subgroup of patients with CNS hemangioblastoma, the ORR = 44% refers to response outcomes observed in CNS hemangioblastomas only.

This is made explicit/clear when these results are reported in detail in section B.2.6 and B.2.7 of the company submission in the text and the table/figure headings where detailed results are presented.

A 23. The population in the decision problem is defined as: "VHL adult patients who require therapy for VHL associated RCC, CNS hemangioblastomas, pNET, and for whom localised procedures are unsuitable or undesirable." However, the MK-6482-004 trial did not specifically require participants to be considered unsuitable or undesirable for localised procedures. Please discuss the implications of this discrepancy.

MSD response:

As described previously in the responses to questions A 10 and A 12, we consider the MHRA indication wording to be both broader and narrower than the trial population. We consider it is broader as patients with a e.g. CNS hemangioblastoma

(with or without other tumours) but no RCC could be treated with belzutifan. Similarly, patients with pNETs (with or without other tumours) but no RCC tumour could be treatment with belzutifan.

We consider it narrower as it is only for patients who “require therapy” but also for whom localised therapies are “unsuitable or undesirable”. As such, these patients are more unwell, have a more severe presentation that requires an intervention than some in the study.

A 24. It appears that the MK-6482-004 study did not estimate overall survival. Please perform an analysis of overall survival, even if data are immature.

MSD response:

As described in section B.2.6 Table 13 of the company submission, only two patients had died by the time of the latest 01-APR-2022 data cut-off date of the MK-6482-004 study. Performing an analysis of overall survival based on only two deaths (one due to suicide and one due to acute fentanyl toxicity, as stated in the section B.2.10 of the company submission) in a study with only 61 participants would be highly inappropriate and clearly will not yield results that could be meaningfully or appropriately used for decision-making on the clinical effectiveness of belzutifan. Therefore, this analysis has not been performed.

A 25. The presentation of baseline and outcome data on subgroups in MK-6482-004 according to different combinations of tumour type is incomplete (as indicated in Section B.2.7, p83 of Document B).

a) Please provide all baseline and outcome data for the subgroup of patients with RCC and CNS haemangioblastomas and pNETs (n=17)

b) Please provide all baseline and outcome data for the subgroup of patients with RCC and CNS haemangioblastomas but not pNETs (n=33)

c) Please provide all baseline and outcome data for the subgroup of patients with RCC and pNETs but not CNS haemangioblastomas (n=5)

MSD response:

These data will not be provided as:

- Such data would be of extremely limited value to decision making due to the small sample sizes involved (as low as n=5 as described in the question).
- These highly specific subgroups do not reflect the GB marketing authorisation or NICE decision problem for this indication which are for a patient population that is not restricted to these subgroups or excludes these subgroups.
- There is no evidence or scientific rationale to expect that the treatment effect of belzutifan in any one tumour is affected by the presence of any other tumours in the same patient at the same time, so such subgroup data would not provide any additional information useful for the assessment of the clinical effectiveness of treatment with belzutifan.
- The cost-effectiveness analysis in the company submission is based on the treatment of effect of belzutifan on individual tumours, so such subgroup data will not provide any additional information useful for assessing the cost-effectiveness of treatment with belzutifan in this indication.

A 26. The term “primary tumour” has different definitions within the documentation. E.g., “…tumour with the greatest burden on the patient…” (p132) and “…highest risk tumour site…where tumours are likely to be most progressed” (p133). In addition, on p27 it is stated that “…..the primary tumour….is not necessarily the first tumour.” Overall, this amounts to an unclear definition of “primary tumour”.

a) Please provide a clear definition of the term “*primary tumour*”

b) In relation to subgroups including patients with more than one type of tumour, please explain which is the primary tumour for all patients.

MSD Response:

a) We would like to apologise if there has been confusion in the term “primary tumour”. In the CS we refer to the primary tumour as the VHL-RCC, VHL-CNS Hb or VHL-pNET with the greatest burden on the patient (as stated in the *Model structure* section of the CS) and greatest concern to the clinician. This “primary tumour” drives treatment decisions, and unlike other oncology therapy areas, the primary tumour is not necessarily the *first* tumour manifestation (as stated in the *Health condition* section of the CS). We have provided a couple of examples for clarity:

- A patient whose first tumour manifestation was a pNET which is under active surveillance but has since had multiple RCC surgeries and is now in need of a full nephrectomy. In this case, the primary tumour is RCC – the pNET is of a lesser concern and the nephrectomy would lead to end-stage renal disease requiring dialysis.

- A patient who has a large pNET and now requires a Whipple’s procedure. A CNS Hb has also been identified, but under active surveillance. In this case, the primary tumour is pNET which is of the greatest concern and requires immediate intervention.

b) In the context of the clinical effectiveness evidence from the MK-6482-004 study on the subgroups of patients with RCC who also had CNS hemangioblastomas and patients with RCC who also had pNETs, as presented in section B.2.7 of the company submission, results are presented in those sections explicitly for specific tumour types of relevance. In the “Central nervous system hemangioblastomas” subsection of section B.2.7, results for each outcome are explicitly reported to be “for CNS hemangioblastomas”, similarly in the “Pancreatic neuroendocrine tumours” subsection of B.2.7, results for each outcome are explicitly reported to be “for pancreatic neuroendocrine tumours”. For patients in the trial, we are unable to distinguish which is the primary tumour by our definition in patients who have more than one tumour type.

A 27. In terms of baseline data:

a) Please provide the number of each type of tumour per patient.

b) Please provide the number and type of VHL-associated tumours per patient

c) In Table 11 of Document B, the information on the number of patients with pancreatic lesions is discrepant with the published paper (Jonasch et al. NEJM 2021): n=32 versus n=61 patients respectively. Please clarify the correct number.

MSD response:

a) We are not in a position to provide detailed individual patient data (i.e. the "per patient" data) from the MK-6482-004 study. Data relevant to this appraisal and the decision problem with regard to number and types of tumours patients in the MK-6482-004 study had at baseline are already provided in the company submission in Appendix P, with the treatment effect of belzutifan on relevant tumours per patient already provided in Figure 4 of section B.2.6 of the company submission.

b) In patients with VHL disease, all tumours present in the patient are necessarily VHL-associated tumours, given that VHL disease is a genetic disease that affects every cell in the body, so these data would be the same as those requested as part of question A 27 part a), and we are not in a position to provide additional detailed individual patient data (i.e. the "per patient" data) from the MK-6482-004 study.

c) The number of patients with pancreatic lesions reported in the Jonasch et al. NEJM (2021) study of N=61 refers to the patients with pancreatic lesions as determined via investigator assessment, whereas the n=32 reported in Table 11 of Document B is the number of patients with pancreatic lesions according to independent review committee determination which is a smaller number.

A 28. Section B.2.4 (Table 12) and Section B.2.7 ("*Other tumours*") mentions participants recruited to the MK-6482-004 study having tumour types other than RCC, CNS haemangioblastomas and pNETs. These include: non pNET pancreatic lesions (number of participants not stated); retinal haemangioblastomas (17 participants, 12 evaluable); adrenal lesions (n=3 participants); endolymphatic sac tumours (n=1 participant); and epididymal cystadenomas (n=16 participants).

a) Please confirm the number of evaluable and non-evaluable participants for each of the above tumour types.

b) Please define "*evaluable*" and "*non-evaluable*" in the context of these tumour types.

c) Please explain how these other tumour types relate to "*primary*" and "*non-primary*" tumours that are mentioned elsewhere in the submission.

d) Please explain the impact on treatment effect of the distribution of these other tumours within the study population.

MSD response:

a) "Evaluable" and "non-evaluable" tumours were only relevant/a consideration for the retinal hemangioblastomas, and not relevant or a differentiation that was made for any of the other tumour types, in the MK-6482-004 study. Specifically for retinal hemangioblastomas, "not evaluable" refers to 1) Overall poor quality of images; or 2) Ancillary findings suggest a retinal hemangioblastomas may be present; however, tumour was not visible or cannot be assessed due to poor quality of images received.

b) For the retinal hemangioblastomas, which was the only type of tumour where evaluable/non-evaluable in this context was considered in the MK-6482-004 study, whether a tumour was evaluable or not was determined during independent review based on the quality of imaging/scan data they received in order to make their assessment of tumour response. If the independent review committee determined that the imaging data they received for a tumour was not of sufficient quality to evaluate tumour response, the tumour was determined to be non-evaluable.

c) We apologise for any confusion here. Primary tumours referred to in the company submission refers specifically to the VHL-associated RCC, CNS Hb or pNET with the greatest burden on the patient. Therefore, "non-primary tumours" refer to non-RCC tumours in the VHL-RCC cohort, non-CNS Hb tumours in the VHL-CNS Hb cohort and non-pNET tumours in the VHL-pNET cohort. For example, non-RCC tumours refers to CNS Hb, pNET, retinal Hb, adrenal lesions etc. "Other tumours" refers to any tumour outside of RCC, CNS Hb or pNET regardless of which cohort a patient

belongs to. Hence, all “other tumours” are “non-primary tumours”, but not all “non-primary tumours” are “other tumours”. In an VHL-RCC patient, a CNS Hb and an adrenal lesions are both considered non-primary tumours, but an adrenal lesion would also be considered “other tumours” in the context of the subgroup analysis.

d) There is no evidence to suggest that the simultaneous presence of any of these other tumours in a patient would affect the treatment effect of belzutifan on RCC, CNS hemangioblastomas, and/or pNETs in the same patient.

However, as data from the MK-6482-004 study suggest that treatment with belzutifan may have a beneficial effect on these tumours, where they are present in addition to RCC, CNS hemangioblastomas, and/or pNETs, then it may be likely that a patient with RCC, CNS hemangioblastomas, and/or pNETs plus one or more of these other tumours may experience a greater overall beneficial treatment effect due to the additional beneficial effect on these other tumours (i.e. the more tumours a patient has, the greater the total beneficial effect, even though the level of effect for each tumour will not be affected).

Indirect treatment comparison (ITC)

A 29. Priority. The company state that: “As a limitation, it was not feasible using the available Natural History Study data to identify whether patients in these subsets had CNS hemangioblastoma and pNET at the patient-level index date (i.e., it was only feasible to identify patients with a recorded history of CNS hemangioblastoma or pNET at some point prior to the patient-level index date).” Please clarify that whether a patient had pNET or CNS was not recorded for each patient.

MSD response:

As noted in the quoted text, whether a patient had pNET or CNS hemangioblastoma at the patient-level index date was not originally recorded in the source data from the United States National Cancer Institute’s Urologic Oncology Branch. For the purposes of the MAIC documented in section B.2.9 of the company submission this had to be inferred/deduced based on whether the patient had a recorded history of

CNS hemangioblastoma or pNET at some point prior to the patient-level index date in that source data.

A 30. Priority. Table 37 states that the outcome for the ITC was: *"Exponential rate parameter for the cause-specific hazards of pre-surgery → 1st surgery"*. However, it is not clear whether the data used to estimate this rate includes 2nd surgery (according to Figure 8, one patient had two CNS surgeries). Also, the economic model includes 2nd and 3rd surgeries.

a) **Please provide the data for both intervention and comparator used for the ITC, including whether 1st or 2nd or 3rd surgery.**

b) **If 2nd or 3rd surgeries included, then please redefine the outcome as rate of surgery (any number).**

MSD response:

a) The data on both the intervention and comparator used in the ITC are provided in Table 37 in section B.2.9 of the company submission (the raw source individual patient data cannot be shared). Only one surgery was considered, further details are provided in the response to question B 2 part a) later in this document.

b) Not applicable, only one surgery was considered.

A 31. Priority. The company stated that they performed a MAIC. However, the comparator (natural history) data appear to have been adjusted (see Tables 34, 35 and 36), which can only be done by access to the individual patient data of the comparator study. It also appears to be the case that the company has access to IPD for the intervention, as would be expected. According to NICE DSU TSDs 17 and 18, population adjustment (MAIC is one method) i.e., of data for one treatment (comparator or intervention) is necessary if there is no access to IPD for both comparator and intervention and that methods of adjustment using both IPD for both comparator and intervention are preferred.

a) **Please explain why population adjustment was chosen instead of IPD – based analyses, which also include propensity-score weighting. In doing so, refer to TSDs 17 and 18.**

b) **Please explain why a MAIC was chosen as the method of adjustment as opposed to other methods such as simulated treatment comparison (STC).**

c) **Please explain why the list of covariates adjusted for the VHL-CNS Hb and pNET cohorts did not include tumour size, as was the case for the VHL-RCC cohort and why the number of all surgery types was not included in all subgroups.**

d) **Please follow the recommendations of TSD 17 in conducting and deciding on the methodology of IPD-based adjustment analyses.**

   i. **Please provide all validity check information e.g. degree of overlap, as required by the QuEENS checklist.**

   ii. **Please consider the use of more than one methodology, depending on assessment of validity.**

   iii. **Please consider all subgroups as determined by the answer to questions A10d and all covariates as determined by prognostic or treatment effect.**

MSD response:

a) IPD from the VHL Natural History Study and IPD from MK-6482-004 were stored, managed, and analysed separately (at IQVIA and MSD, respectively). Statisticians at Merck, who did not have access to Natural History Study IPD, analysed MK-6482-004 IPD to estimate parameter inputs needed for the belzutifan model arm (e.g., transition probability parameters via parametric multistate modelling). The VHL Natural History Study and MAIC was performed by the research group at IQVIA, who had access to patient-level data for the Natural History Study cohort but only summary-level results from MK-6482-004. After obtaining MAIC-based propensity score weights for each target population of interest, IQVIA used the reweighted

sample to estimate an analogous set of parameter inputs for the SOC model arm. It was thus feasible to conduct parallel IPD analyses of each source individually, with close coordination between the two organizations to ensure consistency of statistical approaches and study variable definitions. However, adjustment methods requiring pooled IPD from both sources, such as traditional inverse probability of treatment weighting (IPTW) (i.e., propensity score weighting based on pooled IPD), could not be conducted.

MAIC is closely related to traditional IPTW in that, under both approaches, weights are derived from a propensity score equation that predicts whether a given patient originates from the intervention or comparator cohort as a function of observed baseline characteristics (2-4). Whether using MAIC or traditional IPTW, the objective of our adjustment approach would remain the same: to reweight the external control cohort from the Natural History Study to match the distribution of key baseline characteristics of the MK-6482-004 trial population, and the same set of matching variables would have been selected. The reweighting scheme was successful under the MAIC approach (as illustrated by the comparison of baseline characteristics after reweighting in Tables 34-36 of the CS); a traditional IPTW method would have yielded the same distribution of baseline characteristics in the reweighted sample as that displayed in Tables 34-36 of the CS.

b) MAIC was preferred over STC as the adjustment method for this economic evaluation based on several considerations:

- MAIC is more practical than STC when the number of comparators is small and the number of outcomes is large (4). To populate the present economic model, multiple parameter inputs were required for the SOC treatment strategy (the sole comparator of interest). Within each target population (VHL-RCC, -CNS Hb, and -pNET), the same set of MAIC-based weights could be reused for each population-adjusted parameter input being estimated for that population from the Natural History Study. With STC, each of these parameter inputs (including each specific transition probability, the incidence of non-primary tumour surgeries, etc.) would represent a distinct “outcome” requiring its own prediction model.

- MAIC is also more practical for analyses of time-to-event endpoints requiring non-linear prediction models (4). STC under non-linear prediction models are generally biased when average baseline covariates are plugged into the prediction models, as the mean of individual-level predicted outcomes do not equal the predicted outcome evaluated at mean covariate values (3). Meanwhile, even if individual-level baseline covariates were used to calculate the mean of individual-level predicted outcomes, the implementation of this approach would pose significant challenges for the economic model, particularly in light of probabilistic sensitivity analysis requirements (e.g., see discussion in section 3.2 of Remiro-Azócar 2021 (2)).

- As an adjustment method, MAIC was compatible with well-established, commonly used methods for transition probability estimation. The present model aligns with precedence from several other recent Markov cohort models submitted to NICE (e.g., TA830, TA837, and TA766) that calculated transition probabilities via parametric multistate modelling, an approach that is also similar to the standard parametric modelling methods that are widely used in partitioned survival models for advanced cancers.

c) The size of the largest CNS tumour at the patient-level index date was unavailable in the Natural History Study VHL-CNS Hb cohort, as was the size of the largest pancreatic tumour at the patient-level index date in the Natural History Study VHL-pNET cohort. More generally, the presence/absence of CNS Hb and pNET at the patient-level index date could not be identified within the Natural History Study data; it was only possible to determine whether a patient ever had a pre-index history of these tumour manifestations in the Natural History Study. As discussed in the CS, due to this limitation, surgery-related inputs for SOC in the VHL-CNS Hb and VHL-pNET populations were derived from MK-6482-004 pre-treatment period data, including: parameter estimates for the pre-surgery → surgery transition; and the incidence and percentage breakdown of non-primary tumour surgeries.

As noted in the CS, in the MAIC for each target population, the adjustment of prior surgeries for the primary tumour type led to a sizeable decrease in effective sample size. Given the limited sample sizes available even prior to matching, and the expectation that number of primary tumour surgeries was the most pertinent

covariate in each population, total number of VHL-related surgeries was not included as an additional matching variable.

The use of MK-6482-004 pre-treatment data for key SOC parameters in the VHL-CNS Hb and VHL-pNET populations helps mitigate concerns about potential confounding from unadjusted covariates in the MAIC for these cohorts. As described in response to B 7. , the Excel model has now also been updated to enable a scenario analysis using MK-6482-004 pre-treatment data for analogous SOC parameters in the VHL-RCC population.

d) As described in response to part (a) above, adjustment analyses requiring full IPD (pooled from both sources) could not be conducted for the present model.

A 32. Priority. Only the VHL Natural History study was used for the MAIC.

a) **Please perform all analyses requested in question A31 using the pre-treatment phase of MK-6482-004 and the Optum Clinformatics Data Mart claims study as a source of comparator data.**

b) **Please compare and contrast the results of these analyses.**

MSD response:

MAIC using the pre-treatment phase of MK-6482-004 and the Optum Clinformatics Data Mart claims study as a source of comparator data have not been performed.

As described in the response to question A16, the patient population of the retrospective analysis of the pre-treatment phase of the MK-6482-004 is necessarily composed of the same patients as the MK-6482-004 study' post-treatment-initiation phase and so no matching or adjustment is required.

The Optum Clinformatics Data Mart claims study was not designed for comparability with the MK-6482-004 study. There would be significant limitations with using the Optum database to construct such an external control arm for the purposes of the cost-effectiveness analyses (which would be the primary purpose/use of such an MAIC in the context of this appraisal). The limitations include the more limited availability of matching variables (e.g., baseline renal tumour size would not be available, and the ability to measure the number of prior surgeries would be

dependent on each patient's length of continuous enrolment in the database) compared to the VHL Natural History Study. The data from the Optum study were used to inform adjustment factors in the cost-effectiveness analysis along with short- and long-term surgical complications rates in the model. The data were not used to inform the comparative effectiveness/transition probabilities (before real-world adjustment). Like many other sources (e.g., published literature) used in the analysis for non-efficacy inputs, population matching for all input sources was not required.

Adverse events

A 33. In Appendix F the company states: "Two deaths due to AE occurred during the study. One was a suicide and the other was due to fentanyl toxicity that was reported 127 days after the participant started belzutifan and was reported as not related to study drug by the investigator".

e) Please explain how the company (investigator) can be confident that the toxicity was not caused by the study drug, and was as implied, a result of other sources. Especially considering that the company states adverse events leading to discontinuation of study treatment included "*toxicity to various agents*". Please state which agents specifically were relevant here.

f) Furthermore, if the toxicity was a result of a non-study drug, would this imply other participants could perhaps have been taking additional medication, and if so, would this contaminate overall findings?

MSD response:

a) As a result of poor health, patients with VHL disease may experience constraints in physical functioning (5), they tend to experience emotional distress associated with the wide number of possible manifestations (5, 6), variable age of onset (from early childhood into adulthood) (5, 6), and the 'watch and wait' approach of active surveillance (periodic imaging check for growth of the existing tumours or occurrence of new tumours) (7). These are significant risk factors for substance abuse and/or committing suicide.

According to FDA IND safety reporting guidance (page 3-4) [https://www.fda.gov/media/150356/download], to assess causality for an adverse event, the principle is "there is a reasonable possibility that the drug caused the adverse event". The following examples provided in the IND safety reporting regulation (312.32(c)(1)(i)) illustrate the meaning of reasonable possibility with respect to a determination that there may be a causal relationship between the drug and the adverse event:

- A single occurrence of an event that is uncommon and known to be strongly associated with drug exposure (e.g., angioedema, hepatic injury, Stevens-Johnson Syndrome).

- One or more occurrences of an event that is not commonly associated with drug exposure but is otherwise uncommon in the population exposed to the drug (e.g., tendon rupture).

- An aggregate analysis of specific events observed in a clinical trial, indicating that they occur more frequently in the drug treatment group than in a concurrent or historical control group. Such events may be known consequences of the underlying disease or condition or events that commonly occur in the study population independent of drug therapy. Such events could also be related to an intervention or therapy that is standard of care for the disease (e.g., background treatment).

FDA considers the application of the reasonable possibility causality standard to be consistent with the discussion about causality in the International Council for Harmonisation (ICH) E2A guideline for industry (the ICH E2A guidance). [ICH guidance for industry E2A Clinical Safety Data Management: Definitions and Standards for Expedited Reporting (March 1995), pages 6–7] [https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-15.pdf]

In summary, the underlying VHL-disease is a significant confounding factor for the concerned events. There is no evidence to suggest a possible causal association

between the event and belzutifan, which is in line with the guidelines of FDA and ICH on causality assessment.

b) While participants in the MK-6482-004 study may have received concomitant medications (as described in section B.2.3 of the company submission) during the study period (as is the case for nearly all clinical trials), these are very unlikely to have contaminated the overall findings of the study as none of the concomitant medications taken would be ones that could have an antitumour effect (as described in section B.2.3 of the company submission).

Section B: Clarification on cost-effectiveness data

Economic analysis

B 1. Priority. To reflect the population in the decision problem in terms of each tumour type combination subgroup, as requested in question A10d, please repeat the cost-effectiveness analyses using for each subgroup of patients, subgroup-specific model parameters (e.g., baseline characteristics and rates of surgery for each type of tumour). This should include subgroup specific QALY weightings based on disease severity.

MSD response:

As described in response to A 10. d), distinguishing by different tumour type combinations was not feasible; therefore, we cannot conduct cost-effectiveness analyses in this manner.

Conceptual model

B 2. Priority. Please answer the following questions about the (conceptual) model structure:

a) Please provide more description about how the health states in the model were defined (e.g., if these are based on primary tumour or not) and align this with the health states in the electronic model (there is a

mismatch between the model described in the CS and the model implementation in terms of number of health states – model engine sheets).

b) **Please clarify how disease progression is defined, how it is included in the model, what is the relationship with metastatic disease and if it is one of the reasons for having surgery.**

c) **On page 132 of the CS, it is mentioned that "Each primary tumour site i.e. the VHL-RCC, VHL-CNS Hb, or VHL-pNET tumour with the greatest burden on the patient is modelled as a separate cohort using the same model structure". This suggests that in the model, subgroups are based on primary tumour, whereas it is unclear whether primary tumours have been used to define subgroups in the MK-6482-004 trial as questioned in clarification question A10. If there is a mismatch between the subgroups in the trial and the subgroups in the model, the trial results might not be applicable to the model and vice versa. Please explain the implications on the CE analysis related to this potential mismatch.**

d) **Also, on page 132 of the CS, it is mentioned that "Although the incidence of non-primary tumours, and therefore related surgeries, is captured in the model, the additional burden on costs and quality of life of having multiple tumour manifestations simultaneously is not specifically captured". Please clarify how the incidence of non-primary tumours is captured in the model and why the additional costs and QoL are not. Please explain what would be needed to capture these additional costs and QoL and why it is not possible to include them in the model.**

e) **Please clarify the following sentences on page 133 of the CS: "The cost and health implications of surgeries for non-primary tumours as well as their associated complications were reflected as per-event costs and QALY decrements applied on incidence of each non-primary tumour surgery. This approach to modelling primary and non-primary tumours differently is used to reflect typical disease progression in VHL:**

patients may require surgeries at multiple sites, but clinicians focus on the highest risk tumour site, as this is where tumours are likely to be the most progressed. Given the incidence of non-primary surgeries is not reflected by explicit state transitions within the model, the cumulative health impact of having undergone multiple non-primary surgeries is therefore not captured". Please explain how what's described in the first two sentences is included in the model. Regarding the third sentence, please clarify what is explicitly meant here since as mentioned in the previous question, the company indicated that the incidence of non-primary tumours is captured in the model.

f) **On page 134 of the CS, it is mentioned that "the pre-surgery health state describes patients who have not had surgery since belzutifan trial initiation, and for the purposes of the economic analysis, the treatment decision point. The majority of participants in the MK-6482-004 trial had multiple surgeries prior to trial initiation (mean 5.5)". Please discuss the implications, if any, of patients having multiple surgeries prior to trial initiation.**

g) **On page 135 of the CS, it is mentioned that "To mitigate this limitation while accounting for practical data constraints, functionality was incorporated into the Markov structure to track the occurrence of certain important clinical events. Specifically, surgery and event-free after surgery health states were used to capture perioperative mortality and surgical complications to which costs and disutilities could be applied accordingly. Separate health states were also defined to track patients' treatment discontinuation status over time (i.e., on-treatment vs. off-treatment)". Please clarify how this was included in the model.**

h) **Given the lack of data and the large number of assumptions needed to populate the model, please justify why a simpler model structure (e.g., partition survival model) was not considered.**

MSD response:

a) The model health states were defined to align with important outcomes in the disease progression of VHL, namely surgery and metastatic disease. These health states correspond to surgery and metastases of the primary tumour type only. Incorporation of non-primary tumour surgeries are not captured through model transitions.

To further clarify the initial model structure its subsequent development, we would like to explain the marketing authorisation process. The model was initially developed when the expectation for the marketing authorisation was for VHL-associated RCC only (in line with the population recruited into MK-6482-004 trial) and without restriction to those patients for whom localised procedures are unsuitable or undesirable, as stipulated by the final MHRA marketing authorisation. The FDA approved belzutifan in August 2021 for not only VHL-associated RCC, but also VHL-associated CNS Hb and pNET (with no restriction around localised procedures). The model structure was then adapted to account for the two additional VHL patient populations. The MHRA approved belzutifan in May 2022 and similarly expanded the indication to include CNS Hb and pNET; however, added a further restriction “for whom localised procedures are unsuitable or undesirable”. When the model was adapted to include VHL-associated CNS Hb and pNET cohorts, it initially included first, second or third surgery states and corresponding event-free after surgery states with the expectation that the MHRA label would match the FDA label. However, as the MHRA label restricts belzutifan use to those for whom surgery is unsuitable or undesirable, the model was adapted to permit for one surgery only as a ‘last resort’ intervention. Transition probabilities to subsequent surgeries and corresponding event-free after surgery states were set to 0 to reflect the intent of the MHRA indication wording; 2nd and 3rd surgery states and their corresponding event-free after surgery states are not used. Hence, there are more health states in the electronic model than described in the CS.

b) Surgery and metastases are modelled independently based on the incidence of these events in the MK-6482-004 trial and RW data sources. Surgery can occur at the discretion of the patient and clinicians, including reasons beyond metastases as described in the *Burden of VHL disease* in section B.1.3 of the CS. Disease

progression is part of the overall response primary outcome in the MK-6482-004 trial (defined based on RECIST 1.1) and informs HRQoL in the economic.

c) In an attempt to model the MHRA indication, a simplifying assumption was made: patients in the MK-6482-004 trial with a pNET represent a cohort of patients with a pNET that is the 'primary tumour' as described in response to A 10. (i.e., the tumour that is driving treatment decisions). This is not a perfect representation of the populations as we cannot determine in patients with more than one tumour manifestation which is the 'primary tumour' within the available datasets. Without this assumption it would not be possible to model all populations included in the MHRA indication. As described in response to A 10. , a patient with an RCC and pNET tumour would be modelled in both the RCC and pNET cohorts, allowing that either tumour manifestation could be the primary tumour (and therefore the alternative manifestation is a non-primary tumour).

In the MK-6482-004 trial, 100% of patients have VHL-related RCC at baseline visit; other VHL-related tumour manifestations were common among the included patients but were not required for study inclusion. This is consistent with the modelled VHL-RCC cohort, in which 100% of patients have RCC at model entry and many patients have other VHL-related tumour manifestations as well).

For the VHL-CNS Hb and VHL-pNET model cohorts, clinical inputs from MK-6482-004 were based on subgroups of the trial population that had CNS Hb and pNET as of the baseline visit, respectively. However, because trial participants were required to have RCC, 100% of patients in the CNS Hb subgroup of MK-6482-004 have both CNS Hb and RCC, and 100% of patients in the pNET subgroup of MK-6482-004 have both pNET and RCC. In contrast, the VHL-CNS Hb and VHL-pNET *model* cohorts were defined to include patients with or without concurrent RCC. Due to this discrepancy, the observed incidence rates of non-primary tumour surgeries in these subgroups of MK-6482-004 were adjusted downwardly using the proportions of patients with VHL-associated CNS Hb and pNET who do not have RCC (as reported from the cross-sectional VHL RW QoL Disease Burden Study [2022]). This adjustment is described in the *Surgery* sheet of the Excel model. Additionally, in both arms, the percentage breakdowns of non-primary tumour surgeries in the CNS Hb and pNET cohorts were derived using pre-treatment period data from MK-6482-004,

and were similarly adjusted to account for the proportions of patients with VHL-associated CNS Hb and pNET who do *not* have RCC (i.e., the proportion of non-primary tumour surgeries attributable to RCC was decreased and the proportions attributable to other VHL-related tumours were proportionally increased).

Although this difference between the CNS Hb and pNET subgroups of MK-6482-004 vs. the VHL-CNS Hb and VHL-pNET model cohorts is a known limitation, the impact on the model results for these cohorts is likely minor, given the following considerations:

1. This discrepancy would mainly impact the raw incidences and percentage breakdowns of non-primary tumour surgeries in the CNS Hb and pNET subgroups of MK-6482-004, and (as noted above) these inputs have been appropriately adjusted.

2. While the true prevalence of RCC in the VHL-CNS Hb and VHL-pNET model cohorts is <100%, data from the VHL RW QoL Disease Burden Study showed that the prevalence of RCC is very high in these cohorts: 63.2% of patients with VHL-CNS Hb and 71.3% of those with VHL-pNET in this survey study had concurrent RCC.

3. The same limitation applied to both belzutifan and SOC in the VHL-CNS Hb and VHL-pNET model cohorts since both the MK-6482-004 trial (pre- and post-treatment periods) and the VHL Natural History Study were restricted to patients with VHL-related RCC at the baseline visit. Consequently, this limitation should not differentially introduce bias for one arm vs. the other.

Given the limited nature of the data package and the highly complex and varied nature of the disease, there are obvious limitations with how subgroups have been modelled. We have data for time-to-surgery, specifically the primary tumour surgery of interest for the cohort; however, we cannot categorically determine, for patients with more than one tumour type in the trial, which tumour was associated with the greatest burden. We request the EAG suggest a more appropriate approach given the limitations in the dataset and the evolving population (MK-6482-004 trial vs MHRA label).

d) The incidence, costs and QALY decrements of all non-primary tumour surgeries *are* captured in the economic model. The costs and QALY decrements associated with these surgeries are calculated in each weekly cycle based on the weekly incidence of non-primary surgeries and the proportion of patients still alive in each cycle. This is then layered (additively) onto the costs and QALYs estimated based on patient distribution across primary tumour-related health states (see response to (e) below for further details of this calculation). Although the costs and QALY impact of non-primary tumour surgeries are captured individually in an additive manner, the combined impact of multi-systemic disease (which may be greater than the additive effect of individual surgeries), including metastases and mortality impact linked to non-primary tumours, is not explicitly captured. This is where the value of belzutifan of reducing symptoms across multiple systems is underestimated in the economic model. Given the heterogenous nature of VHL and the diverse outcomes across its various manifestations, it would be of interest to consider how a patient-level simulation model might better capture the impact of multi-systemic tumour manifestations and the value of belzutifan at reducing this burden.

e) We would like to apologise for any confusion caused here. The weekly incidence rate of non-primary tumour surgeries was calculated as events per person-week (reported as events/person-year for intuitive numerical interpretation in the CS). For the belzutifan arm, this was calculated directly from the MK-6482-004 trial data for the VH-RCC cohort and adjustments to RCC surgery incidence based on the VHL RW QoL Disease Burden Study for the CNS Hb and pNET cohorts. For the SOC arm, this was calculated directly from the VHL Natural History Study for the RCC cohort and adjusted for the CNS Hb and pNET cohorts using the same methods as the belzutifan arm. The calculated weekly incidence rate was then multiplied by the percentage of people still alive in each weekly cycle to estimate the number of non-primary tumour surgeries occurring in that cycle. Risks of complications from non-primary tumour surgeries were accordingly calculated in each weekly cycle. Costs of surgeries as well as costs and QALY decrements of non-primary tumour surgery complications were then calculated in each cycle and were layered (additively) onto the costs and QALYs estimated based on patients' distribution across primary tumour-related Markov health states. What is meant by the third sentence ("Given the incidence of non-primary surgeries is not reflected by explicit state transitions

within the model, the cumulative health impact of having undergone multiple non-primary surgeries is therefore not captured") is that the impact of multiple tumour manifestations and multiple surgeries is "more than the sum of its parts". Therefore, considering the impact of non-primary tumour surgeries only additively does not capture the full impact of multi-systemic disease.

f) The majority of patients in the MK-6482-004 trial had undergone multiple surgeries prior to trial initiation (mean for the overall population was 5.5 previous surgeries). This aligns with what is expected for patients in the real world and is furthermore aligned with the MHRA label for belzutifan. Surgery is an all-too-common intervention for VHL patients. Prior to initiating treatment with belzutifan, given the eligibility criterion specifying they are no longer suitable for localised procedures, patients are likely to have had multiple prior surgeries. The only impact of this characteristic of patients recruited in the MK-6482-004 trial is that it ensures alignment to the likely patient population who will receive belzutifan in practice.

g) The implementation of treatment discontinuation status in the economic model is described in the *Treatment effect waning following belzutifan discontinuation* section in B.3.3. In the first instance, treatment discontinuation status is used to estimate belzutifan drug costs. When treatment effect waning is assumed, treatment discontinuation status is used to further define "off-treatment" versions of the model's health state for all health states except metastatic disease (given patients receive subsequent therapies treating metastatic disease) and death. State membership in an "off-treatment" state is based on the discontinuation rate in a given cycle based on the modelled ToT for belzutifan. In these off-treatment health states, the clinical efficacy parameters of patients in the belzutifan arm were assumed to gradually converge over time towards those of SOC. These can be seen in the Markov Trace sheets (*Trace_TxReg1_XXX*) in row 13.

h) A partitioned survival model (PartSA) is a suitable and simple model structure for cancer. VHL disease is not cancer, and the fundamental value of a partitioned survival model would not capture the key clinical outcomes for VHL patients: maintaining organ function, preventing advanced disease, and minimising burdensome CNS Hb symptoms. Furthermore, within-trial OS curves were very immature (close to horizontal) in MK-6482-004, which limited the suitability of a

PartSA. A Markov cohort structure is more suitable for decision-making given the chronic, non-metastatic nature of VHL disease and the inability to directly model OS with the available trial data.

Unfortunately, the situation we face is highly typical of rare, genetic diseases: very limited data availability, highly complex disease and highly heterogenous presentation. A PartSA model does not allow for straightforward adjustment of outcomes modelled after an intermediate event. Given the range of outcomes faced by VHL patients at the decision point in the current analysis, a simpler model such as PartSA would not be able to account nor adjust for the complexity of the disease nor the value that belzutifan offers to patients with VHL.

B 3. Priority. On page 153 of the CS it is stated that the TP from pre-surgery → metastatic disease for the VHL-CNS Hb cohort (i.e. metastases for non-primary tumours) in the belzutifan arm was estimated by assuming the percentage reduction (belzutifan vs SOC) in the hazard rate of pre-surgery → metastatic disease to be equal to the percentage reduction in the hazard rate of pre-surgery → surgery (belzutifan vs SOC). For the VHL-RCC and VHL-pNET cohorts respective TP for the metastatic disease health state were estimated using the HR of pre-surgery → surgery (for belzutifan vs VHL Natural History Study) to the hazard of developing metastatic disease estimated for SOC. Please confirm this is correct and explain the reasoning behind this discrepancy between cohorts on the estimation of TPs from pre-surgery → metastatic disease.

MSD response:

As clarification, the approach for estimating the pre-surgery → metastatic disease rate in the belzutifan arm was consistent across all three cohorts. In the CS excerpts above, the description of the approach is worded differently for the CNS Hb cohort as compared with the RCC and pNET cohorts, but the underlying meaning is the same: In each cohort, the relative treatment effect of belzutifan vs. SOC on pre-surgery → metastases was assumed equal to the relative treatment effect of belzutifan vs. SOC on pre-surgery → surgery in that cohort. (We use the terms “HR”

or “percentage reduction in hazards” interchangeably to refer to the relative treatment effect, e.g., a hazard ratio of 0.2 would equate to an 80% reduction in hazards.) In each cohort, the hazard rate of pre-surgery → metastases under SOC is proportionally reduced according to the relative treatment effect of belzutifan vs. SOC on pre-surgery → surgery, thereby estimating the hazard rate of pre-surgery → metastases under belzutifan.

Comparator

B 4. Priority. The CS on page 141 for SoC states: “For VHL-RCC and VHL-pNET cohorts, immediate surgery is assumed for 90% of patients. For VHL-CNS Hb, immediate surgery is assumed for 50% of patients; however, the outcomes associated with surgery is assumed for 100% of the cohort due to tumour burden creating neurological disability for the remaining 50% not operated on, which is assumed to have similar impact on HRQoL as the serious complications from CNS surgery.”

a) **Please provide a clear and detailed description of the comparators in each subgroup that should be used in the model with reference to the answers to question A13.**

b) **Please justify the need for immediate surgery in the context of the answers to questions A12 and A13.**

c) **Please clarify whether there is active surveillance for CNS Hb or not, and why it is claimed that the risk of metastatic disease and/or other symptoms of tumour burden is particularly increased in CNS Hb tumours.**

d) **Please provide objective evidence for the percentages mentioned above. If any evidence is lacking, then please provide clinical expert opinion including a report of elicitation methods.**

e) **The patients who receive SoC are described as those where *“immediate surgery is necessary”* (p. 157). If the surgery results in *“loss of organ***

***function and/or problematic sequalae*" (p. 157), but must be given immediately, then it must be the case that the patient would suffer some harm otherwise. If that is the case and the SoC population is meant to be the same as those who receive belzutifan, then please explain why no one in the belzutifan arm receives immediate surgery and the rate of first surgery is informed solely based on trial evidence. Does this mean that patients in the belzutifan arm are effectively left to suffer the harm entailed by not receiving immediate surgery?**

f) Please conduct the CEA as per question B1, treating the belzutifan and SoC arms as identical in terms of need for immediate surgery or justify why the need is different.

g) Please provide scenario analyses based on the objective evidence or expert opinion, assuming different percentages between 0% to 100%, and including one where the TP for surgery is determined wholly by TTS, as opposed to where a percentage is assumed.

MSD response:

a) The comparator arm in the economic model is not a specifically defined treatment or surgery as stated in the response to A 13. above. Hence, we refer to the comparator arm as standard of care (SOC). This SOC comprises of immediate surgery and its associated outcomes, and active surveillance (in 10% of the VHL-RCC and VHL-pNET cohorts). Further description of how surgical procedures are incorporated in the model are included in response to A 13. above.

b) In the context of the MHRA label and as stated in response to A 12. c) above, patients are at a 'fork in the road' where they have run out of alternative treatment options yet still "require therapy". Therefore, they have a requirement for immediate surgery to treat their primary tumour of significant burden in the absence of belzutifan as a treatment option.

c) We believe this question refers to the following statements made in the *Treatment decision point* section of the CS:

“In routine clinical practice, the decision point for a patient meeting the criteria of belzutifan eligibility would have three options: 1) surgery that is unsuitable or undesirable because it results in loss of organ function, 2) active surveillance to monitor a tumour that is above 3cm (RCC) or 2cm (pNETs) and therefore there is an increased risk of metastatic disease and/or other symptoms of tumour burden (particularly in CNS Hb tumours), or 3) belzutifan.”

We would like to clarify that this statement was looking at VHL disease as a whole before focusing in on the specific decision points for individual primary tumour manifestations. To clarify, patients in the CNS Hb cohort can have active surveillance but not without experiencing significant sequelae associated with tumour burden which would otherwise be alleviated through localised procedures. CNS Hb tumours do not metastasise; the point around increased risk of metastatic disease referred specifically to the VHL-RCC and VHL-pNET cohorts. The point around increased risk of other symptoms of tumour burden was more specifically focussed on the VHL-CNS Hb cohort. However, the population stipulated by the MHRA label “for whom localised procedures are unsuitable or undesirable” are patients experiencing either debilitating sequelae as a result of surgery or debilitating sequelae as a result of not undergoing needed surgery. We request the EAG consult an expert specialising in the treatment of VHL patients with CNS Hb manifestations to understand the situation for patients with tumours that cannot be resected but which still cause intrusive and burdensome symptoms.

As stated in *The place of belzutifan* section in B.1.3, patients who can have surgery with minimal complications should have surgery and are therefore not eligible to receive belzutifan. To this effect, patients with CNS Hb who have peripherally located tumours which can be operated on are by definition not eligible for belzutifan. In the CNS Hb cohort, belzutifan is reserved for patients “who require therapy” for tumours that are causing debilitating symptoms and “for whom localised procedures are unsuitable or undesirable” i.e. the tumour is in a location which cannot be operated on.

d) Formal elicitation methods were not used to estimate the proportions of patients requiring immediate surgery. This parameter can be tested in the economic model,

as was done in deterministic sensitivity analysis. The results are presented below in Table 4 (also in Appendix J1.4 of the CS).

Table 4 Tabular Deterministic sensitivity analysis of the proportion to receive immediate surgery

	ICER vs. comparator: Belzutifan vs. SOC *(£/QALY)*	
	Low input value	*High input value*
VHL-RCC cohort base case (90% receive immediate surgery)	**73,095**	
Proportion receiving immediate RCC surgery under SOC in VHL-RCC cohort: 80-100%	79,006	67,768
VHL-CNS Hb cohort base case (100% receive immediate surgery or sequelae)	**56,933**	
Proportion receiving immediate CNS Hb surgery or sequelae under SOC in VHL-CNS Hb cohort: 90-100%	60,013	56,933
VHL-pNET cohort base case (90% receive immediate surgery)	**77,649**	
Proportion receiving immediate pNET surgery under SOC in VHL-pNET cohort: 80-100%	83,857	71,897

e) This question identifies the challenge in implementing the MHRA indication wording within the economic model. The need for immediate surgery reflects the "*who require therapy*" indication wording and the loss of organ function and/or problematic sequelae reflects the "for whom localised procedures are unsuitable or undesirable" wording. The MHRA granted this marketing authorisation (MA) in this indication based on the results of the MK-6482-004 trial, the same trial used to inform the belzutifan arm of the economic model. That is to say, the MA was granted based on which patients would benefit from treatment with belzutifan, based on evidence reported in this very study. Misalignments between MA wording and supporting clinical trial patient population characteristics are not unusual or rare for highly specialised indications such as this one. However, this misalignment poses challenges for economic modelling and the data availability challenges from this single-arm phase II trial of belzutifan make this even more challenging.

It is logical that a patient would suffer some harm if needed surgery were not provided immediately. The harm in this case would be risk of metastatic disease due

to tumour growth (for RCC and pNET) or symptomatic burden (in all cohorts but particularly in CNS Hb). Belzutifan works by shrinking tumours and therefore reducing the risk of these two types of “harm”. This benefit is reflected in the economic model through the transitions within the health states as informed by the trial evidence for belzutifan. No belzutifan-treated patients receive immediate surgery given they have recourse to an effective therapy which provides an alternative to surgery. Therefore, no belzutifan-treated patient is assumed to be left to suffer the harm of not receiving surgery as they are receiving a treatment shown to reduce these two types of harms.

f) It would be inappropriate to consider the need for immediate surgery equal between arms. Both arms “require therapy” as the MHRA label stipulates; for the SOC arm this would equate to immediate surgery as they have progressed beyond the point of which active surveillance is a manageable option. The alternative to immediate surgery is therefore belzutifan; the aim of both surgery and belzutifan is to prevent progression to metastatic disease and/or relieve symptomatic burden. Surgery does this via resection of the tumour; belzutifan does this via shrinking tumour size.

g) As described above, patients ‘who require therapy’ are not eligible for active surveillance. In the RCC and pNET cohorts, however, we have made an allowance for 10% of patients to not have immediate surgery, instead receiving active surveillance to allow some flexibility in this assumption and align the treatment pathway with what could be expected in real-world practice. In the CNS Hb cohort, 50% do not have immediate surgery but to have outcomes associated with surgery which is effectively active surveillance which does nothing to address uncontrolled sequelae. It would not be clinically plausible to test scenarios as low as 0% of patients receiving immediate surgery as they would therefore not meet the eligibility criteria of the target population in “requiring therapy” in the absence of belzutifan. Deterministic sensitivity analyses have tested a range of 80-100% of patients receiving immediate surgery in RCC and pNET cohorts (i.e. allowing up to 20% of patients receiving active surveillance) and 90-100% of patients receiving immediate surgery or its outcomes in CNS Hb cohorts (i.e. allowing up to 10% of patients

receiving active surveillance without sequelae). The results of this sensitivity analysis are presented in response to B 4. d) above.

B 5. Priority. On page 142 of the CS it is stated that "The VHL Natural History Study collected data from US-based centres of excellence and patients in the study may therefore have received a different SOC compared to standard UK clinical practice. It is expected that patients treated at these sites had better access to surgery, and as a result, higher rates of surgery and therefore lower rates of metastasis were observed than would be expected in UK clinical practice." Please justify the basis of this expectation and explain if there has been any action taken or evidence collected to validate this assumption. The section "Aligning risk of surgery and metastatic disease to real-world SOC" of the CS reiterates this point, but again no evidence basis is provided to support this expectation/consideration.

a) **Please present the results of scenario analyses where these rates are varied within a range of realistic values, providing justification for the selected ranges.**

b) **In reference to question A12, if the MK-6482-004 trial population is less severe than the UK target population, please clarify why the risks observed in the trial were not aligned to represent real-world risks as it was done for SoC, and please perform the same kind of alignment for the belzutifan arm as the SoC arm.**

MSD response:

Based on evidence from a real-world study conducted on a dataset from the Optum Clinformatics Data Mart, a clear difference in surgery rates for VHL-related tumours is observed compared to the rates estimated based on data from the VHL Natural History Study. In discussions with clinicians, the cause of this difference was posited as being due to less proactive surveillance. The effect of this would naturally be in less disease control and therefore higher rates of metastatic disease. The Optum study was therefore used to adjust surgery and metastases rates from those calculated in *both* the MK-6482-004 trial and the VHL Natural History Study. As

described in *Aligning risk of surgery and metastatic disease to real-world SOC* of the CS, the first step of the adjustment was to use the Optum study to adjust transition probabilities calculated based on the VHL Natural History Study downward to match Optum for the estimation of these clinical parameters in the SOC arm. Then for the belzutifan arm, the cause-specific hazard rates of these transitions were adjusted by applying the original *ratios* (belzutifan vs SOC) of the exponential rates of these transitions in order to maintain the relative treatment effect of belzutifan.

a) Due to unavoidable data limitations, there is no particular range that can be considered realistic or plausible with which to present scenario analyses. The alternative scenario which can be presented is where these adjustments to account for real-world SOC are removed. The ICERs are presented below (also presented in Appendix J1.4 of the CS).

Table 5 Scenario analysis results removing alignment of surgery and metastases risk to real-world SOC

	ICER vs. comparator: Belzutifan vs. SOC *(£/QALY)*
VHL-RCC cohort base case	**73,095**
Do not adjust surgery and metastases rates to account for real-world standard of care	75,814
VHL-CNS Hb cohort base case	**56,933**
Do not adjust surgery and metastases rates to account for real-world standard of care	49,901
VHL-pNET cohort base case	**77,649**
Do not adjust surgery and metastases rates to account for real-world standard of care	55,768

b) We apologise for any confusion here. The risks observed in the trial were also aligned with real-world SOC in the base case, i.e. the belzutifan arm was also adjusted to account for real-world SOC consistent with the adjustment performed for the SOC arm. The method of alignment is described in the *Aligning risk of surgery and metastatic disease to real-world SOC* of the CS and further clarified in the first part of the response to this question.

Transition probabilities

B 6. Priority. On page 147 of the CS, it is mentioned that "Parametric models were fitted to time-to-event data to estimate the cause-specific

hazards of each transition starting from the pre-surgery state (i.e., pre-surgery → surgery, pre-surgery → metastatic disease, and pre-surgery → death) and event-free after surgery state (i.e., event-free after surgery → metastatic disease, and event-free after surgery → death) over time within the belzutifan and SOC arms".

a) **Please provide the detailed survival analyses for all TPs that were estimated based on different parametric models as per the NICE DSU technical support document.**

b) **Please include all parametric models (for all transition probabilities) in the model. Based on Table 46 and 47, it seems that only the Exponential distribution was considered. If this is the case, please explain why it was restricted to this distribution only.**

c) **Please explain how all rates in Tables 46 and 47 were derived. In particular, please clarify and justify why after surgery there is a benefit associated to belzutifan: it would seem reasonable to assume that if patients in the belzutifan arm received surgery is because the treatment is not optimally working in those patients. Therefore, assuming a benefit at all might be questionable and assuming that it is equal to the one observed pre-surgery might be an overestimation.**

d) **Please present the results of a scenario analysis where the transition probabilities after surgery are equal in both arms.**

MSD response:

a) Seven candidate distributions were considered to model the cause-specific hazards of pre-surgery → surgery in each target population. Consistent with methodological guidance from the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14, base-case parametric functions were selected such that the same functional form was used to model this transition as estimated from the MK-6482-004 trial and the VHL Natural History Study. Base-case parametric distributions for pre-surgery → surgery were chosen based on the following criteria:

- Visual assessment of fit vs. observed time to surgery, metastases, or death: Predictions generated by different parametric functions were visually verified against the observed data in each model arm. Specifically, modelled curves for time to surgery, metastases, or death (a composite endpoint determined by all three transitions from the pre-surgery state) were compared with observed Kaplan-Meier (KM) curves based on original data sources. KM curves were obtained from the MK-6482-004 trial (data cutoff: 01 April 2022 for belzutifan in all target populations), the reweighted Natural History Study cohort (for SOC in the VHL-RCC cohort), and pre-treatment period data from MK-6482-004 trial participants (for SOC in the VHL-CNS Hb and -pNET cohorts). (Note: Because, by definition, patients had to be alive and metastases-free throughout the pre-treatment period of MK-6482-004, the KM curves from this data source represent time from pre-surgery → surgery in the absence of any competing risks from pre-surgery → metastatic disease or pre-surgery → death. Therefore, in order to make interpretable comparisons between the model predictions versus observed KM data from the MK-6482-004 pre-treatment period, the cause-specific hazards of pre-surgery → metastatic disease and pre-surgery → death were temporarily set to zero when assessing visual fit in the SOC arm of the VHL-CNS Hb and -pNET cohorts.)

- Statistical fit vs. observed time to surgery, metastases, or death: Akaike information criterion (AIC) and Bayesian information criterion (BIC) fit statistics commonly used in PartSA models are not suitable measures of fit with observed data when modelling competing risks. This is because, in the presence of competing risks, fit is determined by a combination of different parametric distributions (in this case, pre-surgery → surgery, pre-surgery → metastatic disease, and pre-surgery → death) rather than a single distribution. Mean squared error (MSE) was therefore used as an alternative diagnostic test to assess fit between modelled vs. observed time to surgery, metastases, or death curves in the MK-6482-004 trial (post-treatment period) and the reweighted Natural History Study cohort. AIC and BIC were used to assess statistical fit between modelled versus observed KM curves in the MK-6482-004 pre-treatment period, as there were no competing pre-surgery → metastatic disease or pre-surgery → death events in this data source.

- Clinical plausibility of long-term extrapolations: Due to clinical implausibility, parametric distributions that resulted in crossing curves (i.e., longer time to surgery, metastases, or death under SOC than belzutifan) were excluded from consideration for the base case. Additionally, distributions that yielded moderate long-term survival projections for belzutifan relative to other candidate distributions (i.e., projections that fall between the highest and lowest predicted curves among the different candidate distributions for belzutifan) were considered plausible and were therefore favoured as base case.

When applying the above criteria, exponential distributions for pre-surgery → surgery resulted appeared to provide a good balance between visual/statistical fit with observed data and plausible long-term extrapolations in each arm. Table 6 describes the selection process for the base-case distributions for pre-surgery → surgery in each target population. The accompanying fit statistics and visual assessments of fit are shown in

Table 7 and Figure 1, respectively.

Table 6 Summary of process to select base-case distribution of pre-surgery → 1st surgery in each model cohort

Description of criterion applied at each step	# of distributions that meet criterion
VHL-RCC cohort	
0 **All candidate parametric functions for pre-surgery → 1st surgery** • Included a total of 7 distributions (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma, and generalized gamma).	7
1 **Initial exclusions based on clinical implausibility** • Excluded 1 distribution (Gompertz) that yielded crossing curves for time to surgery, metastases, or death between the belzutifan and SOC arms.	6
2 **Visual assessment of fit vs. observed time to surgery, metastases, or death** • Based on visual inspection, all distributions yielded a close visual fit between predicted vs. observed curves for time to surgery, metastases, or death in the belzutifan arm (Figure 1a), while only a subset of distributions produced a close visual fit in the SOC arm (Figure 1b). The selection of the base-case distribution therefore prioritized achieving a close visual and statistical fit in the SOC arm. • Three distributions (Weibull, log-normal, and log-logistic) were accordingly excluded due to poor visual fit in the SOC arm.	3
3 **Statistical fit based on MSE vs. observed time to surgery, metastases, or death**	1

	• MSE statistics aligned with findings from visual assessment, with all 3 remaining distributions producing MSEs of comparable magnitude. • The exponential distribution produced the best statistical fit (according to MSE) with the mature Kaplan-Meier curve for time to surgery, metastases, or death in the SOC arm. The exponential distribution was also preferred as base case given the small number of pre-surgery → 1st surgery events in the belzutifan arm.	
4	**Clinical plausibility of long-term time to surgery, metastases, or death** • Due to the application of treatment effect waning in the belzutifan arm, the long-term trajectory of predicted time to surgery, metastases, or death in this arm was comparable between the exponential, gamma, and generalized gamma distributions, which further supported the base-case choice of exponential. • The gamma distribution (2nd-best fitting in the SOC arm) was considered as a scenario analysis.	1 Base case: Exponential
VHL-CNS Hb cohort		
0	**All candidate parametric functions for pre-surgery → 1st surgery** • A total of 7 distributions (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma, and generalized gamma); however, 1 of the 7 distributions (generalized gamma) was excluded due to non-convergence in the belzutifan arm.	7 (6 of the 7 distributions converged)
1	**Initial exclusions based on clinical implausibility** • No exclusions at this step.	6
2	**Visual assessment of fit vs. observed time to surgery, metastases, or death** • No exclusions at this step: Based on visual inspection, all distributions yielded similarly close visual fit between predicted vs. observed curves in both arms (Figure 1b-c).	6
3	**Statistical fit based on MSE vs. observed time to surgery, metastases, or death** • The exponential distribution produced the 2nd-best statistical fit (according to MSE) in the belzutifan arm, and the 3rd- or 1st-best statistical fit (according to AIC or BIC, respectively) in the SOC arm. • Based on good visual and statistical fit in both arms, the one-parameter exponential distribution was preferentially selected as base case over other distributions considering the small number of pre-surgery → surgery events in the belzutifan arm.	1
4	**Clinical plausibility of long-term time to surgery, metastases, or death** • In both arms, there was minimal differentiation between long-term survival predictions under the different candidate distributions (Figure 1b-c), which further supported the choice of exponential as base case.	1 Base case: Exponential
VHL-pNET cohort		
0	**All candidate parametric functions for pre-surgery → 1st surgery** • Included a total of 7 distributions (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma, and generalized gamma) were fitted for SOC in the VHL-pNET cohort using pre-treatment period data from the subgroup of MK-6482-004 participants with pNET.	7
1	**Initial exclusions based on clinical implausibility or non-convergence** • No exclusions at this step.	7
2	**Visual assessment of fit vs. observed time to surgery, metastases, or death**	7

	• Based on visual inspection, all distributions yielded indistinguishably similar visual fit between predicted vs. observed curves for SOC (Figure 1f).	
3	**Statistical fit based on MSE vs. observed time to surgery, metastases, or death** • The exponential distribution produced the best statistical fit (according to both AIC and BIC) in the SOC arm. • Based on good visual and statistical fit in both arms, the one-parameter exponential distribution was preferentially selected as base case over other distributions, particularly given considering the small number of pre-surgery → surgery events in the belzutifan arm.	1
4	**Clinical plausibility of long-term time to surgery, metastases, or death** • Long-term extrapolations were indistinguishably similar under the different candidate distributions, which further supported the choice of exponential as the base-case distribution.	1 Base case: Exponential

Table 7 Fit statistics for parametric models fitted to pre-surgery → 1st surgery in each treatment arm and model cohort

a. Belzutifan (VHL-RCC cohort)

Distributions fitted to pre-surgery → 1st surgery in MK-6482-004 (data cutoff date: 01 April 2022)	**MSE** [1] *(predicted vs. observed time to surgery, metastases or death)*	**Rank by MSE**
Exponential – *base case*	**0.0005611**	**7**
Weibull [2]	0.0002093	3
Gompertz [3]	0.0002006	1
Log-normal [2]	0.0002278	6
Log-logistic [2]	0.0002107	4
Gamma – *scenario analysis*	**0.0002125**	**5**
Generalized gamma	0.0002059	2

Abbreviations: MSE, mean squared error.

Notes:

[1] Statistical fit is assessed by MSE rather than AIC/BIC in the presence of competing risks, i.e., if any patients in the analytical sample experienced a competing transition (pre-surgery → metastatic disease or pre-surgery → death) before transitioning from pre-surgery → 1st surgery.

[2] Excluded due to poor visual/statistical fit in the SOC arm.

[3] Excluded due to implausible crossing of predicted time to surgery, metastases, or death in the belzutifan vs. SOC arms.

b. SOC (VHL-RCC cohort)

Distributions fitted to pre-surgery → 1st surgery in the Natural History Study (2021), reweighted to match the MK-6482-004 population [1]	**MSE** [2] *(predicted vs. observed time to surgery, metastases or death)*	**Rank by MSE**
Exponential – *base case*	**0.0004029**	**1**

Weibull [3]	0.0038250	7
Gompertz [4]	0.0004497	3
Log-normal [3]	0.0031616	6
Log-logistic [3]	0.0030896	5
Gamma – *scenario analysis*	**0.0004294**	**2**
Generalized gamma	0.0004598	4

Abbreviations: MAIC, matching-adjusted indirect comparison; MSE, mean squared error.
Notes:
[1] Parametric distributions were fitted after reweighting the Natural History Study sample to the MK-6482-004 trial population. Details on the MAIC are described in Appendix A.
[2] Statistical fit is assessed by MSE rather than AIC/BIC in the presence of competing risks, i.e., if any patients in the analytical sample experienced a competing transition (pre-surgery → metastatic disease or pre-surgery → death) before transitioning from pre-surgery → 1st surgery.
[3] Excluded due to poor visual/statistical fit in the SOC arm.
[4] Excluded due to implausible crossing of predicted time to surgery, metastases, or death in the belzutifan vs. SOC arms.

c. Belzutifan (VHL-CNS Hb cohort)

Distributions fitted to pre-surgery → 1st surgery in MK-6482-004 (data cutoff date: 01 April 2022)	**MSE** [1] *(predicted vs. observed time to surgery, metastases or death)*	**Rank by MSE**
Exponential – *base case* [2]	**0.0000587**	**2**
Weibull	0.0000635	4
Gompertz	0.0000370	1
Log-normal	0.0000653	6
Log-logistic	0.0000632	3
Gamma	0.0000638	5
Generalized gamma	Did not converge	-

Abbreviations: MSE, mean squared error.
Notes:
[1] Statistical fit is assessed by MSE rather than AIC/BIC in the presence of competing risks, i.e., if any patients in the analytical sample experienced a competing transition (pre-surgery → metastatic disease or pre-surgery → death) before transitioning from pre-surgery → 1st surgery.
[2] The exponential distribution was selected based on good statistical fit in both arms and the small number of pre-surgery → 1st surgery events in the belzutifan arm. Additionally, in the long-term, there was little differentiation between the different candidate distributions in both arms.

d. SOC (VHL-CNS Hb cohort)

Distributions fitted to pre-surgery → 1st surgery using MK-6482-004 pre-treatment period	**AIC**	**BIC**	**Rank by AIC**	**Rank by BIC**
Exponential – *base case* [2]	**592.7**	**594.6**	**3**	**1**
Weibull	593.0	596.8	5	5
Gompertz	592.6	596.4	2	3
Log-normal	591.2	595.1	1	2

Log-logistic	592.8	596.6	4	4
Gamma	593.6	597.4	7	6
Generalized gamma	593.1	598.8	6	7

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; ToT, time on treatment.
Notes:
[1] Statistical fit is assessed by AIC/BIC in the above table because, for SOC in the VHL-CNS Hb cohort, distributions of pre-surgery → 1st surgery were fitted using retrospectively collected data from MK-6482-004 trial participants before they initiated belzutifan. Because patients needed to be alive and metastases-free to enroll in MK-6482-004, there were no competing pre-surgery → metastatic disease or pre-surgery → death events in this data source.
[2] The exponential distribution was selected based on good statistical fit in both arms and the small number of pre-surgery → 1st surgery events in the belzutifan arm. Additionally, in the long-term, there was little differentiation between the different candidate distributions in both arms.

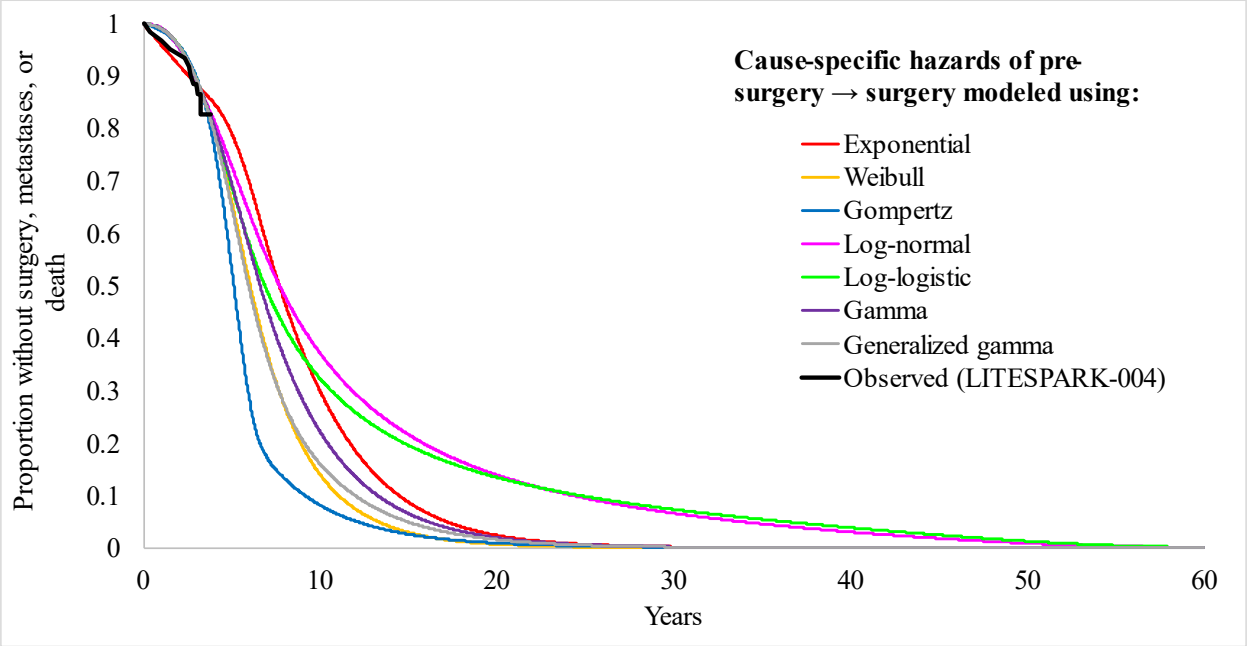
e. SOC (VHL-pNET cohort)

Distributions fitted to pre-surgery → 1st surgery using MK-6482-004 pre-treatment period	**AIC**	**BIC**	**Rank by AIC**	**Rank by BIC**
Exponential – *base case* [2]	**60.2**	**61.2**	**1**	**1**
Weibull	62.0	64.0	4 *(tie)*	4 *(tie)*
Gompertz	62.0	64.0	4 *(tie)*	4 *(tie)*
Log-normal	61.6	63.6	2	2
Log-logistic	61.9	63.9	3	3
Gamma	62.6	64.6	6	6
Generalized gamma	63.2	66.2	7	7

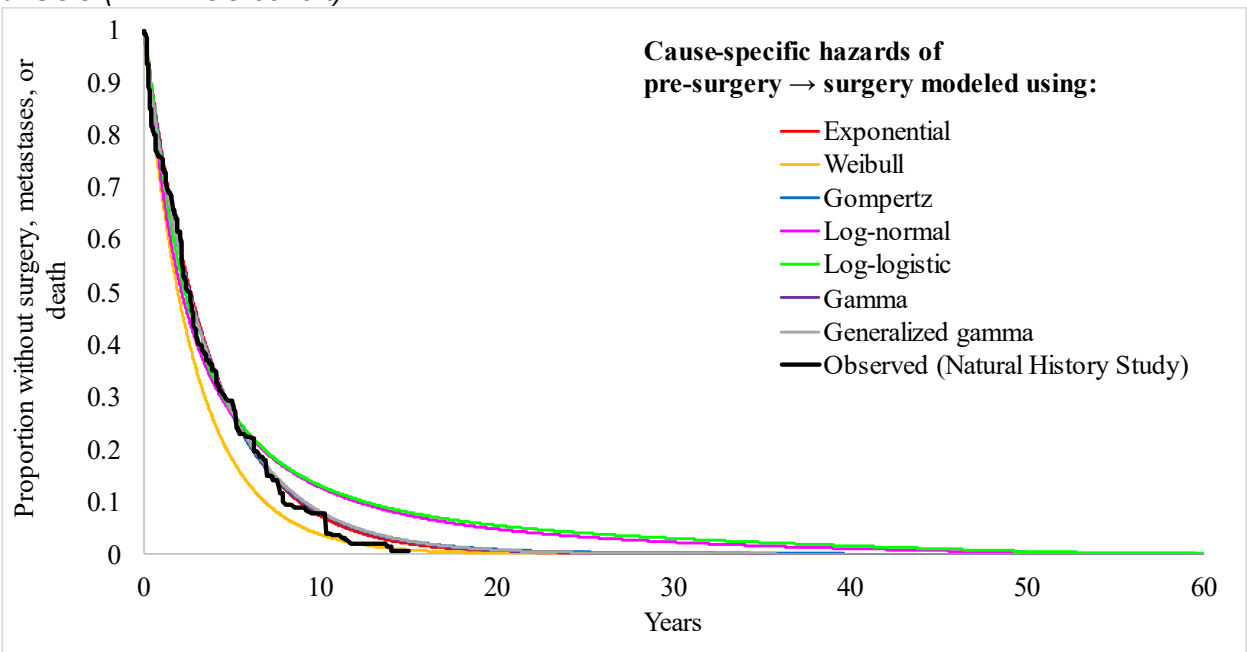
Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; ToT, time on treatment.
Notes:
[1] Statistical fit is assessed by AIC/BIC in the above table because, for SOC in the VHL-pNET cohort, distributions of pre-surgery → 1st surgery were fitted using retrospectively collected data from MK-6482-004 trial participants before they initiated belzutifan. Because patients needed to be alive and metastases-free to enroll in MK-6482-004, there were no competing pre-surgery → metastatic disease or pre-surgery → death events in this data source.
[2] The exponential distribution was selected based on best statistical fit according to both AIC and BIC, the indistinguishably similar visual fit and long-term extrapolations under all 7 candidate distributions, and the small number of pre-surgery → 1st surgery events in the SOC arm for this cohort. (There were no pre-surgery → 1st surgery events after treatment initiation in the MK-6482-004 subgroup with pNET at baseline; therefore, distributions could not be directly fitted for the belzutifan arm of the VHL-pNET cohort.)

Figure 1. Visual assessments of fit between modeled vs. observed time to surgery, metastases, or death in each model cohort and treatment arm

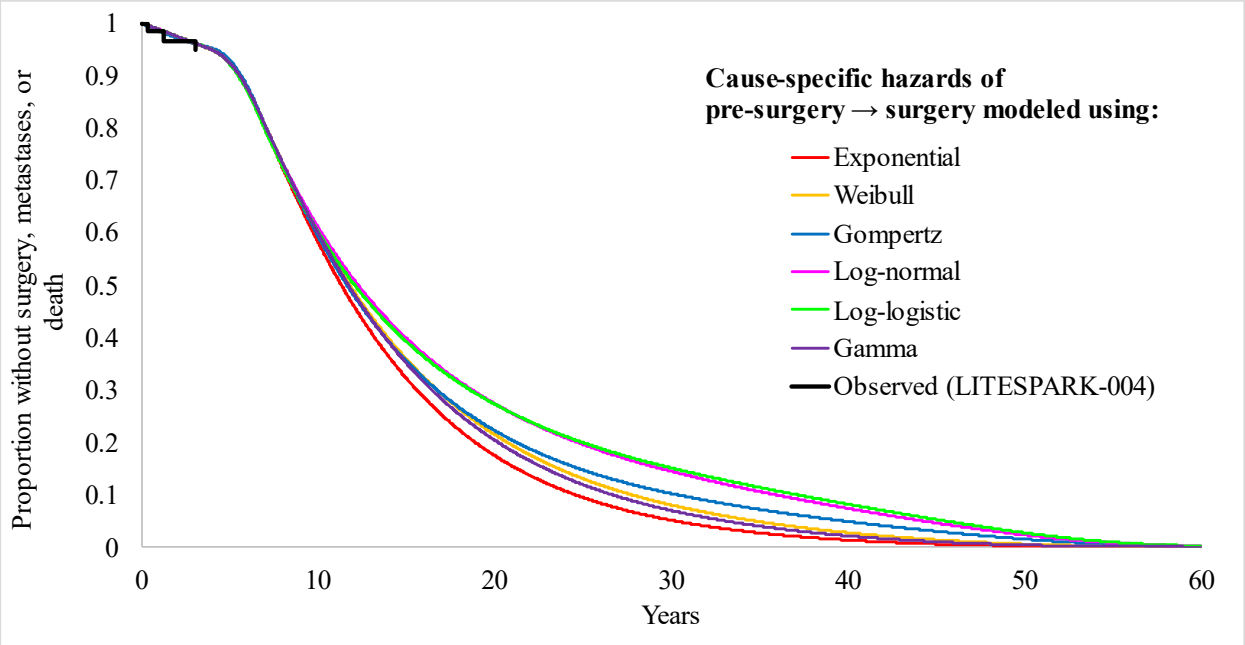
a. Belzutifan (VHL-RCC cohort)



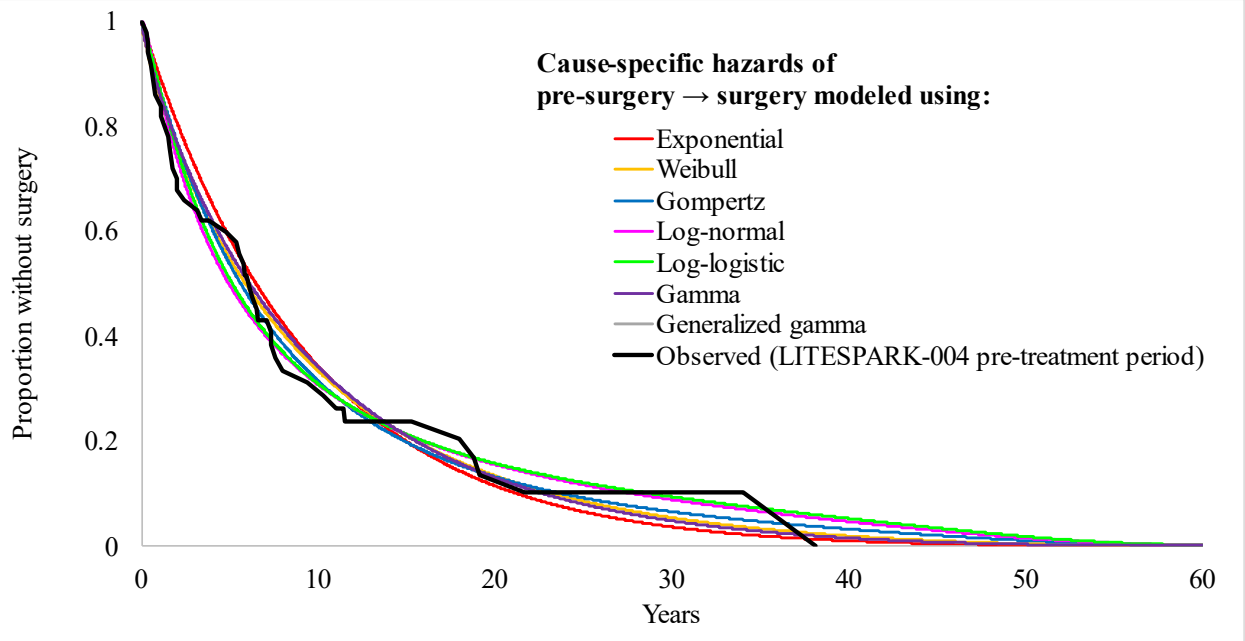
b. SOC (VHL-RCC cohort)



c. Belzutifan (VHL-CNS Hb cohort)

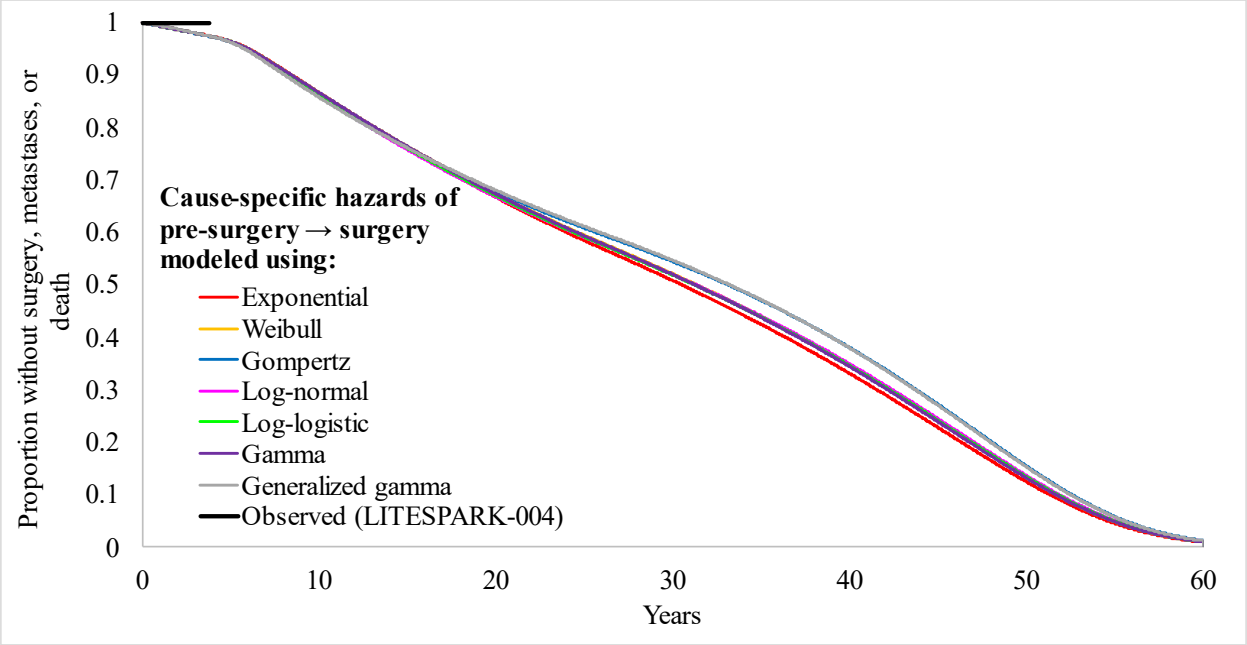


d. SOC (VHL-CNS Hb cohort)



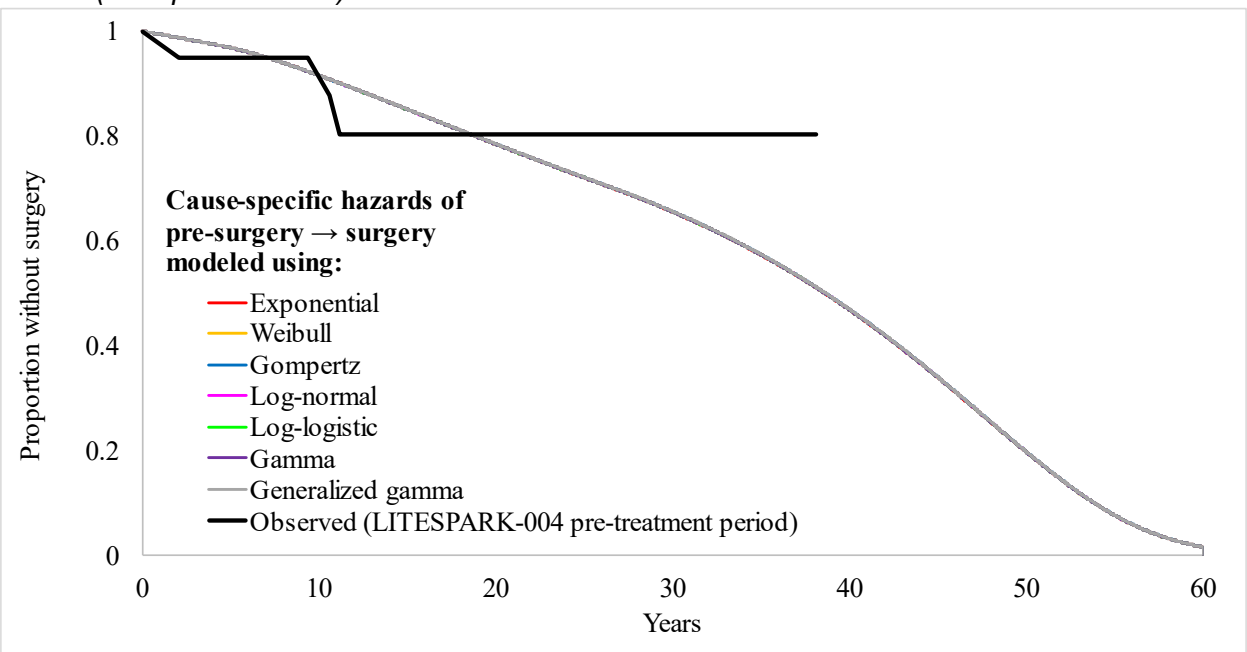
Note: The cause-specific hazards of pre-surgery → metastases and pre-surgery → death are temporarily set to 0 when generating validation figure (d), as the Kaplan-Meier curve from the MK-6482-004 pre-treatment period represents time from pre-surgery → surgery in the absence of any competing risks from pre-surgery → metastatic disease or pre-surgery → death.

e. Belzutifan (VHL-pNET cohort)



Note: There were no observed transitions from the pre-surgery state in MK-6482-004 as of the 01 April 2022 data cutoff date. Therefore, the 7 candidate distributions were fitted to pre-surgery → surgery in the SOC arm, and a HR approach was applied in the belzutifan arm.

f. SOC (VHL-pNET cohort)



Note: The cause-specific hazards of pre-surgery → metastases and pre-surgery → death are temporarily set to 0 when generating validation figure (f), as the Kaplan-Meier curve from the MK-6482-004 pre-treatment period represents time from pre-surgery → surgery in the absence of any competing risks from pre-surgery → metastatic disease or pre-surgery → death. Due to the small number of pNET surgeries observed during the pre-treatment period, all fitted distributions for pre-surgery → surgery appear similar to the exponential distribution, as shown above.

b) As detailed above, the exponential distribution was most appropriate to model time to surgery, metastases or death. The rationale for choosing the exponential distribution for the pre-surgery → surgery transition is provided above; nevertheless, the model includes the functionality to choose alternative parametric distributions and the gamma distribution was chosen as an alternative for the RCC cohort to be explored in scenario analyses. For the pre-surgery → metastatic disease and pre-surgery → death transitions, the exponential distribution was also selected because the number of these transitions were small in both arms and therefore the exponential distribution prevents overfitting or convergence issues. The default use of exponential distributions for infrequent health state transitions is consistent with several prior Markov cohort models submitted to NICE. Exponential distributions were used to model direct transitions from disease-free → death in TA830 and recurrence-free → death in TA837 and TA766, as these direct transitions from the starting state to death were infrequently observed. The use of exponential distributions for infrequent transitions such as pre-surgery → metastatic disease and pre-surgery → death was expected to have minimal impact on the model results. (As an example, there were relatively few pre-surgery → surgery events for SOC in the VHL-pNET population and changing the distribution for this transition from exponential to any of the other six candidate distributions has a trivial impact on the cost-effectiveness results.)

c) The rates in Tables 46 and 47 of the CS are yearly exponential rates converted from weekly exponential rates (which are reported in the economic model) for ease of interpretation. The calculation of transition probabilities derived from cause-specific hazards are as follows:

For each individual transition starting from the pre-surgery state, transition probabilities in each weekly cycle were calculated within the model as a function of the cause-specific hazards for all three transitions from this state. The following calculation steps were performed:

1. For each cause *k* of transitioning away from pre-surgery (i.e., surgery, metastatic disease, or death), the average cause-specific hazard within the cycle from week *(t-1)* to *t* was calculated as:

$\overline{h}_k(\mathrm{t}) = H_k(\mathrm{t}) - H_k(\mathrm{t}-1),$

where $H_k(.)$ is the cause-specific cumulative hazard of cause *k* (based on the parametric function selected to model cause *k*).

2. The average hazard of any transition from pre-surgery within the cycle from week *(t-1)* to *t*, denoted $\overline{h}_{total}(\mathrm{t})$, was calculated as the sum of the average cause-specific hazard for all three causes within that cycle. This hazard was converted into a probability using the formula:

$1 - e^{-\overline{h}_{total}(\mathrm{t})}$

3. In each cycle, the relative contribution of each cause *k* to the overall hazard of transitioning from pre-surgery was derived as:

$\frac{\overline{h}_k(\mathrm{t})}{\overline{h}_{total}(\mathrm{t})}$

This represents the probability of having had a transition of type *k* given that a transition from the pre-surgery has occurred within the cycle. The relative contribution of cause *k* was then multiplied by the probability of any transition from pre-surgery within the cycle to obtain the transition probability corresponding to cause *k*.

The treatment benefit of belzutifan following surgery can be maintained as surgery does not necessitate treatment discontinuation. Duration of treatment in the trial was until unacceptable treatment-related toxicity or unequivocal disease progression. There will undoubtedly be some patients who are on the threshold of requiring surgery at treatment initiation with belzutifan. These patients may go on to have surgery but continue taking belzutifan for other tumour manifestations and therefore retain treatment benefit. Furthermore, even if a patient was to discontinue belzutifan at the point of surgery, there is no evidence that the treatment effect ceases. Conceptually, there is no reason a priori that surgery should affect belzutifan efficacy.

d) It would be inappropriate to consider transitions after surgery to be equal in each treatment arm given that surgery does not necessitate belzutifan discontinuation nor does it render belzutifan treatment ineffective.

B 7. Priority. TPs from pre-surgery → surgery health states for the VHL-pNET and VHL-CNS Hb cohorts in the SoC arm are stated to be informed by the pre-treatment period from the MK-6482-004 trial as the natural history data could not be used to inform these TPs. This is explained by lack of information in the VHL NHS on whether patients had pNET or CNS.

a) Please clarify if the explanation above is correct.

b) Please explain the implications and potential biases resulting from the estimation of TPs from pre-surgery → surgery health states based on the pre-treatment period of the MK-56482-004 trial.

c) Please provide estimates of the respective TPS for the VHL-RCC cohort using the pre-treatment data from MK-6482-004 trial instead of the VHL NHS. Comment on any difference between TPs estimated using both methods and provide CE estimates in a scenario analysis which should be fully based on pre-treatment data for all arms. In an additional scenario analysis, please consider making an adjustment (applying a HR) to the estimated TPS for the VHL-pNET and VHL-CNS cohorts based on any difference observed in the VHL-RCC cohort.

MSD response:

a) We can confirm that the above explanation is correct. It was not feasible to identify whether patients in the VHL Natural History Study had CNS Hb or pNET at the patient-level index date.

b) There are potential biases in using two different data sources for transitions within the same cohort: the pre-treatment period of the MK-6482-004 trial and the VHL Natural History Study may have had different treatment options, disease management and pathways. There may also be implications in comparing these cohorts with the RCC cohort which used the VHL Natural History solely to inform the

SOC arm. However, in the absence of alternative data sources, the pre-treatment period data was the best available data source to estimate time to surgery in the VHL CNS Hb and pNET cohorts; furthermore, using this data source uses the same patients as were treated with belzutifan.

c) Although MK-6482-004 pre-treatment period data represented the best available evidence for the pre-surgery → surgery transition under SOC in the VHL-CNS Hb and VHL-pNET cohorts, this data source is subject to one important limitation: by definition, patients in MK-6482-004 were alive and metastases-free prior to belzutifan initiation, given the trial enrolment criteria. Consequently, pre-treatment period data could not be used to generate a full set of transition probabilities from the pre-surgery state. In contrast, for the VHL-RCC cohort, a full set of transitions probabilities from the pre-surgery state under SOC could be reliably generated based on VHL Natural History Study results; the Natural History Study was thus considered the best available evidence for these transitions under SOC in the VHL-RCC cohort and was used as the base-case data source for these transitions.

Nevertheless, we have updated the Excel model to enable a scenario analysis using MK-6482-004 pre-treatment period data to model pre-surgery → surgery under SOC in the VHL-RCC cohort, per the approach used for SOC in the other two cohorts. Under this scenario, the incidence rate and distribution of non-RCC surgeries under SOC in the VHL-RCC cohort is also based on MK-6482-004 pre-treatment data. The table below summarizes the model inputs that differ under the base case vs. this new scenario analysis. As shown, the inputs were not dramatically different under these data sources; when using pre-treatment period data, the rate of pre-surgery → first RCC surgery decreased, while the incidence of non-RCC surgeries increased. Additionally, the ICER improves when using this data source.

Table 8 Parameter values for SOC in the VHL-RCC cohort that change in the scenario analysis using MK-6482-004 pre-treatment data

Clinical inputs for SOC in the VHL-RCC cohort	**Base case** (*Data source:* VHL Natural History Study RCC cohort)	**Scenario analysis** (*Data source:* MK-6482-004 pre-treatment period data)
Weekly exponential* rate of pre-surgery → RCC surgery	0.00487	0.00207
Weekly incidence rate of non-RCC tumour surgeries	0.00344	0.00438

Distribution of surgeries for non-RCC tumours		
CNS Hb surgery	52.4%	67.4%
pNET surgery	3.4%	7.0%
Adrenal lesion surgery	25.1%	2.3%
Endolymphatic sac tumour surgery	4.5%	0.0%
Epididymal cystadenoma surgery	0.1%	2.3%
Retinal Hb surgery	14.5%	20.9%
ICER	**£73,095**	**£42,622**

*<u>*Note:*</u> In the Excel model, all 7 candidate distributions for pre-surgery → surgery (i.e., exponential, Weibull, Gompertz, log-normal, log-logistic, generalized gamma, and gamma) are available for selection under both data sources.

A new dropdown menu ("Select data source for surgery risks under SOC in the VHL-RCC cohort") has been added to the *Specifications* sheet to toggle between the above scenarios.

B 8. Priority. On page 151 of the CS it is stated that "given the absence of evident VHL-tumour related deaths in MK-6482-004 and the low mortality rates observed in the VHL Natural History Study, the per-cycle TP from pre-surgery → death was set equal to the maximum of (i) the background mortality, using national mortality rates based on the age and sex distribution of the model cohort in each cycle, and (ii) the mortality rate of the VHL Natural History Study RCC cohort." Please provide further details on the mortality rate estimated from the VHL Natural History RCC cohort, as it is also not explained in the description of the mortality rate of the VHL-RCC cohort of the SoC arm (page 158 of the CS). Similarly, please provide details on the mortality rate of the VHL Natural History Study for the CNS-Hb cohort.

MSD response:

Within the VHL Natural History Study RCC cohort (which was reweighted to match key baseline characteristics of MK-6482-004 trial population), a parametric multistate modelling approach was used to estimate the cause-specific hazards of the pre-surgery → death transition. As described in response to question B 6. , an exponential distribution was used for this transition due to the small number of direct transitions from pre-surgery to death. In order to fit an exponential distribution to this specific health state transition, the two competing transitions from the pre-surgery

state were treated as censoring events. Namely, patients who experienced their first post-baseline RCC surgery or metastatic disease prior to death were censored when modelling the pre-surgery to death transition and were thus treated as lost to follow-up at the time of the earlier competing event. After these additional censoring criteria were applied to the patient-level time-to-event data for each transition, parametric curve fitting was performed using the survival analysis package flexsurvreg in R software (R Development Core Team, Vienna, Austria), similar to the process for fitting parametric functions for a partitioned survival model.

An analogous approach was used to fit an exponential distribution to the cause-specific hazards of pre-surgery → death in the VHL Natural History Study CNS Hb cohort (which was reweighted to match key characteristics of the MK-6482-004 subgroup with CNS Hb at baseline). When modelling this transition in the VHL Natural History Study CNS Hb cohort, patients who experienced a competing event (i.e., their first post-baseline CNS Hb surgery or metastatic disease) prior to death were censored. An analogous approach was similarly used to fit an exponential distribution to the cause-specific hazards of pre-surgery → death in the VHL Natural History Study pNET cohort (which was reweighted to match key characteristics of the MK-6482-004 subgroup with pNET at baseline). When modelling this transition in the VHL Natural History Study pNET cohort, patients who experienced a competing event (i.e., their first post-baseline pNET surgery or metastatic disease) prior to death were censored.

<u>References for the multistate modelling method:</u>

Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness Analysis in R Using a Multi-state Modeling Survival Analysis Framework: A Tutorial. Med Decis Making. 2017;37(4):340-352.

Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of Survival Probabilities for Use in Cost-effectiveness Analyses: A Comparison of a Multi-state Modeling Survival Analysis Approach with Partitioned Survival and Markov Decision-Analytic Modeling. Med Decis Making. 2017;37(4):427-439.

National Institute for Health and Care Excellence. DSU Technical Support Document 19: Partitioned survival analysis for decision modelling in health care: a critical review. June 2 2017.

Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007;26(11):2389-2430.

B 9. Priority. The risks of short- and long-term complications following surgery for the VHL-RCC cohort are in the majority doubled when considering the MHRA label population than estimated from the Optum study (Tables 53-59 of the CS).

a) **Please justify this assumption with appropriate evidence. Furthermore, only for chronic kidney disease the risk is lower in Table 57 of the CS. Please justify the lower risk for chronic kidney disease as compared to other complications.**

b) **Please also explain why the risks of specific long-term complications in Tables 57-59 of the CS (i.e. complications related to end stage renal disease and/or dialysis, cerebral vasculature occlusion or stroke, secondary diabetes or exocrine pancreatic insufficiency) have been assumed to be substantially higher as compared to the risks estimated from the Optum study.**

c) **Considering the uncertainty around these risks, please run multiple scenario analyses varying the risk of short and long-term complications using appropriate ranges (and justify these ranges).**

MSD response:

a) The risks of short- and long-term complications following surgery in the VHL-RCC cohort are increased compared to the estimates from the Optum study due to the nature of the population specified by the GB label. The decision points for the target population and for patients in the Optum study are different. The target population of the current appraisal include patients with more severe manifestations of VHL disease for whom surgery is unsuitable or undesirable. This implies that they would experience a significantly increased risk of complications as a result of these

surgeries. The risk is lower for chronic kidney disease as it is expected after a 'last-resort' RCC surgery that all patients would have some form of renal impairment, of which 80% would have end-stage renal disease (ESRD) requiring dialysis and the remainder (20%) would have chronic kidney disease, an assumption which was informed by clinical expert opinion.

b) The risk of ESRD (in the VHL-RCC cohort), secondary diabetes or exocrine pancreatic insufficiency and immunocompromisation (in the VHL-pNET cohort) are increased compared with the Optum study. This is a result of surgery being a 'last resort' for the assessed patient population, and such a surgery is understood to lead to absent/limited organ function. These complications are therefore the metabolic consequences of surgery resulting from partial/complete removal of the organs with primary tumour burden. The complications of cerebral vasculature occlusion or stroke in the VHL-CNS Hb cohort is increased versus Optum as there is a substantially higher risk of this in the target population given their unsuitability for surgery. It should be noted that patients in the Optum study were not classified as "surgery unsuitable or undesirable" as is stipulated by the MHRA label. Therefore, the risk of these specific complications is much higher in the target population than in patients within the Optum study.

c) Please see Table 9 below for scenario analyses around the risks of surgical complications. A scenario has been included in which the risk of surgical complications and perioperative mortality are set equivalent to those reported in the Optum study, with the exception of the metabolic consequences of RCC & pNET surgery and stroke for CNS Hb surgery. The remaining scenarios explore these metabolic consequences and stroke complications, including ICERs associated with scenarios where the complication risk is set to 10% lower than in the base case (in absolute terms).

Table 9 Scenario analyses for the risk of surgical complications

	ICER (£/QALY)		
	VHL-RCC cohort	*VHL-CNS Hb cohort*	*VHL-pNET cohort*
Base case	**73,095**	**56,933**	**77,649**
Risk of surgical complications and perioperative mortality as Optum study (except the metabolic consequences in	74,881	64,124	82,773

VHL-RCC and VHL-pNET cohorts and stroke in VHL-CNS Hb cohort)			
Risk of end-stage renal disease and/or dialysis 70% (-10% from base case) and chronic kidney disease 30% (adjusted to account for 100% renal impairment) as a complication of RCC surgery	77,839	Base case	Base case
Risk of cerebral vascular occlusion or stroke at 75% as a complication of CNS Hb surgery (-10% from base case)	Base case	57,912	Base case
Risk of secondary diabetes or exocrine pancreatic insufficiency and immunocompromisation at 90% as a complication of pNET surgery (-10% from base case)	Base case	Base case	81,548

B 10. Priority. On page 167 of the CS, it is mentioned that "to align with the surgery-unsuitable or -undesirable population, the perioperative mortality risks were adjusted by a factor of 2.0 (i.e. doubled) for each cohort to reflect the increased risk of perioperative mortality as surgical procedures are a 'last resort' option in the MHRA label population in line with clinical expert." Please explain the reasoning behind the adjustment and comment on the appropriateness of the sources used to inform perioperative mortality risk by VHL cohort.

MSD response:

In the absence of alternative data sources, the Optum study was used to estimate perioperative mortality risk. As the MHRA label stipulates a patient population for whom surgery is unsuitable or undesirable, and in line with clinical expert opinion, it is expected that surgery in this group is associated with higher risk and therefore the likelihood of perioperative mortality is increased compared to observed data in the Optum study. Nevertheless, to explore the impact of this assumption, a scenario which sets perioperative mortality risk equivalent to that of the Optum study is presented in the scenario analysis reported in Table 9 showing a slight increase in the ICER.

B 11. Priority. Please explain if and how progression-free survival was included in the economic model (for all subgroups). For example, based on Figure 7, please explain too what happened after month 34,

since between month 34 and 36 the remaining patients at risk was halved.

MSD response:

We can confirm that progression-free survival, as informed by KM plots reported in the CS for the VHL-RCC, VHL-pNET and VHL-CNS Hb subgroups, was not included directly in the economic model. As detailed in the HRQoL section of the CS, disease progression (per RECIST 1.1 criteria) along with other response levels, informs the HRQoL of VHL patients. However, progression-free survival as a continuous outcome is not directly included in the economic model.

Maximum follow-up in the MK-6482-004 trial was 3.84 years (46.04 months). Given that patients in the clinical trial did not all initiate belzutifan simultaneously, after month 34 fewer and fewer patients have sufficiently long follow-up to provide data at subsequent time points. These patients are therefore censored beyond the time of last observation, which is why the number of patients at risk begins to drop sharply after month 34.

B 12. Priority. Please conduct scenario analyses where the assumptions around the derivation of the transition probabilities in the belzutifan arm are plausibly varied.

MSD response:

Scenario analyses around the estimated transition probabilities are presented in Appendix J1.4. For ease, they are presented in Table 10 below for each cohort. Additional scenarios around the parametric models for the pre-surgery → surgery transition are also presented demonstrating the lack of variability in the ICER. For this transition in the VHL-pNET cohort, an alternative parametric distribution in the SOC arm is presented as TTS in the belzutifan arm was estimated using a HR approach as no pNET surgeries were observed in the trial. The assumptions on how these transition probabilities are derived do not have plausible alternatives to consider; therefore, we vary the efficacy inputs in deterministic sensitivity analyses to explore parameter uncertainty.

Table 10 Deterministic sensitivity analyses for efficacy and transition probabilities in the belzutifan arm

	ICER vs. comparator: Belzutifan vs. SOC *(£/QALY)*	
	Low input value	*High input value*
VHL-RCC cohort base case	**73,095**	
Distribution for pre-surgery → surgery in the belzutifan arm (VHL-RCC cohort): Gamma	76,127	
Distribution for pre-surgery → surgery in the belzutifan arm (VHL-RCC cohort): Gen Gamma	81,465	
Exponential rates of EF→MD under belzutifan +/- 10%	73,111	73,080
Exponential rates of EF→Death under belzutifan +/- 10%	73,125	73,066
VHL-CNS Hb cohort base case	**56,933**	
Distribution for pre-surgery → surgery in the belzutifan arm (VHL-CNS Hb cohort): Gompertz	56,578	
Exponential rates of EF→MD under belzutifan +/- 10%	56,934	56,933
Exponential rates of EF→Death under belzutifan +/- 10%	56,938	56,929
VHL-pNET cohort base case	**77,649**	
Distribution for pre-surgery → surgery in the SOC arm (VHL-pNET cohort): Log-normal	77,771	
Exponential rates of EF→MD under belzutifan +/- 10%	77,649	77,649
Exponential rates of EF→Death under belzutifan +/- 10%	77,649	77,648

Time on treatment

B 13. Priority. Please explain in detail how time to treatment discontinuation has been included in the model. Please clarify if it is expected to have an impact on both costs and effects, why and where to see this in the economic model. Please present the results of scenario analyses based on selection of the different time on treatment parametric curves.

MSD response:

In the base case, treatment discontinuation rates in each non-metastatic health state are based on the parametric curve fitted to observed time on treatment (ToT) in the trial. Treatment discontinuation rates represent the transition probabilities from a given on-treatment health state (e.g., pre-surgery, on-treatment) to the corresponding off-treatment state (e.g., pre-surgery, off-treatment). ToT was thus modelled based on time spent in the on-treatment health states, which was tracked as part of the Markov model structure. These can be seen in the electronic model Markov traces in sheets *Trace_TxReg1_RCC*, *Trace_TxReg1_CNSHb*, *Trace_TxReg1_pNET* for each respective cohort. There are no associated treatment costs in the off-treatment health states. When applying treatment effect waning, clinical efficacy inputs for patients who have discontinued belzutifan (including the efficacy of belzutifan in preventing transitions to surgery, metastases, or death, reducing the incidences of non-primary tumour surgeries, and inducing primary tumour response) were assumed to linearly converge to those of the SOC arm between *[Wane_Start]* and *[Wane_End]* years.

An alternative scenario with a Weibull distribution for time on treatment is presented in *Scenario analyses* in B.3.11 of the CS and produced in Table 11 below. Weibull was the only alternative plausible distribution when considering statistical fit and clinical plausibility of its long-term projections of ToT.

Table 11 Scenario analysis exploring Weibull distribution for ToT

	ICER (£/QALY)		
	VHL-RCC cohort	*VHL-CNS Hb cohort*	*VHL-pNET cohort*
Base case	**73,095**	**56,933**	**77,649**
Distribution for belzutifan ToT: Weibull	91,265	71,497	96,471

MSD acknowledge that long-term ToT is an area of uncertainty, which future data collection may help resolve.

B 14. Priority. Please explain in detail the assumptions behind residual treatment effect (waning) after discontinuation. It seems that this has been implemented as a fixed time, however it is unclear why waning

is not dependent on time on treatment (e.g., patients discontinuing early should have less or none residual benefit of treatment). Please explain where this has been included in the model. Please present the results of scenario analyses based on different assumptions on treatment effect waning (e.g., no residual benefit, dependent on time on treatment, etc.). The current scenarios seem insufficient to capture the uncertainty associated to this aspect of the model.

MSD response:

Residual treatment benefit relates to an assumption that is relevant for treatment discontinuations beyond the time of the maximum follow-up in the MK-6482-004 trial (3.84 years). The impact of treatment discontinuation on belzutifan efficacy for discontinuations that occur up to 3.84 years since treatment initiation *has already been captured* in the data collected with the clinical trial. In simple terms, data reported in the clinical trial on time to surgery, metastases or death is informed by a range of treatment discontinuations (e.g. early, after several years, or not at all) observed in the clinical trial. There is therefore limited uncertainty as to the affect in belzutifan efficacy of discontinuations that occur within 3.84 years from treatment initiation.

What is less certain is the impact of treatment discontinuations that occur beyond 3.84 years from treatment initiation. To reflect this uncertainty, an assumption of treatment effect waning has been included to assess the impact of discontinuation on belzutifan efficacy. To model this, a period of residual treatment benefit equivalent to the maximum duration of trial follow-up (3.84 years) has been included to reflect the fact that deviating from the belzutifan data soon after treatment discontinuation would be implausible: this data (up to 3.84 years) already accounts for discontinuations. Only once this duration has elapsed does treatment effect waning begin to influence belzutifan efficacy.

Furthermore, applying treatment waning during the observed trial period would worsen the alignment between observed vs. predicted curves for time to surgery, metastatic disease, or death in the belzutifan arm. It would therefore be inappropriate to consider a treatment effect waning assumption of no residual benefit (i.e. before the maximum follow-up period of the trial is complete). After the 3.84-year period has

elapsed, waning *is* dependent on time on treatment as it is not initiated until patients discontinue treatment and move to the "off-treatment" version of the respective non-metastatic health state. These can be seen in the electronic model Markov traces in sheets *Trace_TxReg1_RCC*, *Trace_TxReg1_CNSHb*, *Trace_TxReg1_pNET* for each respective cohort and more specifically column G in each sheet shows the proportion of waning among patients who have discontinues belzutifan. Alternative scenarios on the time point to initiate treatment waning would only be accurate beyond the trial follow-up period (i.e. greater than 3.84 years) and these scenarios would result in a lower ICER.

Adverse events

B 15. Priority. In the model, only anaemia and fatigue are included as adverse events in the economic model. Please include in the model (both on costs and HRQoL sides) all adverse events meeting the criteria for inclusion in the economic analyses:

d) Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan: based on Table 117 in the CS, it seems that hypertension should have been included in the model too.

e) Grade ≥3 TRAEs occurring in >0% of patients treated with belzutifan: based on Table 116 in the CS, it seems that hypoxia and urinary tract infection should have been included in the model too.

MSD response:

We apologise for the confusion here as there is a mistake in the CS. In the *Adverse events* section of B.3.3 the first sentence on page 188 should read:

"The model considers Grade ≥3 AEs occurring in ≥5% of patients treated with belzutifan, **and** Grade ≥3 TRAEs occurring in >0% of patients treated with belzutifan."

Therefore, the criteria for inclusion in the model is only met for anaemia and fatigue.

B 16. Priority. Adverse event rates are sourced from the MK-6482-004 trial, but as mentioned in the CS, this population is not reflective of

the licensed population. Please clarify whether the adverse event rates observed in the MK-6482-004 trial are expected to be comparable between the two populations. Please present the results of scenario analyses where these rates are varied within a range of realistic values, providing justification for the selected ranges. Alternatively, if the trial population is expected to be less severe than the licensed population, the criteria for including adverse events in the model could be relaxed, and for example include all adverse events in Table 118 and the two observed deaths, to reflect that the licensed population is more severe.

MSD response:

The licensed population has higher risk with respect to the likelihood of experiencing surgical complications due to the eligibility restriction to patients for whom surgery is unsuitable or undesirable. This does not mean that they are more likely to experience adverse events from belzutifan than the trial population; therefore, it would be inappropriate to include all adverse events.

B 17. Priority. Please clarify what the consequences of treatment interruptions due to adverse events are and whether this has been included in the model or not. Since this was observe in a large proportion of participants, it might impact the model results.

MSD response:

As belzutifan efficacy is estimated from the MK-6482-004 trial data, any impact from treatment interruptions during the trial period would already be accounted for. Furthermore, there is no evidence to suggest that belzutifan efficacy would cease during a brief interruption in the treatment period.

Health-related quality of life

B 18. Priority. Please provide full details of the VHL patient survey used to source health-related quality of life data for the cost-effectiveness analyses. Please clarify whether a) the population in the of the VHL patient survey and in the KEYNOTE-564 trial are comparable, b) these

are comparable to the MK-6482-004 trial population and c) they are representative for the UK patient population for all three subgroups.

MSD response:

a) VHL patients in the non-metastatic health states with a CR in the MK-6482-004 trial have a HRQoL comparable to disease-free patients in the KEYNOTE-564 trial. This is because both groups have had previous surgery for RCC tumours, and their tumours are being monitored for recurrence (in KEYNOTE-564) or progression (in MK-6482-004). Comparability between KEYNOTE-564 and patients in the VHL patient survey is not necessary, it is only necessary for each of these to be comparable to the respective response levels of patients in the MK-6482-004 trial. Patients in the disease-free (without toxicity) state in KEYNOTE-564 need to be reflective of CR in the MK-6482-004 trial (which is justified above) and patients in the VHL patient survey need to be reflective of the remaining response levels (with the exception of PD in the VHL-CNS Hb cohort). Furthermore, utility values for response levels need to be consistent with a better response being associated with a higher utility value which is also reflected in the use of these sources in the economic model.

b) Please see response to a) above.

c) As described in the response to question A 21, due to the rarity of VHL disease in the UK (the prevalence of VHL disease is between 1 in 77,340 and 1 in 68,493, with between 55 and 120 patients in England likely to be eligible for treatment with belzutifan, as described in section B.1.3 of the company submission), there is a lack of published data on the population characteristics of patients with VHL disease in the UK, much less data stratified specifically for the subgroups of patients with different tumour manifestations. The most up-to-date published information on the UK VHL disease patient population is that presented in the publication of the national audit of VHL disease in the UK by Maher et al. (https://www.nature.com/articles/s41416-022-01724-7), which does not report detailed/tabulated population characteristics. Therefore, it is not feasible to provide a detailed quantitative comparison between the characteristics of the UK VHL disease patient population and the patient population included in any study.

B 19. Priority. Please discuss the (face) validity of the EQ-5D values presented in Table 40, 41 and 42 (e.g., compare the values presented in this submission with other sources of utilities for this or similar diseases – e.g., studies retrieved by the SLR, and with the utility values for the general population – also indicate if these were validated with clinical experts and how). In addition, please answer the following questions:

a) **The text above Table 40 mentions n = 16 patients with metastatic disease whereas on Table 39 the number indicated seems to be n = 58. Please clarify this point too.**

b) **Please clarify the differences between metastatic, progressive and advanced disease and how these are differentiated in the economic model.**

c) **The number of observations in Table 40 and 41 are in general small leading to large standard deviations. Please show what probability distributions were assumed for the PSA and what range of values were sampled in the model (e.g., please provide probability distribution parameters and/or confidence intervals).**

MSD response:

The EQ-5D values described in Tables 40-42 of the CS show internal consistency: worsening disease is consistently associated lower EQ-5D scores:

- Table 40: Among all patients (n=220), patients with metastatic disease have worse utility scores than patients without metastatic disease
- Table 41: Among patients with metastatic disease, whilst acknowledging the small sample (n=16), patients with progressive disease have worse utility scores than those with stable disease
- Table 42: Among patients without metastatic disease (n=195), patients with progressive disease have worse utility scores than patients with stable disease

In the SLR conducted to identify studies reporting on HRQoL in VHL, no studies were identified which reported EQ-5D scores in a VHL population.

The utility values reported in Tables 40-42 are all lower than age- and sex-match utility values estimated for the general population in the UK, which is consistent with the disease burden faced by patients with VHL.

Utility values were presented and discussed in engagements with clinical experts conducted for the preparation of the current appraisal. The utility values were deemed to be plausible, whilst acknowledging the differences in terms of severity of tumour manifestations in the VHL RW QoL Disease Burden Study and the population eligible for belzutifan per the GB marketing authorisation.

Comparisons of utilities values included in the base case to similar diseases has inherent challenges given the rarity of VHL. We recognize there are limitations in the dataset providing E5-5D scores in its ability to capture the disease severity in the target population of the current appraisal. Where necessary, utility values previously accepted in appraisals in disease areas which could be used as a proxy to represent specific VHL health states (e.g. motor neurone disease as a proxy for the event-free post-surgery state in patients VHL-CNS Hb for whom localised procedures are unsuitable or undesirable). The EQ-5D scores obtained for patients with metastatic disease in the VHL RW QoL Disease Burden Study (0.412) do show some differences when compared with utility values which have been accepted in metastatic RCC (0.772; TA830) and metastatic pancreatic cancer (0.67; TA476). However, as discussed repeatedly in the CS, the target population of the current appraisal has severe manifestations of VHL, including multi-systemic disease which is associated with disease burden above and beyond that which is associated with metastatic disease related to one of the affected organs.

a) Table 40 in the CS refers to 16 patients with metastatic disease, while Table 39 does not refer to patients with metastatic disease but rather refers to disease status.

b) In the VHL patient survey, patients with metastatic disease necessarily had tumours that had metastasised. Progressive disease in the metastatic disease state of the model (shown in Table 41 of the CS) reflects the post-progression aspect of the metastatic disease state. Progressive disease in the non-metastatic health states

of the model (shown in Table 42 of the CS) reflects the response level self-reported by patients i.e. CR, PR, SD or progressive disease (PD), prior to metastatic disease. In the economic model, metastatic disease is sometimes referred to as advanced disease i.e. advanced RCC/pNET.

c) A Beta distribution was assumed for the utility values and the parameters are shown in Table 12 below (note: these can be found in the *PSA Setup* sheet of the economic model). The standard errors are based on the original sources of utility inputs and percentile matching is used to preserve rank of utility values from best to worst health state.

Table 12 Utility value input parameters and distributional assumptions considered in the PSA

Input parameter	Distribution	Base-case value	Alpha	Beta	Standard error
Utility in pre-surgery, surgery, and EF after surgery states (with CR)	Beta	0.868	3919.66	598.31	0.005
Utility in pre-surgery, surgery, and EF after surgery states (with PR)	Beta	0.754	367.12	119.78	0.019
Utility in pre-surgery, surgery, and EF after surgery states (with SD)	Beta	0.754	367.12	119.78	0.019
Utility in pre-surgery, surgery, and EF after surgery states (with PD, in RCC and pNET cohorts)	Beta	0.665	115.48	58.17	0.036
Utility in pre-surgery, surgery, and EF after surgery states (with PD, in CNS Hb cohort)	Beta	0.550	213.20	174.44	0.025
Utility of pre-progression metastatic disease	Beta	0.525	2.59	2.34	0.205
Utility of post-progression metastatic disease	Beta	0.412	2.83	4.03	0.176

B 20. Priority. On page 201 of the CS the company states that "resultant weighted averages of the CR, PR, SD, and PD utilities (shown in Table 77) were used in all non-metastatic health states, rather than just the pre-surgery state, as patients can continue to receive belzutifan and chieve/maintain CR, PR, or SD following surgery." Please explain in detail when a patient is considered to stop treatment with belzutifan (all possible causes) and how this is implemented in the model.

MSD response:

As per the MK-6482-004 trial protocol, treatment was continued until unacceptable treatment-related toxicity or unequivocal disease progression. In the economic model, ToT was modelled by fitting a Gompertz curve to patient-level data on time-to-treatment discontinuation in the MK06482-004 trial. This approach therefore accounts for any treatment discontinuation during the trial period. Overall response rate and time on treatment are not linked in the economic model.

B 21. Priority. Please answer the following HRQOL-related questions:

a) **On page 198 of the CS, it is mentioned that the "licensed population has more severe disease (and hence would be expected to have worse utility scores) than the population informing the utility data in the economic analysis". Please clarify whether this refers to all health states, adverse event disutility, etc. and explain why. In any case, please conduct scenario analyses to show the impact of using other utility values (which should be properly justified) on the model results.**

b) **On page 198 of the CS, it is also mentioned that "*the effect of belzutifan on HRQoL is underestimated in the economic analysis*". Please clarify why the company thinks this is the case, since the overall effect on HRQOL may depend on other assumptions as well.**

c) **Please justify why in the model an HRQOL benefit has been included since the first model cycle. Initially, it would be expected HRQoL (and other outcomes such as risk of surgery) to be equal in both arms until belzutifan starts to show an effect, which is unlikely to be immediate (seeing for example time to response observed in the MK trial). If**

initially HRQoL (and other outcomes such as the risk of surgery) is assumed to be better in the belzutifan arm compared to SOC, that would indicate that patients in SOC are more severe and therefore not equal to those in belzutifan arm.

d) **In line with the previous question, on page 199 of the CS, it is mentioned that for "the pre-surgery, surgery, and event-free after surgery health states, a better response is associated with a higher utility value, as a better response avoids the complications associated with tumour growth and the greater risk of metastases resulting from disease progression, which would reduce patients' HRQoL". Please clarify whether this rationale could be applied to other model inputs such as transition probabilities (linking response and transition probabilities in a similar was as it was linked to utilities), and indicate if this has been included in the model and how.**

e) **Also, on page 199 of the CS, it is mentioned that the "economic analysis therefore uses response-adjusted utility values for each primary tumour site population in the pre-surgery, surgery, and event-free after surgery health states". Please clarify how this is applicable to the SOC arm, in which there is no treatment and, therefore, it is unclear how response can be measured.**

f) **Please clarify why there is "*high potential for misclassification amongst the PR and SD categories based on patient responses*" and why it was decided to pool values "*across the PR and SD categories*" (page 200 of the CS). Please provide the utility values before pooling.**

g) **On page 200 of the CS, it is mentioned that "Because patients with CNS Hb in the VHL RW QoL Disease Burden Study were not selected for being unsuitable or undesirable for localised procedures, the utility value estimated for VHL CNs Hb patients in this study was not considered representative of the population eligible to receive belzutifan per the MHRA label". Please clarify why this is not applicable to all subgroups included in the economic analyses since being**

unsuitable or undesirable for localised procedures is not exclusive for the CNS Hb subgroup.

h) **Please also clarify why "Patients in this trial [KEYNOTE-564] were considered representative of VHL patients with the most favourable prognosis and HRQoL" (page 201 of the CS).**

i) **In the VHL RW QoL Disease Burden Study spontaneous reduction in tumour was observed. This is in line with the expectation of the clinicians of being highly unlikely since it was observed in a small proportion. Please justify why this was not included in the model and instead all patients in SOC were assumed to have 0% chance to achieve CR/PR.**

j) **Table 76 in the CS shows the distribution of objective response level used to calculate utility values in the pre-surgery, surgery, and event-free after surgery states, which are further reported in Table 77. These weighted averages were used in all non-metastatic health states since the first model cycle. It is unclear whether this approach is correct, as suggested in sub-question c) above. In particular, it is questionable whether the timing of response and progression should be taken into account. The EAG considers that at baseline, because treatment has not started yet, such distribution and therefore, the utilities, should be the same in both arms. It seems irrational to model response right from the first cycle. Please justify why the model implements a QOL benefit right from the start based on response.**

k) **Furthermore, it also seems irrational to assume that 23% of patients in SOC have progressed disease at baseline (compared to 0% in belzutifan). That would imply that patients are not equal in both arms and indeed more severe in SOC. Please justify why the model implements a QOL benefit right from the start based on progression status.**

l) **Also, please clarify whether surgery would bring some sort of benefit to patients as opposed to not receiving surgery. If SOC patients get**

surgery right at the beginning, would it be expected that for some time these patients would have a better HRQOL than at baseline or than those patients in belzutifan who did not get surgery and the drug has not responded yet?

MSD response:

a) More severe disease manifestations in the assessed VHL population at baseline would plausibly be reflected in lower utility values in the pre-surgery health state, given that once surgery occurs, its impact on HRQoL will supersede the baseline utility values. This highlights how consequential a surgery event is for a patient for whom surgery is unsuitable or undesirable. Similarly, in the event patients develop metastatic disease, this is likely to lead to significant impacts on HRQoL. Of note, the additional impact of the presence of severe VHL disease manifestations combined with metastatic disease has not explicitly been captured in the economic model.

There are no alternative data sources to source utility values in VHL disease which could provide plausible alternatives to the values included in the economic analysis. Nevertheless, the uncertainty around utility and disutility values is explored in the sensitivity and scenario analyses presented in Appendix J1.4. For ease, the results are reproduced belowTable 13. Despite the fact that utility values in the non-metastatic health states are one of the more sensitive parameters of those varied in the DSA, their parameter uncertainty does not cause large variations in the ICER.

Table 13 Deterministic sensitivity analyses and scenario analyses of utility and disutility values

	ICER vs. comparator: Belzutifan vs. SOC *(£/QALY)*	
	Low input value	*High input value*
VHL-RCC cohort base case	**73,095**	
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	79,887	67,991
Utility in pre-progression metastatic disease state +/- 10%	73,172	73,020
Utility in post-progression metastatic disease state +/- 10%	73,245	72,950
Disutilities of short-term complications +/- 10%	73,073	73,117
Disutilities of long-term complications +/- 10%	75,237	72,016
Disutility from AEs +/- 10%	73,093	73,097

Assume same utility for CR as PR/SD	73,710	
Apply caregiver disutility	69,940	
Do not apply age-adjusted disutility	70,192	
Do not apply AE-related disutility	73,074	
VHL-CNS Hb cohort base case	**56,933**	
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	65,329	51,322
Utility in pre-progression metastatic disease state +/- 10%	57,108	56,758
Utility in post-progression metastatic disease state +/- 10%	57,319	56,546
Disutilities of short-term complications +/- 10%	56,911	56,955
Disutilities of long-term complications +/- 10%	56,300	58,854
Disutility from AEs +/- 10%	56,932	56,935
Assume same utility for CR as PR/SD	57,392	
Apply caregiver disutility	53,496	
Do not apply age-adjusted disutility	54,042	
Do not apply AE-related disutility	56,921	
VHL-pNET cohort base case	**77,649**	
Utility in pre-surgery, surgery, and event-free after surgery states +/- 10%	84,036	72,203
Utility in pre-progression metastatic disease state +/- 10%	77,562	77,745
Utility in post-progression metastatic disease state +/- 10%	77,461	77,890
Disutilities of short-term complications +/- 10%	77,651	77,646
Disutilities of long-term complications +/- 10%	80,328	75,340
Disutility from AEs +/- 10%	77,646	77,651
Assume same utility for CR as PR/SD	81,086	
Apply caregiver disutility	72,876	
Do not apply age-adjusted disutility	75,002	
Do not apply AE-related disutility	77,625	

b) Overall response rates are used to produce a weighted average utility value for the non-metastatic health states for both the belzutifan and SOC arms. Overall response rates do not necessarily capture all dimensions of the EQ-5D (mobility, usual activities, self-care, pain and discomfort, and anxiety and depression). Therefore, the benefit of belzutifan in these dimensions is likely underestimated, more specifically in areas where the overall response rate is not an indicator of

overall utility as captured by the EQ-5D. Furthermore, belzutifan will work on multiple tumour sites simultaneously in patients that have multiple tumours; however, response is solely linked to the primary tumour. Hence, the additional utility benefit of belzutifan working on multiple tumour sites simultaneously is not captured.

c) In the base-case analysis, utility in the non-metastatic health states is linked to patients' distribution across different levels of tumour response. For simplicity, the model applies fixed proportions of patients at each response level under belzutifan and SOC (with the exception that, after discontinuing belzutifan, the proportions of patients at each response level are assumed to eventually converge to those in the SOC arm). Modelling time-varying proportions of patients at each response level during belzutifan treatment would have added significant structural complexity, and the added precision from such an approach was expected to have limited impact on the model results.

Nevertheless, we understand the concern and have updated the Excel model with two new scenario analysis options to explore the impact of this limitation:

Scenario 1: Utility linked to proportion of all patient-assessments at each response level in MK-6482-004

In the base-case analysis, patients' distribution across response levels during belzutifan treatment is based on best overall response in MK-6482-004. A limitation of this approach is that is does not adjust for the time required to achieve response. To address this limitation, the model now includes an alternative option to calculate patients' distribution across response levels as the proportions of all patient-assessments of response that had a CR, PR, SD, PD, or not evaluable (NE) result in MK-6482-004. Because response assessments in MK-6482-004 occurred at regular intervals of approximately every 12 weeks, this approach closely approximates the total proportion of patient-time spent at each response level, thereby accounting for time to response and duration of response. (As shown in the *Raw_Response scenario…* sheets of the model, the proportion of patients at each response level under this new scenario was calculated based on swimmer plots of patients' response trajectories over time in MK-6482-004.)

Results from scenario 1 are presented in Table 14 below alongside the base case. This scenario can be replicated in the revised Excel model by selecting "Proportion of all patient-assessments at each response level in MK-6482-004" from the first dropdown menu on the *Utility* sheet.

Table 14 Scenario analysis results of utility linked to proportion of all patient-assessments at each response level

Population	**Base case:** *Utility linked to best overall response*			**Scenario 1:** *Utility linked to % of patient-assessments at each response level*		
	Inc. costs	**Inc. QALYs**	**ICER (£/QALY)**	**Inc. Costs**	**Inc. QALYs**	**ICER (£/QALY)**
VHL-associated RCC	[redacted]	[redacted]	73,095	[redacted]	[redacted]	73,673
VHL-associated CNS Hb	[redacted]	[redacted]	56,933	[redacted]	[redacted]	57,343
VHL-associated pNET	[redacted]	[redacted]	77,649	[redacted]	[redacted]	79,290

Scenario 2: No differentiation of utility by response

As an additional conservative scenario, the revised model provides the option to set the utility for all response levels equal to the same value, i.e., the average EQ-5D-5L utility (scored using the crosswalk onto the UK EQ-5D-3L value set) among patients without metastatic disease in the VHL RW QoL Disease Burden Study (2022), irrespective of their self-reported response status.

Results from scenario 2 are presented below alongside the base case results in Table 15 below. This scenario can be replicated in the revised Excel model by selecting the option "VHL RW QoL Disease Burden Study (2022), without differentiation by response" from the 2nd (row 15), 3rd (row 17), 4th (row 19), and 5th (row 21) dropdown menus in the "Utility" tab.

Table 15 Scenario analysis results of no differentiation of utility by response

Population	**Base case:** *Utility linked to best overall response*			**Scenario 2:** *No differentiation of utility by response*		
	Inc. Costs	**Inc. QALYs**	**ICER (£/QALY)**	**Inc. Costs**	**Inc. QALYs**	**ICER (£/QALY)**
VHL-associated RCC	[redacted]	[redacted]	73,095	[redacted]	[redacted]	77,081

VHL-associated CNS Hb	[redacted]	[redacted]	56,933	[redacted]	[redacted]	57,893
VHL-associated pNET	[redacted]	[redacted]	77,649	[redacted]	[redacted]	84,949

d) With sufficient data, response rates could plausibly be linked to transition probabilities (e.g. pre-surgery → surgery); however there is a lack of data on response status from the VHL Natural History Study informing disease progression for SOC patients. Furthermore, given the low number of surgery events in the trial, it would not have been feasible to differentiate the pre-surgery → surgery transition at different response levels. Therefore, this linkage was not included in the economic model.

e) Response is measurable (and is relevant) in SOC patients. Response status is measured via a scan to assess tumour size and growth and can be done irrespective of whether a patient is receiving treatment or not; patients would undergo frequent monitoring regardless of treatment received. Therefore, it is relevant to include this for the SOC arm in the model. SOC cannot plausibly be associated with CR, however, through active surveillance and symptom management it can plausibly be associated with SD (and PD). In the absence of an alternative data source to estimate response rates (and associated utility values) in the SOC arm, the self-reported overall response rates from the VHL RW QoL Disease Burden Study were used. Only the response rates from those not receiving VHL treatment were used in the economic analysis to reflect the fact that patients in the SOC arm do not receive systemic treatment.

f) As the VHL RW QoL Disease Burden Study provides data self-reported by patients, there is potential for patients to confuse SD for PR in the recollection of their current disease status. Unfortunately it was not possible to validate these self-reports with clinicians nor their case notes. Furthermore, a spontaneous reduction in size of VHL-related tumours is very unlikely in the absence of active VHL treatment. Given the lack of face validity of PR in the absence treatment and the small patient numbers in this category, these patients were reclassified as having achieved SD. Due to time constraints, we are unable to rerun analyses in the VHL RW QoL Disease Burden Study to retrieve utility values before pooling the PR and SD categories. Furthermore, this is expected to have limited impact on the cost-

effectiveness results as shown by the scenario presented above in Table 13 (and also in Appendix J1.4) where the same utility value is assumed for CR as PR/SD producing a change in ICER of <£1,000 in the VHL-RCC and VHL-CNS Hb cohorts and <£4,000 in the VHL-pNET cohort.

g) Patients in the VHL-CNS Hb cohort for whom localised procedures are unsuitable or undesirable are different from the other two cohorts. Patients who need surgery of RCC or pNET tumours but for whom surgery is unsuitable or undesirable may not be immediately distinguishable from a patient who is suitable for surgery. However, a CNS Hb patient who is unsuitable for surgery and *requires therapy* has an *urgent* need for surgery because they are experiencing neurological symptoms that are by definition worse than the risks associated with surgery. These symptoms can manifest in a myriad of functions in a way that is wholly different from an RCC or pNET patient who may have lower or less urgent symptom burden. In RCC and pNET patients, there is the trade-off between the risks affecting organ function with localised procedures versus the risk of metastatic disease. Whereas in CNS Hb patients, this trade-off is between risks of complications from surgery versus their current experience of debilitating neurological symptoms.

h) Patients within the disease-free state among in the KEYNOTE-564 trial were considered suitably representative of VHL patients who achieved CR status based on the definition of CR according to RECIST v1.1 criteria used in the belzutifan clinical trial (i.e., disappearance of all target lesions, with any pathological target or non-target lymph nodes reduced in short axis to <10 mm). The disease-free (without toxicity) utility was previously estimated for NICE TA830 based on EQ-5D-5L data from KEYNOTE-564 during patient-visits in which patients remained disease free. Hence, these patients were considered suitable representative of VHL patients with CR as explained in the response to B 18. a).

i) The assumption that a spontaneous reduction in tumour size can occur without active treatment lacks face validity. Nevertheless, assuming that CR can occur in the SOC arm using the response rates as reported from the VHL RW QoL Disease Burden Study (i.e. CR n=1; PR n=4) has minimal impact on the ICER as shown in Table 16.

Table 16 Scenario analyses assuming overall response rates as the VHL RW QoL Disease Burden Study without reclassification to SD

	ICER (£/QALY)		
	VHL-RCC cohort	*VHL-CNS Hb cohort*	*VHL-pNET cohort*
Base case	**73,095**	**56,933**	**77,649**
Allowing CR/PR in the SOC arm using ORR reported from the VHL RW QoL disease burden study without reclassification	73,099	56,892	77,665

j) Please see response to part (c) above.

k) For simplicity, the model applies fixed proportions of patients at each response level under belzutifan and SOC as stated in response to part (c) above. Patients who 'require therapy' by definition have a form of progressive disease that has passed the point at which active surveillance is manageable and now requires intervention. Nevertheless, as described in part (c) there are 2 additional scenarios in the model that explore utility linked to proportion of all patient-assessments at each response level in the belzutifan arm and no differentiation of utility by response both of which demonstrate a minor impact on results.

l) These patients who receive surgery as a 'last-resort' would be at high risk of significant complications given this treatment modality is not suitable for them. Therefore, the HRQoL benefit is not realised in the same way as with belzutifan-treated patients. Given the risks of surgical complications and their associated disutilities, it would be expected that a change in the baseline utility due to a benefit from surgery would have minimal impact on the ICER.

Costs and resource use

B 22. Priority. The average cost of Belzutifan treatment is indicated to be [redacted]. Using the list price of belzutifan is £11,936.70 for a 90 tablet pack of Belzutifan 40mg, an average time on treatment (ToT) of [redacted], a mean relative dose intensity (RDI) of [redacted]%, and 3 tablets per day schedule, the average cost is higher than that indicated by the company. Please provide the calculations of the average cost of treatment.

MSD response:

The average cost of belzutifan treatment is calculated assuming a 4-week interval between each dispensing of a new pack of belzutifan. This can be calculated by using the drug cost per pharmacy dispensing in the *Drug & Admin Costs* sheet cell *P65* in the model which accounts for RDI and multiplying by 13 (52 weeks divided by 4-weekly dispensing of each pack) and by median ToT. The belzutifan drug acquisition costs reported as part of the disaggregated results reflect treatment discontinuations based on the estimated ToT curve and it is therefore logical that drug acquisition costs would likely be different than when the median ToT is used (unless an exponential model was used to extrapolate ToT).

B 23. Priority. On page 212 of the CS it is mentioned that "a subset of patients with advanced RCC or pNET are assumed to receive no active metastatic disease treatment, as not all patients with metastatic disease receive active treatment." Please provide further details on the numbers of metastatic patients not receiving active treatment, the source used to inform these parameters, where these parameters can be found in the electronic model, and on whether any validation steps were taken for these parameters.

MSD response:

Market shares were estimated based on the subsequent treatment market shares used in the NICE appraisal of pembrolizumab as an adjuvant treatment of RCC post-nephrectomy (TA830) in the VHL-RCC cohort, and European Society for Medical Oncology (ESMO) clinical practice guidelines for gastroenteropancreatic neuroendocrine neoplasms and input from clinical experts in the VHL-pNET cohort. The market shares of metastatic treatments are presented in Table 89 of the CS. For ease, they are also presented in Table 17 below. Market shares were based on data collected in November 2021. Patients are eligible to receive no active treatment only after receiving a *first*-line therapy in the metastatic RCC setting, consistent with assumptions in NICE TA830. These market shares can be found in the *Market Shares* sheet of the economic model.

Table 17 Metastatic treatment market shares

Metastatic therapy regimens	Market share by treatment arm

	Belzutifan	SOC
First-line metastatic therapy (metastatic RCC)		
Sunitinib	30.0%	30.0%
Tivozanib	14.0%	14.0%
Pazopanib	29.0%	29.0%
Cabozantinib	13.0%	13.0%
Nivolumab / ipilimumab	14.0%	14.0%
Avelumab / axitinib	0.0%	0.0%
No active treatment	0.0%	0.0%
Second-line metastatic therapy (metastatic RCC)		
Nivolumab	0.0%	0.0%
Axitinib	7.0%	7.0%
Cabozantinib	32.0%	32.0%
Lenvatinib / everolimus	0.0%	0.0%
Pazopanib	4.0%	4.0%
Sunitinib	0.0%	0.0%
Tivozanib	0.0%	0.0%
Everolimus	7.0%	7.0%
Sorafenib	0.0%	0.0%
Cytokines (interferon)	0.0%	0.0%
No active treatment	50.0%	50.0%
First-line metastatic therapy (metastatic pNET)		
Streptozocin / 5-fluorouracil	0.0%	0.0%
Streptozocin / doxorubicin	0.0%	0.0%
Temozolomide / capecitabine	0.0%	0.0%
Everolimus	0.0%	0.0%
Sunitinib	0.0%	0.0%
Interferon a2B	0.0%	0.0%
Lanreotide	50.0%	50.0%
Octreotide	50.0%	50.0%
No active treatment	0.0%	0.0%
Second-line metastatic therapy (metastatic pNET)		
Cisplatin / etoposide	0.0%	0.0%
Everolimus	25.0%	25.0%
FOLFIRI	0.0%	0.0%
FOLFOX	0.0%	0.0%
Streptozocin / 5-fluorouracil	25.0%	25.0%
Streptozocin / doxorubicin	25.0%	25.0%
Sunitinib	25.0%	25.0%
Temozolomide / capecitabine	0.0%	0.0%
Interferon a2B	0.0%	0.0%
Lanreotide / everolimus	0.0%	0.0%
Octreotide / everolimus	0.0%	0.0%
Lanreotide	0.0%	0.0%
Octreotide	0.0%	0.0%
No active treatment	0.0%	0.0%

B 24. Priority. Page 219 of the CS states that “discontinuation rates for first-line metastatic treatments for advanced RCC and advanced pNET are approximated from exponential rates of PFS failure”. Please

provide further details on the estimation of the discontinuation rates per treatment for the first line metastatic treatments and indicate where these parameters can be changed in the model.

MSD response:

Metastatic treatment discontinuation rates are approximated using the exponential rates of PFS failure. Two separate network meta-analyses (NMA) were used to estimate mean PFS for each metastatic treatment regimen versus sunitinib (in advanced RCC) and versus no active treatment (for pNET). The exponential PFS rates and HRs are presented in Tables 65-68 of the CS. Discontinuation rates can therefore be estimated by assuming treatment until progression and/or using the maximum treatment duration according to dosing schedules recommended by NICE where applicable. Estimated PFS of included metastatic disease therapies can be modified in the *Effectiveness* sheet of the economic model.

B 25. Priority. The base case cost-effectiveness analysis considers social care costs associated with stroke and neurological dysfunction as a complication of surgery associated with VHL. Additionally, for PD patients in the VHL-CNS Hb cohort, social care costs associated with disease management have also been included. Please indicate how the parameters in Table 95 of the CS are incorporated in the cost effectiveness calculations (in the model) for each of the three cohorts and provide examples of previous appraisals in which similar social costs have been considered in base case calculations.

MSD response:

The following social care costs can be found in the economic model:

- Stroke: In the *Surgery* sheet in the respective cost cells *I52* and *I100*. The estimated social care cost was £2,833 in 2015 which has been inflated to 2021 prices.

- Neurological complications: In the *Surgery* sheet in the respective cost cells *I102*. The study providing this cost reports a per-patient cost of brain disorders of €3126 at 2013 prices. A proportion of this (26.8%) was attributed to direct

non-medical costs (i.e. social services). This was then converted into 2013 GBP (€1 = £0.8492) and then inflated to 2021 prices.

- Disease management of patients with PD in the VHL-CNS Hb cohort: In the *HCRU* sheet in row 26 & 27. A three-monthly cost of £250 (in 2017 prices) was identified, which was then converted to a weekly cost and inflated to 2021 prices.

Examples of appraisals where costs of rehabilitation relating to stroke have been reflected (which encompasses social care costs) include TA275 and TA256. Costs of social care are frequently included in multiple sclerosis TAs; for example, TA767 which includes costs of community services (e.g. nurse visit, home helper) and major investments (e.g. purchase of a wheelchair, transform the house or car). As mentioned in the CS, multiple sclerosis is a proxy condition for VHL-CNS Hb PD given the extent of disease burden and the fact that its symptom manifestation includes neurological complications.

All NICE Tas are conducted from the perspective of the NHS and PSS. Although many TAs do not include social care costs, they should be included where they are relevant in line with the NICE methods guide, and hence are included in this TA.

Cost-effectiveness results

B 26. Priority. Please provide a plot of the Markov traces for the base-case results. Include this in the model too and indicate where it can be found.

MSD response:

In the revised Excel model, Markov trace graphs have now been added to the top of each Markov trace tab (sheets named "*Trace_TxReg…*").

B 27. Priority. Please present the cost-effectiveness results using the appropriate QALY weighting for each subgroup (note the weighting may differ per subgroup).

The QALY shortfall analysis respective to the STA process is reported in B.3.6 Severity section of the CS. The table with a summary of the analysis and QALY weights are reproduced in Table 18 below.

Table 18 Summary of QALY shortfall analysis

Cohort	Expected total QALYs for the general population	Total expected QALYs for people with VHL on current SOC	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
VHL- GB marketing authorisation population (weighted cohort)	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated RCC	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated CNS Hb	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated pNET	18.15	[redacted]	[redacted]	[redacted]	1.2

QALY: quality adjusted life year; SOC: standard of care; VHL: Von Hippel Lindau

B 28. Priority. Please discuss the cost-effectiveness results (base-case and uncertainty) in the context of the appropriate UK cost-effectiveness thresholds.

MSD response:

VHL is a complex disease to model. The trial data are very typical of an ultra-rare disease: a single-arm study with very few patients. It is clear from the clinical results that belzutifan is transformative. This has been recognised by both the FDA and the MHRA in the granting of the MA and the 'expanded indication'. However, this has made what was already a complex modelling challenge even more challenging.

As expressed in the CS, MSD is disappointed that NICE did not route this into the HST process. HST is more able to consider this type of complexity. However, MSD chose to move forward within the STA process as we needed to prioritise the patients, who are waiting for this treatment. The result is this TA falls into the well-recognised 'chasm' between STA thresholds and HST thresholds.

MSD is confident that belzutifan is a transformative treatment option for patients with VHL-associated tumours. We are also confident it offers good value for money to the NHS; the decision risk is small due to the very small patient numbers. Applying large population statistical and economic methodologies to this dataset, as is common in the STA process, is the wrong way of measuring its value.

B 29. Priority. In the electronic model when changing the parameters defining the time points treatment waning is initiated and completed on the Tx Duration sheet, the results for the VHL-CNS Hb cohort remain unaltered. On the other hand, in the scenario analyses presented by the company, the ICER in the VHL-CNS Hb cohort is lower compared to the base case results when no waning is assumed. Please explain this discrepancy in the results of the VHL-CNS Hb cohort compared to the other two cohorts or correct the error if there is an error in the model.

MSD response:

We do not find an error in the model. To alter the treatment effect waning time points, these must be done for each cohort. For the CNS Hb cohort, these should be changed in cells *I44* and *I45* of the *Tx Duration* sheet. When changes are made to these cells, the results do change. Please see screenshots from the Excel model provided below.

Base case assumptions & results

Select assumption regarding the relationship between effectiveness and ToT of belzutifan:	Assume treatment effect waning: Model efficacy and ToT together with on/off treatment states			
If applicable, specify timeframe to apply treatment effect waning:	*VHL-RCC cohort*	*VHL-CNS Hb cohort*	*VHL-pNET cohort*	*Notes:*
Time point to initiate treatment waning (years):	3.84	3.84	3.84	[1]
Time point to complete treatment waning (years):	6.55	6.55	6.55	[2]

Altered assumptions & results

Select assumption regarding the relationship between effectiveness and ToT of belzutifan:	Assume treatment effect waning: Model efficacy and ToT together with on/off treatment states			
If applicable, specify timeframe to apply treatment effect waning:	*VHL-RCC cohort*	*VHL-CNS Hb cohort*	*VHL-pNET cohort*	*Notes:*
Time point to initiate treatment waning (years):	7.68	7.68	7.68	[1]
Time point to complete treatment waning (years):	10.39	10.39	10.39	[2]

B 30. Priority. Tables 103-105 of the CS presenting the scenario analyses show that removing the adjustment parameters used for surgery and metastases rates to account for real-world standard of care increases the base case ICER in the VHL-RCC cohort, while the ICERs decrease in the VHL-CNS Hb and VHL-pNET cohorts after this omission.

a) Please explain the reason behind this discrepancy.

b) Please comment on the rationale behind the use of the Optum study to estimate these adjustment parameters and indicate if any validation exercise has taken place for the inclusion of the adjustment parameters in the calculations.

MSD response:

a) As stated in response to B 5. above, adjustment parameters for surgery and metastases rates were applied to both arms of the model. Adjustments to surgery and metastases rates were conducted for the VHL-RCC cohort whilst adjustments to metastases rates only were made for the VHL-CNS Hb and VHL-pNET cohorts. Surgery rates for the latter two cohorts were modelled based on data from the pre-treatment period of MK-6482-004 trial. We do not have evidence that the patients in the pre-treatment period of study received an elevated standard of care during this

time period. Hence, for the CNS Hb and pNET cohorts the upward adjustment to metastases rates results in a higher ICER (than if the adjustment was removed) as increased metastases means a reduced benefit of belzutifan is realised. For the RCC cohort, the downward adjustment of surgery rates dominates any worsening in the ICER from the upward adjustment of metastases rates resulting in a lower ICER (than if the adjustment was removed) so an increased benefit of belzutifan is realised.

b) Please see the response to B 5.

Validation

B 31. Priority. Please clarify if and how the conceptual model was validated. Please consider discussing here face validity (e.g., if experts considered the model structure appropriate, justify the choice of the health states, etc.) and cross validity (e.g., if this model has been compared to similar models in the literature). Please also clarify if input parameters, other than transition probabilities, were validated and how.

MSD response:

The initial model was conceptualised at the point that the expected indication wording would be for VHL-related RCC tumours only, prior to clinical trial data read out. At that model design stage, a clinical panel with one external expert physician, Dr. Eric Jonasch, MD (Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center), supplemented by an internal therapy area clinician, validated the conceptual model and other aspects of the model development plan. Key topics and feedback from this discussion are summarized in Table 19 below.

Table 19 Feedback received on the conceptual model and inputs

Topic	Summary of feedback
Conceptual model / specific health states	• Experts confirmed that a Markov model structure reasonably reflects the VHL disease trajectory and its impact on costs and health outcomes over time. • They confirmed that tumour reduction surgeries are the most relevant clinical events to represent in the disease process given the substantial morbidities and disabilities

	associated with surgical management of VHL disease (e.g., chronic pain, organ dysfunction, renal failure, neurological changes, etc.). • They confirmed the relevance of the metastatic disease state, while noting the metastases should be relatively infrequent for patients who are appropriately managed by SOC.
Estimation of clinical inputs for SOC	• To inform the set of baseline variables to use for matching in the MAIC, expert input was sought regarding baseline characteristics that are likely to be prognostic of transition probabilities starting from the pre-surgery state, or that may modify the effect of belzutifan on these transition probabilities. • *Baseline risk adjustment through propensity score reweighting* in section B.2.9 of the CS describes the feedback received and the final set of baseline variables used for matching.
Residual efficacy of belzutifan after discontinuation	• Experts were asked whether any residual treatment effect is expected following treatment discontinuation. Both agreed that time to surgery will likely continue to be delayed (relative to SOC) for some time after a patient discontinues belzutifan. This is because, at the time of discontinuing belzutifan, the size of patients' tumours is likely to be smaller than it would have been if they had not been treated with belzutifan up to that point. • Consequently, it will take a longer time for the largest tumour to reach the threshold size that warrants surgery, even if the linear growth rate (mm/year) immediately reverts to pre-treatment levels after discontinuation of belzutifan.
Approach for estimating survival after metastases	• Experts agreed that it would be reasonable to use evidence from first-line drug trials in metastatic RCC to estimate transition probabilities from metastatic disease → death, even though these trials were conducted in a general metastatic RCC population (not specifically among patients with VHL disease who developed metastatic RCC).
VHL-related tumour manifestations that drive costs and/or quality of life impairment	• Experts mentioned that the following VHL-related manifestations are among the most important drivers of quality of life impairment or costs: surgeries (particularly for RCC, CNS Hb, and retinal Hb) and related complications, as well as disability due to CNS Hb lesions. (*Note:* The final model ultimately included a comprehensive set of VHL-related tumour surgeries and related complication risks based on real-world data, and considered the direct impact of tumours on quality of life by linking utility in the non-metastatic health states to levels of tumour response.)

In addition to transition probabilities, other parameter inputs and assumptions were validated by clinicians at various stages of the model development. In particular:

- Surgical complication risks: Risks of specific surgical complications per surgery for each type of surgery were informed by a parallel study within a real-world retrospective claims database (Optum Clinformatics Data Mart, 2000 – 2020). The list of relevant surgical complications for each surgery type, as well as the specific diagnosis and procedure codes used to identify each complication, were confirmed by input from clinical experts at Merck and an academic medical center (Dr. Eric Jonasch from MD Anderson). These were adjusted to align with the UK MHRA label.

- Utility value for VHL-CNS Hb patients with PD: The utility value of CNS Hb patients with PD being particularly worse in patients for whom surgery is unsuitable and the use of a proxy condition to source this value was validated with a UK consultant endocrinologist who runs a VHL MDT.

- Definition of "unsuitable or undesirable for surgery": The notion that patients for whom surgery is unsuitable or undesirable being synonymous with a 'last resort' surgery leading loss of organ function and/or extremely poor outcomes was validated with a UK consultant endocrinologist who runs a VHL MDT.

As previously discussed, following readout of the trial data and belzutifan efficacy for non-RCC tumours, the FDA marketing authorisation expanded the expected indication to include pNETs and CNS Hb. At that point the model was adapted to have an additional two cohorts. MSD acknowledges we have found it challenging to further adapt the model to include the specifics of the MHRA model: 'requires therapy' and 'localised procedures are unsuitable or undesirable'.

Electronic model

B 32. Priority. Page 165 of the CS states that costs of non-primary tumour surgeries, as well as costs and QALY decrements due to non-primary tumour surgery complications, were calculated in each cycle, and were layered (additively) onto the costs and QALYs estimated based on patients' distribution across primary tumour-related Markov health states. Please adapt the model to allow for a distinct estimation

of the costs and QALYs due to non-primary tumour surgeries (not incorporated in the primary tumour-related health states).

MSD response:

While we would like to implement the request, and have certainly considered it ourselves, we unfortunately do not have the appropriate data analysis for *other tumours* (e.g. retinal Hb) to accurately do so. More specifically, we do not have the TTS curves for these in the same way we do for the RCC, CNS Hb and pNET tumour types in both arms of the model. Given the additional model complexity this would introduce, we have applied the transparent and simplifying assumption to apply its incidence as background per cycle.

This is an uncertainty that we believe undervalues belzutifan as what is missing from the model as a result is the 'snowballing effect' due to multiple tumours (i.e. the impact of tumours is more than the sum of their parts). As a company, we are going back to the dataset to see if we can find a way to model these surgeries; however, we do not yet have the data.

B 33. Priority. Please check the model implementation of:

a) **Discounting: the difference between discounted and undiscounted QALYs seems oddly small. In case, this is correct, please explain why this happens.**

b) **Vial sharing: the impact on costs seems oddly small too. In case, this is correct, please explain why this happens.**

c) **Terminal costs: they seem to be always higher for SOC.**

MSD response:

We can confirm there are no implementation errors on any of the above points.

a) Table 20 below presents the difference in discounted and undiscounted QALYs in the disaggregated results in the VHL-RCC cohort, as an example. When discounting is applied, QALYs are primarily accrued in the pre-surgery and event-free after surgery health states and QALY losses occur mainly through surgical complications. Without discounting, the QALY accrual increases in the pre-surgery and event-free

after surgery health states. However, the QALY losses via surgery complications are much greater than when discounting is applied. Therefore, the net change in total QALYs is marginal. This is expected because the surgical complications with the greatest impact are long-term and therefore are highly sensitive to discounting.

Table 20 Undiscounted and discounted disaggregated QALYs in the VHL-RCC cohort

Outcomes	Discounted		Undiscounted	
	Belzutifan	*SOC*	*Belzutifan*	*SOC*
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for other tumours	[redacted]	[redacted]	[redacted]	[redacted]
AE-related disutility	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutlity	[redacted]	[redacted]	[redacted]	[redacted]
Age-related disutility	[redacted]	[redacted]	[redacted]	[redacted]

b) Vial sharing has minimal impact, as this assumption only impacts subsequent treatments for metastatic disease. Belzutifan is a tablet formulation administered orally.

c) Terminal care costs are (marginally) higher for SOC as there are more deaths that occur earlier in the SOC arm; therefore, the terminal care costs of these earlier deaths are less impacted by discounting.

B 34. Priority. Please explain why the impact of age on the ICER is not equal in all three subgroups (it seems to decrease the ICER for the RCC subgroup and increase for the other two) and whether this is in line with expectations.

MSD response:

We can confirm there are no implementation errors here. Changing age can either increase or decrease the ICER in a given population, depending on a variety of

inputs in the model. Transition probabilities are estimated and/or adjusted from various sources which differ between cohorts and therefore the impact of age on the cost-effectiveness results is not expected to necessarily be similar in magnitude or direction in all the cohorts.

B 35. Priority. Worksheet "Specifications":

a) **For all options where a HR approach is applied, please provide evidence that proportional hazards can be assumed.**

b) **For transitions from pre-surgery to metastatic disease or death, please clarify why only the Exponential distribution is possible to select. Please conduct a full survival analysis and include other probability distributions as in case of time to surgery.**

MSD response:

a) In the current CE model, HR approaches were used to estimate transition probabilities in the belzutifan arm that, due to few or no events, could not be directly estimated through parametric modelling. Table 21 below summarizes the specific HR assumptions used for these transition probabilities and their corresponding rationale.

Regarding the assumption of *proportional* hazards (i.e., time-constant HRs) for these transition probabilities, this assumption follows from our choice of exponential distributions to model pre-surgery → surgery in the base case. In the Excel model, if the distribution for pre-surgery → surgery transitions are changed from exponential to a non-exponential distribution, the HR of pre-surgery → surgery for belzutifan vs. SOC is re-calculated accordingly and will be time-varying. In this case, all other transition probabilities that depend on the HR of pre-surgery → surgery will also reflect time-varying HRs. For example, when gamma distributions are selected for pre-surgery → surgery in each arm of the VHL-RCC cohort, the HR of pre-surgery → surgery for belzutifan vs. SOC in this cohort will be time-varying (see column BS of the *TP_TxReg1_RCC* sheet); consequently, other transition probabilities in the belzutifan arm (e.g., pre-surgery → metastatic disease) that depend on the HR of pre-surgery → surgery for belzutifan vs. SOC will be estimated using time-varying HRs (e.g., see column AY of the *TP_TxReg1_RCC* sheet).

One additional remark: for all transitions from the pre-surgery and event-free after surgery states in all three populations, patients in the belzutifan arm who discontinue belzutifan are assumed to eventually face the same transition probabilities as in the SOC arm. This treatment waning assumption effectively relaxes the proportional hazards assumption, i.e., because the proportional hazards assumption no longer holds once a patient has discontinued belzutifan and is subject to treatment effect waning.

Table 21 Overview of transition probabilities from the pre-surgery state that are estimated using a HR approach:

Transition	Description of HR approach	Rationale
Pre-surgery → surgery for belzutifan *(VHL-pNET cohort)*	As of the 1 Apr 2022 data cutoff date, there were no pNET surgeries among the 22 patients with pNET at baseline. Rather than assume zero risk of pNET surgeries while patients are being treated with belzutifan, the HR of pre-surgery → surgery with belzutifan (vs. SOC) in the VHL-pNET population was assumed equal to the HR of pre-surgery → surgery with belzutifan (vs. SOC) in the VHL-RCC population multiplied by $(1\text{-}ORR_{pNET})/(1\text{-}ORR_{RCC})$, where ORR_{pNET} and ORR_{RCC} are the objective response rates of belzutifan with respect to pNET and RCC tumours, respectively.	This assumption was considered reasonable and appropriate, as it accounts for the higher objective response rate of belzutifan with respect to pNET tumours (91%) (compared to 64% for RCC tumours) in the MK-6482-004 trial. Because both ORR and the need for surgery are determined by tumor size, the higher ORR with respect to pNET should translate to a greater percentage reduction in the hazard rate of pre-surgery → surgery.
Pre-surgery → metastatic disease for belzutifan *(all cohorts)*	No pre-surgery → metastases transitions have been observed for belzutifan in the MK-6482-004 trial as of the 1 April 2022 data cutoff date. In order to apply some nominally positive risk of pre-surgery → metastatic disease (rather than assume zero risk of metastases) while patients are treated with belzutifan, the HR of pre-surgery → metastatic disease with belzutifan (vs. SOC) was assumed equal to the HR of pre-surgery → first surgery with belzutifan (vs. SOC) in each target population.	This assumption was considered clinically appropriate because belzutifan would reduce the risks of surgeries and of metastatic disease through the same mechanism (i.e., by decreasing the size of or halting the growth of tumours); in RCC, for example, there is a well-established link between size of the largest renal tumour and risk of metastases (Duffey et al. 2004), which led to the recommendation that patients undergo surgery once their largest renal tumour reaches 3 cm.
Pre-surgery → death for belzutifan *(all cohorts)*	For transitions from pre-surgery to death, the treatment effect of belzutifan vs. SOC (as estimated for transitions to surgery in the VHL-CNS Hb cohort) was assumed to	According to neurosurgeons who were consulted during the model development, CNS Hb is a major cause of death in VHL disease due to growth of Hb or bleeding of Hb

	only reduce the risk of deaths that are attributable to CNS Hb progression in each cohort. Deaths from the pre-surgery state due to other causes were assumed equal between the belzutifan and SOC arms.	resulting in severe neurological disability and death. They noted that, because CNS Hb-related mortality is usually due to the mass effect, reducing tumour size or halting tumour growth would reduce the risk of death due to CNS Hb. In MK-6482-004, belzutifan was effective in reducing the size and growth rate of CNS Hb, and is thus expected to reduce the risk of death due to CNS Hb progression: Among patients in MK-6482-004 who had CNS Hb at the baseline visit (N=50), the ORR was 44.0% and the disease control rate was 90.0% with respect to CNS Hb. The overall median LGR after treatment initiation for participants with CNS Hb was negative (-1.64 mm/year), indicating inhibition of tumour growth over the course of the study. In this population, the weekly exponential hazard rate of pre-surgery → surgery following belzutifan initiation was 0.00010 (vs. 0.00202 during the pre-treatment period of MK-6482-004), implying a 95% reduction in the hazards of this transition.

b) Details of the full survival analysis and why the Exponential distribution was chosen for these transitions are included in the response to B 6. above.

B 36. Priority. Worksheet "Effectiveness":

a) Please explain why the overall survival curves for SOC do not start at 1 and there is a clear separation from the belzutifan curves right from the beginning. As mentioned in previous questions, it seems counterintuitive to assume such a difference when response to treatment is not immediate.

b) The anticipated survival benefit of belzutifan treatment compared to SoC is reflected in the transition probabilities from pre-surgery and event-free-after surgery→death which account for the belzutifan-attributable reduction in the rate of death attributable to secondary CNS

Hb progression. Omitting this survival benefit due to prevention of CNS Hb-related deaths and neglecting the perioperative mortality risk (1st, 2nd, and 3rd surgery) would be expected to allow OS curves on the effectiveness worksheet of the electronic model to overlap between the two arms. However, this is not currently the case. Please confirm this is correct and indicate which are exactly the changes that need to be made in the model to produce similar overall survival curves between the treatment arms (you can only indicate for the VHL-RCC cohort).

MSD response:

a) We would like to highlight that it would be inappropriate for these OS curves to be used for validation purposes. As stated in the *Validation of long-term extrapolation of OS* in section B.3.3 of the CS, there were no available data sources to validate the modelling of OS in the target population stipulated by the MHRA label. To obtain interpretable comparisons between modelled OS and the RWE available, adjustments to align with the GB label population or account for real-world SOC were removed. Therefore, the OS curves to be used for validation can be found in the *Effectiveness Validations_OS* sheet in the economic model. The immediate drop in the SOC arm in the OS curves in question is due to the perioperative mortality associated with surgery which the majority of patients in the SOC arm undergo in the first cycle reflecting the time period of the decision point.

b) As stated above, the OS curves included in this sheet should not be used for OS validation. We can confirm that perioperative mortality and mortality due CNS Hb progression are factors that contribute to overall survival; however, another mechanism of improved survival associated with belzutifan is through the avoidance of metastases since patients face a high mortality upon developing metastases (from non-CNS Hb tumours). This likely accounts for the remaining difference in OS. In the economic model, setting metastases rates equal in both arms is not feasible since the pre-surgery → metastatic disease transition for the belzutifan arm is estimated using an HR approach. (*Note:* the event-free after surgery → metastatic disease transition is also derived from the pre-surgery → metastatic disease transition).

Section C: Textual clarification and additional points

C 1. Input from “clinical experts” is mentioned several times throughout Document B but is usually not referenced. Reference 3 of Document B appears to relevant, but we could not identify this within the papers folder.

a) Please provide a citation each time that input from clinical experts is mentioned in the documentation.

b) Please provide the relevant papers for the above citations.

MSD response:

a) We have based some assumptions on information that clinicians gave us as part of broad discussions about VHL and belzutifan. We have not (yet) been able to go back and re-validate every assumption used in the model based on this information/advice given to us by clinicians. Details of the validation process are reported in B.3.14 Validation section of the CS.

b) We are unable to provide documentation of this as it contains confidential information and we have not sought permission from participating experts.

C 2. Appendix C contains two embedded files. Please provide both as separate files.

MSD response:

These have been provided as separate files along with these clarifications. Please note that the embedded files can also be accessed by double-clicking on them in Appendix C of the company submission.

C 3. On page 154 of the CS, the second part of the formula includes two times the parameter ‘(% of pre-surgery → death transitions attributable to CNS Hb progression in cohort)’. Please clarify if this is an error and indicate where exactly in the model can one find the respective calculations.

MSD response:

We would like to apologise for this typographical error. The calculation should read:

(cause-specific hazard rate of pre-surgery → death under SOC) × [100% – (% of pre-surgery → death transitions attributable to CNS Hb progression in cohort)]

+

(cause-specific hazard rate of pre-surgery → death under SOC) × (% of pre-surgery → death transitions attributable to CNS Hb progression in cohort) × (hazard ratio of pre-surgery → surgery with belzutifan vs. SOC in the CNS Hb cohort)

The respective calculations in the economic model can be found in the *TP_TxReg1_CNSHb* sheet in column AZ.

C 4. On page 193 of the CS, it is stated that to account for waning of the treatment effect of belzutifan, the clinical efficacy parameters of patients in the belzutifan arm were assumed to gradually converge over time towards those of SOC. Please confirm if with ‘gradually’ a linear decline is assumed in the model.

MSD response:

Yes, this is correct. We can confirm that a linear decline assumed in the model.

References

1.Jonasch E, Donskov F, Iliopoulos O, Rathmell WK, Narayan VK, Maughan BL, et al. Belzutifan for Renal Cell Carcinoma in von Hippel-Lindau Disease. N Engl J Med. 2021;385(22):2036-46.
2.Remiro-Azócar A, Heath A, Baio G. Methods for population adjustment with limited access to individual patient data: A review and simulation study. Research Synthesis Methods. 2021;12(6):750-75.
3.Cheng D, Tchetgen ET, Signorovitch J. On the double-robustness and semiparametric efficiency of matching-adjusted indirect comparisons. Res Synth Methods. 2022.
4.Ishak KJ, Proskorovsky I, Benedict A. Simulation and matching-based approaches for indirect comparison of treatments. Pharmacoeconomics. 2015;33(6):537-49.
5.Lammens CR, Bleiker EM, Verhoef S, Hes FJ, Ausems MG, Majoor-Krakauer D, et al. Psychosocial impact of Von Hippel-Lindau disease: levels and sources of distress. Clin Genet. 2010;77(5):483-91.
6.Fraser L, Watts S, Cargill A, Sutton S, Hodgson S. Study comparing two types of screening provision for people with von Hippel-Lindau disease. Fam Cancer. 2007;6(1):103-11.
7.Alliance V. VHLA suggested active surveillance guidelines: VHL Alliance (VHLA); 2020 [Available from: https://www.vhl.org/wp-content/uploads/forms/vhla-active-surveillance-guidelines.pdf.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating clear-cell renal carcinoma caused by von Hippel-Lindau disease [ID3932]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

NICE National Institute for Health and Care Excellence

About you

1.Your name	[redacted]
2. Name of organisation	Action Kidney Cancer
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Action Kidney Cancer started out as a patient support group called the Kidney Cancer Support Network founded in 2006 by cancer survivors Rose Woodward and Julia Black. The group provided practical and bespoke support to individual patients for access to life-extending cancer drugs to treat metastatic kidney cancer. Empowering patients to take an active role in their own health care, and in decisions affecting the choice, provision, and quality of cancer services throughout the UK, remains the top priority for Action Kidney Cancer. Over the years, Action Kidney Cancer has grown considerably, with a membership of over 1400 kidney cancer patients and carers on its closed community forum. In addition, our website regularly has over 300 visits per day from people looking for information about kidney cancer, advice, and support. Action Kidney Cancer is unique; originally it operated as a voluntary organisation, totally patient-led and managed by the patients and carers it represents. Although Action Kidney Cancer remains patient-led, the group is now a registered charity, which enables it to better meet the growing needs of the kidney cancer community in the UK. Funding for the organisation comes from trusts, foundations, and the pharmaceutical industry (around 58%), as well as fundraising activities/events and donations from the public and kidney cancer community (42%).
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment	We received £15,000 from Merck Sharp and Dohme (MSD) towards our multi-funded Ask the Expert series of videos for 2022. MSD were not involved in the planning, production, or implementation of the project.

companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	When gathering the information for this submission, we specifically asked for patient and carer experience of using belzutifan for the treatment of kidney cancer caused by von Hippel-Lindau (VHL) disease through our closed community forum. Over 1400 patients and carers use this facility to communicate on a regular basis, and we receive in the order of 5-600 interactions and comments a day.

Living with the condition

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Action Kidney Cancer is a patient-led kidney cancer charity with the largest and most active patient and carer membership across the UK. As such, we feel we are in a strong position to feedback how kidney cancer affects the day-to-day lives of people living with this disease. VHL is a rare inherited disorder that causes tumours and cysts to grow in certain parts of the body, including the kidney, brain, spinal cord, eyes, inner ear, adrenal glands, pancreas, and reproductive system. More than two thirds of people with VHL disease develop clear cell renal cell carcinoma (RCC). Metastatic clear cell RCC caused by VHL disease affects around 200 people per year in the UK. It is a devastating disease and is currently incurable and can recur after surgery in the remaining kidney. Most patients are forced to give up work because of the disease itself or the debilitating effects of current treatments. This brings enormous financial pressure for the patient and their family, and can precipitate psychological problems, depression, and loss of confidence and self-worth. Because VHL is an inherited condition, this causes stress, worry and guilt for the wider family who may carry the VHL gene, since they are concerned that they will pass the gene on to their children and grandchildren: *"My husband was only 41 years old in 1975 when he died. The doctors had no idea what was wrong. I was then left with 2 small children. I was told that the doctors were starting to take an interest in the VHL disease, and I was going to get the children controlled. I was told that the disease started in about the teenage years. At that age my son started having symptoms and of course I reacted immediately.* *My son was examined, and he had the disease, so I wanted my daughter examined too, but they said that it only affected boys. Of course, I had brought my daughter with me, so I insisted, that they examined her too. I still remember the stool I was sitting on when they told me that ….. also had VHL.* *Then the next generation came, and it became possible to test them. Then I had a grandson, when he was 11 years old, he got sick and had many surgeries.* *We have since had plenty of operations, either one or the other or the third. But it's great that you've come this far, otherwise I wouldn't have had any children."* Patients with metastatic RCC caused by VHL disease may suffer constant pain from metastatic tumours in the brain, bones, lungs, liver, and other rarer sites. Patients with bone metastases are at risk of bone breaks and spinal cord compression. Metastases in the lungs can lead to breathlessness, and persistent coughing. Spread of the cancer to the brain can lead to severe and debilitating headaches, confusion and, in some cases,

	paralysis. Kidney function is often compromised, and patients find daily living difficult, regularly needing periods of rest during the day. Patients tell us that psychological support is very difficult to access, and many patients are prescribed anti-depressant drugs to help manage their mental as well as physical clinical situation. Sexual function is affected for both male and female patients, and family life suffers as a result. The impact of a terminal hereditary diagnosis on the family, as well as the patient, also needs consideration; these families need support during the most difficult time in their lives when a loved one is diagnosed with a hereditary terminal disease. Patients diagnosed with hereditary kidney cancer or rare RCC subtypes, such as RCC caused by VHL disease, currently have very limited treatment options, exacerbating feelings of depression, fear, and low self-worth. Current first-line treatments offer an important, but sometimes short-lived period of stability, but not all patients respond to these treatments and most patients become refractory after a period. Biomarkers for the treatment of RCC caused by VHL disease are yet to be identified, and unfortunately clinicians are not able to predict which patients will respond to which drug. Therefore, a process of elimination is used to select the most effective treatment for individual patients. Clinicians in the UK should have the ability to choose the optimal treatments for individual patients from those available. Without a choice of an effective and tolerable treatment, most patients will face disease progression, including worsening of symptoms, such as severe pain, fatigue, and shortness-of-breath, and eventually death. Patients need to be able to choose their therapy to continue managing their disease, and to maintain quality of life. An increase in the choice of treatments will eventually lead to more personalised therapy, enabling patients and clinicians to tailor care plans to suite individual patient needs. After a partial nephrectomy for kidney cancer, one patient said: *"I have to have CT scans to see if it's going to come back. I do feel like I am constantly living in fear and living from scan to scan. Your mind works overtime, your imagination runs overtime that this dreaded disease is going to come back and rip your life apart. A cancer diagnosis changes your life and your whole outlook on life. It's never going to be easy to hear those words at any point….it massively impacted on my mental health."* Kidney cancer cases are rising year-on-year and there is a need for treatment with better overall survival rates than currently exist, especially for difficult-to-treat rare subtypes of RCC, such as RCC caused by VHL disease.

	Living with kidney cancer takes its toll on patients and their families both physically and psychologically. As a patient-led charity, Action Kidney Cancer encourages patients to ask for help from others to improve their wellbeing. Stress and anxiety can be reduced by talking about feelings with family, friends, a health care professional, or other people who have been through a similar experience. Taking part in activities that the patient enjoys, such as spending time with family and friends, socialising with other patients or carers, or relaxing activities such as walking, meditation or yoga, can also help to reduce stress and anxiety. Patients tell us that psychological support is very difficult to access, and many patients are prescribed anti-depressant drugs to help manage their mental health. Carers and close family members seem to find the psychological impact even harder as they live with the guilt of not being able to do all they can for their loved one. Family members also live with the constant worry that they are carrying the VHL gene and could develop RCC and pass this on to children and grandchildren. Access to an effective and tolerable treatment for RCC caused by VHL disease would enable patients and their families to know that they had tried their best to beat the cancer, leading to better family relationships and a subsequent improvement in quality of life and wellbeing for the patient. Finally, there is an unmet need for an effective treatment for hereditary subtypes, such as RCC caused by VHL disease, which are inherently difficult to treat. Patients diagnosed with hereditary kidney cancer or rare RCC subtypes currently have very limited treatment options, exacerbating feelings of depression, fear, and low self-worth.

Current treatment of the condition in the NHS

7. What do patients or carers think of current treatments and care available on the NHS?	Since there is a significant unmet need for the treatment of RCC caused by VHL disease, most cases are treated as for clear cell RCC. The current treatment pathway for metastatic RCC is surgery (either radical or partial nephrectomy), followed by either sunitinib, pazopanib or tivozanib in the first line setting, and lenvatinib, axitinib, everolimus, cabozantinib, or lenvatinib plus everolimus in the second line setting, all of which are oral medicines and have similar modes of action (vascular endothelial growth factor receptor (VEGFR) inhibitors or mTOR inhibitors that block angiogenesis in the tumour). Nivolumab is also recommended for use within NHS England for second- or third-line treatment of metastatic RCC and is the first third-line treatment in use by the NHS. Nivolumab is an immune checkpoint inhibitor (anti-PD-1), which is administered as a biweekly intravenous infusion, requiring outpatient hospital treatment (chemotherapy chair resources), and the associated travel time and expense for the patient and family member or carer. There are also three combination therapies available in the first line, including immunotherapy combinations (e.g., nivolumab plus ipilimumab), and immunotherapy plus VEGFR inhibitor combinations (e.g., avelumab plus axitinib and pembrolizumab plus lenvatinib). However, side effects for these combination therapies can be debilitating. We have extracted the following details from statements submitted to Action Kidney Cancer by patients living with metastatic RCC. Using currently available drugs, many patients suffer with the following side effects, all of which severely affect quality of life: • Extreme fatigue • Rash and itching • Severe hand and foot syndrome which can leave patients unable to walk • Intestinal problems (chronic diarrhoea) • Pneumonitis requiring hospital treatment and cessation of treatment • Severe mouth ulcers causing problems eating and drinking • Nausea and vomiting, which can also cause problems taking the medication • High blood pressure (hypertension) • Hyperthyroidism • Immune-related adverse events • Muscle pain/joint pain • Constipation • Diarrhoea

	All the above side effects require additional medicines to help patients manage the drugs and/or tumour pain, which often require opioid prescriptions. Costs for additional medicines to mitigate the side effects of these therapies should be considered. Other less serious side effects, which still affect the patient's quality of life, are headache, loss of taste, hair loss and change of hair colour, depression, loss of libido, and inability to drive. In some cases, treatment can affect a patient's quality of life to such an extent that clinicians recommend a dose reduction, and some patients are even advised to stop treatment because of severe side effects. Patients are aware that these treatments are life-extending drugs, but they continue to look for drugs with different modes of action, which can give improved overall survival with better quality of life. For patients that have been on standard first-line treatment with VEGFR inhibitors and experienced severe side effects, a new treatment with a better side effect profile will lead to a dramatic change in quality of life: *"No GI issues at all like I had with Sutent. Some knee and shoulder pain, but I am used to that from arthritis. Food is great, energy is great... I feel cured!! I realise I am not... but I never knew I had kidney cancer until they told me I did... and I never was sick. Start Sutent, and that is all I felt... sick. The surgery to remove my kidney, took me about 8 or 10 months to feel good again... brain met surgery... easy... my hard part was the Sutent side effects."* *"When I began treatment, I was in a state of helplessness. The abdominal tumour was located in such a position that it was growing so fast and caused so much pain I was unable to function. I was taking very high doses of Opiate pain medication with the result that I had no appetite and combined with side effects of Sutent my weight dropped to 139 pounds from 210 pounds. I lost large amounts of muscle. As a result, I was eventually confined to a wheelchair. I was unable to carry out even basic tasks and from being a very physically strong man who was very active and worked on my small ranch, I could do nothing for myself. I was very ill; I was told I had about 12 months to live. Tumours were growing aggressively."* *"I have had three infusions of Nivolumab, and I feel great. So far only minor SE. There was some shoulder, neck, and headaches at first, but none in the past week after my last infusion. I was on Votrient for almost year, and I am so glad to be rid of the GI side effects. My energy is good, my taste buds are back, no more tingling in hands and feet and my hair colour is slowly returning."* Although less serious than some of the side effects to current first-line treatments available via NHS England, some patients find the changes to their appearance caused by these treatments distressing: white, thinning hair,

	and pale skin make them feel nearer to death and singles people out as cancer patients. Some of the current first-line treatments can also cause issues with the thyroid gland, blood pressure, and cholesterol levels. From a psychological point of view, knowing that you have stage 4 cancer and knowing that there are possibly more effective treatments that you are not able to access is very difficult for patients. Carers and family members seem to find this even harder, as they live with a guilt of not being able to do all they can for their loved one. Access to a choice of treatments would enable patients and their families to know that they had tried their best to beat the cancer, leading to better family relationships and a subsequent improvement in quality of life and wellbeing for the patient. Nowadays, kidney cancer patients do not exist in silos. They communicate widely within online patient communities. International discussion forums exist where patients talk to one another daily. Patients are more aware of the experiences of others, including their access to innovative treatments, quality of life, and treatment successes and failures. News about lack of access to effective medicines ripples out to other patients and families, destroying their hope and positivity. Information about treatments is readily available to patients around the world on websites. Patients and clinicians in England expect NICE and the pharmaceutical industry to find a way to develop new and innovative treatments to extend their choices and to improve outcomes.

8. Is there an unmet need for patients with this condition?	There is a significant unmet need for an effective treatment for hereditary subtypes of kidney cancer, such as RCC caused by VHL disease, which are inherently difficult to treat. Patients diagnosed with hereditary kidney cancer or rare RCC subtypes currently have very limited treatment options, exacerbating feelings of depression, fear, and low self-worth.

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	A mutation in a gene called the von Hippel-Lindau (VHL) gene results in high levels of a protein called hypoxia-inducible factor, or HIF-2α in the blood of these patients. This causes changes in the cancer cells resulting in the growth of the tumour. A new, innovative medicine, called a hypoxia-inducible factor 2α (HIF-2α) inhibitor, or belzutifan, is a tablet that blocks the action of HIF-2α. Belzutifan has been proven to be a clinically effective and well-tolerated treatment and has recently been approved by the US Food and Drug Administration (FDA) for people with VHL disease who require therapy for associated RCC. Belzutifan is a first-in-class hypoxia-inducible factor (HIF) inhibitor for the treatment of VHL disease. Patients and carers are hopeful that belzutifan will improve response to treatment and subsequent survival, with minimal side effects and little impact on quality of life, after previous treatments have failed. This is borne out by the results from a study MK-6482-004. All 61 patients in the study had VHL-associated RCC. For RCC caused by VHL disease, overall response rate was 49%. 56% of responders had a duration of response of 1 year or more and the median time to response was 8 months. The most common side effects were anaemia, fatigue, increased creatinine, headache, dizziness, increased glucose, and nausea, which are easily treated and managed. Belzutifan is, therefore, well tolerated by patients with RCC caused by VHL disease. Unlike immunotherapy treatments that require travel to hospital every 2-3 weeks for treatment, needing time off work and the use of chemotherapy chairs, belzutifan is a tablet that can be taken at home. This is an advantage for those patients who may need to travel some distance to regional cancer centres, take time off work, or have a partner travel with them for treatment. These results show that belzutifan can control the cancer in a large proportion of patients with RCC caused by VHL disease who have already been treated with anti-cancer medication. In addition, belzutifan is well tolerated by these patients.

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	The results from the MK-6482-004 trial with 61 people with metastatic RCC caused by VHL disease showed significant improvement in survival and response to treatment with belzutifan in the second or third line compared to the standard of care. In addition to improvement in survival and response to treatment, patients on belzutifan reported more tolerable side effects and an improvement in health-related quality of life. Most side effects are treatable and manageable. The following quotes are taken from patients with RCC caused by VHL disease being treated with belzutifan: *"Despite the high dose of 200mg/day, there are no recognisable side effects - apart from a permanently low but uncritical HB value. I receive the medication every four weeks …… during the examination cycles prescribed by the study.* *Belzutifan enables me to live a largely normal life and to run a small company with 20 employees ….. full-time again. My last and only surgery was the nephrectomy in June 2020."* *"My son and I are still in the trial and are taking 120 mg Belzutifan daily with very few side effects and great effect on kidney cancer and VHL. I had 2 VHL related kidney cancer, one of which has now disappeared. What remains is one kidney cancer, everything else has disappeared.* *My son experiences great reduction of tumour in the brain, as well as reduction and stabilization of kidney cancer after taking Belzutifan.* *It has had a great impact on our quality of life that we now do not have to worry every time we MRI and CT are scanned [sic], as we know that Belzutifan keeps the disease at rest and we do not have to undergo repeated surgery as before. I now work full-time as a health visitor and my son will graduate in 1 year and become a doctor himself."* *"I see from my experience that Belzutifan can help VHL patients live a normal life with a clear reduction in surgery and treatment. The risk profile - as shown by my experience with a high dose over the last almost 2 years, other "real-life" data from the US and available studies - is clearly positive. Without Belzutifan, I would probably no longer be able to work in the way I do and be available to my family. I would be on treatment with classic kidney cancer drugs, which have much more pronounced side effects and would have no effect on my other VHL lesions. Belzutifan saved my life and it will save the lives of other VHL patients."* *"My great concern is my brother and other VHL patients who continue to suffer from VHL with repeated hospitalizations and surgeries, spread of kidney cancer, loss of function as late complications after surgery*

	and long rehabilitation courses. They need Belzutifan who [sic] can change and save their lives, just as my son and I have experienced that Belzutifan has changed and saved our lives." *"My daughter and grandson are part of an experiment with Belzutifan that has given us hope, but my son did not join the Belzutifan project as he was too ill.* *I only have one big wish, that my son and the other VHL patients may have Belzutifan."* Metastatic RCC caused by VHL is a devastating disease and is currently incurable. Most patients are forced to give up work because of the disease itself, and current treatments are very debilitating. This brings enormous financial pressures for patients and their families, sometimes resulting in psychological problems, depression, loss of confidence and self-worth. Patients with VHL disease have multiple tumours in different parts of their bodies, including the kidney, brain, spinal cord, eyes, inner ear, adrenal glands, pancreas, and reproductive system. The current treatment for these tumours is surgery and anti-cancer medication for the tumours that have spread. Treatment with belzutifan will reduce the need for unnecessary surgery and help prevent these people from developing chronic kidney disease and having to undergo dialysis. We understand that the patient population is small, and treatments are expensive, and we appreciate the budgetary constraints of the NHS. Nonetheless, NICE and the manufacturer need to work collaboratively to negotiate an acceptable patient access scheme to ensure the unmet need for an effective treatment for these patients is met and patients can benefit from this first-in-class innovative treatment for RCC caused by VHL disease.

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients with clear cell RCC caused by VHL disease.

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Belzutifan is a first-in-class treatment for clear cell RCC caused by VHL disease. However, there are significant unmet needs for the treatment of other subtypes of non-clear cell RCC. This puts patients with other rare subtypes of non-clear cell RCC at a disadvantage when it comes to treatment options. People from deprived areas of England and Wales, and people whose first language is not English are less likely to visit their GP if they have signs or symptoms of cancer, putting them at a disadvantage for early diagnosis and treatment of their cancer. Also, cultural issues regarding toilet habits may deter some ethnic minorities from discussing symptoms with their GP. Awareness campaigns targeted at these groups of people might help to improve the care and treatment of people from deprived areas or ethnic minorities.

Other issues

13. Are there any other issues that you would like the committee to consider?	Currently, UK cancer survival rates trail about 10 years behind other comparable European countries, including Italy and Austria. If the UK is to improve patient outcomes, including the patient experience as well as overall survival, it is vital that these novel treatments are made available to patients in order that they have the best possible care. If these medicines are not made available, it leaves UK patients at a major disadvantage in terms of the availability of innovative cancer treatments; these patients are likely to die prematurely compared to other kidney cancer patients in the rest of Europe and North America. Poor UK survival rates might possibly be due to the restrictions in clinical choice brought about by UK regulatory authorities. In the absence of biomarkers for the treatment of RCC, clinicians are not able to predict which patients will respond to which drug, and drug selection is accomplished by trial and error. Clinicians should have the ability to choose the most effective treatments for individual patients from those available. Without effective treatment alternatives, most patients will face disease progression. A choice of treatment is paramount for the effective management of the progression of this disease and maintenance of quality of life. Current treatment options are not effective for everyone. Undue restrictions in accessing novel therapies would simply add unnecessary additional burden to patients with a terminal diagnosis. Having more treatment choice would enable patients and oncologists to individualise treatment plans according to specific disease/treatment history and contraindications, thereby enabling the best possible quality of life for the patient. Although unproven, belzutifan could potentially be used for the treatment of patients with other rare or hereditary (non-clear cell) subtypes of RCC where there is currently a significant unmet need for an effective and safe treatment strategy.

14. To be added by technical team at scope sign off. Note that topic-specific questions will be added only if the treatment pathway or likely use of the technology remains uncertain after scoping consultation, for example if there were differences in opinion; this is not expected to be required for every appraisal.] **if there are none delete highlighted rows and renumber below**	

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Belzutifan is a first-in-class innovative treatment for RCC caused by VHL disease. Belzutifan has been proven to be very safe and effective as a second- or third-line treatment for people with advanced RCC caused by VHL disease, and has already been approved for use by the FDA in the USA • Belzutifan is extremely well tolerated, as well as showing significant improvement in survival and response to treatment in the second or third line compared to the standard of care for RCC caused by VHL disease • Belzutifan addresses an area of significant unmet need in the treatment of RCC caused by VHL disease • Adding belzutifan as a choice in the second or third line enables patients and clinicians to individualise treatment plans to better control this disease and maintain a high quality of life • The extended survival, reduction in unnecessary surgeries and relative tolerability of belzutifan enhances quality of life and enables patients to contribute socially and economically to society.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating clear-cell renal carcinoma caused by von Hippel-Lindau disease [ID3932]

Professional organisation submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you

NICE National Institute for Health and Care Excellence

1. Your name	[redacted]
2. Name of organisation	University of Cambridge
3. Job title or position	Regius Professor of Physic and Head of the School of Clinical Medicine
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? Yes or No I am an employee of the University of Cambridge and am responsible for oversight of a substantial number of clinical academics. I am also a long-standing member of the UK Kidney Association, which represents the UK nephrology community and have been asked to make this submission on their behalf A specialist in the treatment of people with this condition? Yes or No I have previously led multidisciplinary clinics for patients with VHL disease (approx 2006 - 2012) at the Hammersmith and Royal Free, but since moving to Cambridge in 2012 I no longer oversee care of patients with VHL disease. A specialist in the clinical evidence base for this condition or technology? No Other (please specify): I have led a scientific program elucidating the physiological operation and pathological consequences of the VHL protein and of HIF-2
5a. Brief description of the organisation (including who funds it).	We are the leading professional body for the UK renal community, dedicated to improving lives by supporting professionals in the delivery of kidney care and research. We have over 1,200 doctors, scientists and multi-professional team members. Funded by membership fees and corporate sponsorship.

5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] **If so, please state the name of manufacturer, amount, and purpose of funding.**	UKKA has received the following amounts: Novartis • £130k - research Pfizer • £2.4k event sponsorship • £175k - research
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	

The aim of treatment for this condition

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	Key aims of treatment are 1) To prevent metastatic spread of renal carcinoma 2) To preserve kidney function 3) To reduce requirements for renal surgery
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Reduced growth rate Reduction of size Reducing rate of breaching the 3cm threshold for operative intervention.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	Annual surveillance imaging of the kidneys from age 16. Once a lesion is identified, it is monitored and when a lesion approached 3cm it is removed surgically by partial nephrectomy or by ablation and at the same time other lesions in the kidney are removed.
9a. Are any clinical guidelines used in the	Guidelines from Maher ER et al European J Human Genetics 2011 19:617-623

treatment of the condition, and if so, which?	
9b. Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Yes. There is a good audit of current practice in the NHS Maher et al Br J Cancer 2022 126 1339-1345
9c. What impact would the technology have on the current pathway of care?	If it reduces the rate of renal tumor surgery/ablation and improves presentation of renal infection that will clearly benefit patients.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	The new technology will certainly change the current approach as it will alter the growth of renal tumors. It will also make patients anaemic which may require transfusion and/or erythropoiesis stimulating agents.
10a. How does healthcare resource use differ between the technology and current care?	Belzutifan would be likely to require additional hospital visits and blood tests. It might also require additional scans (MRI) to monitor tumor size.
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	This should be used in specialist clinics
10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Some training of staff would be necessary. It may be necessary to increase the capacity of VHL clinics.

11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. But it is challenging to know what will be the best time to introduce treatment, and how long it should be for.
11a. Do you expect the technology to increase length of life more than current care?	Possibly although current approaches mean death related to RCC is very rare in VHL disease loss of renal function, and other manifestations of VHL disease driven by HIF-2 probably shorten life expectancy.
11b. Do you expect the technology to increase health-related quality of life more than current care?	Yes.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	VHL is a rare disease – approximately 1 in 50,000 – so this is not applicable.

The use of the technology

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	Current guidelines for monitoring and managering RCC in VHL disease are clear and well understood. How best to use belzutifan will complicate this. Also, additional monitoring will be needed, and anaemia will be a common side effect requiring treatment in some patients.

affecting patient acceptability or ease of use or additional tests, or monitoring needed.)	
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Development of such rules will be difficult but is likely to include a requirement that renal tumors are present and are less than 3cm as in the Phase II study that led to US regulatory approval.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Effects on pancreatic, retinal, CNS, adrenal tumors will almost certainly be helpful. Also, the patients will, I believe, benefit from knowing that they are on a treatment that is modifying the disease rather than simply removing tumors as they crop up.
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	The technology undoubtedly targets the key pathway in a precise and effective way. Potentially it could decrease the burden of kidney surgery, and reduce damage to other organs (e.g., brain and eye).
16a. Is the technology a 'step-change' in the management of the condition?	Yes, potentially.

16b. Does the use of the technology address any particular unmet need of the patient population?	Repeated surgery and ablation to the kidneys is not an ideal way to manage the condition and this could provide a better alternative
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Some patients will need transfusion and/or treatment with ESA's

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, they do
18a. If not, how could the results be extrapolated to the UK setting?	N/A
18b. What, in your view, are the most important outcomes, and were they measured in the trials?	The trials to date show an effect on tumor growth. They do not establish a reduction in the need for surveillance, operation or ablation, or improved presentation of kidney function.
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Long term outcomes in terms of tumor behaviour are uncertain in my view
18d. Are there any adverse effects that were not apparent in clinical	Not to my knowledge

trials but have come to light subsequently?	
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]? [delete if there is no NICE guidance for the comparator(s) and renumber subsequent sections]	N/A
21. How do data on real-world experience compare with the trial data?	I am not able to comment.

Equality

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No.
22b. Consider whether these issues are different from issues with current care and why.	No.

Topic-specific questions

23 [To be added by technical team at scope sign off. Note that topic-specific questions will be added only if the treatment pathway or likely use of the technology remains uncertain after scoping consultation, for example if there were differences in opinion; this is not expected to be required for every appraisal.] **if there are none delete highlighted rows and renumber below**	

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Belzutifan is the first non-surgical treatment for VHL-related RCC • It is effective in slowing tumor growth • It is less clear whether this will reduce burden of surgery and hence delay loss of renal function • Well tolerated, but will result in anaemia • Effect on other VHL manifestations (e.g., eyes, CNS) likely to be beneficial

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Belzutifan ~~for untreated renal cell carcinoma caused by~~ for treating tumours associated with von Hippel-Lindau disease [ID3932]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Prepared by a working committee of Members and Trustees of VHL UK/Ireland.**

2. Name of organisation	**VHL UK/Ireland (Registered Charity Number 1160381)**
3. Job title or position	**Members and Trustees**
4a. Brief description of the organisation (including who funds it). How many members does it have?	Funded purely by donations from the public, the aim of the VHL UK/Ireland charity is to serve the charitable needs of persons who have **von Hippel- Lindau Syndrome (VHL)** and similar genetic conditions such as **Hereditary Leiomyomatosis and Renal Cell Cancer (HLRCC)** and **Birt-Hogg- Dubé Syndrome (BHD)**, their families and carers in the UK and Ireland for the public benefit in particular, but not exclusively by: 1. Providing support and advice regarding these genetic conditions 2. Providing funds for research into the genetic conditions and for equipment to assist with such research 3. To advance the education of the public in all matters concerning these genetic conditions The charity is made up of 13 charity members and has a public following of approximately 560 persons with either VHL, HLRCC, BHD, their carers, family members and friends. VHL/UK Ireland is closely affiliated with the VHL Alliance in the USA which has been in existence for 30 years and has approximately 4,500 followers. The VHL Alliance produces handbooks for both adults and children who suffer from VHL.
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Our trustees have an aggregate total of over 50 years' experience volunteering for the charity, many of whom are patients and/or carers themselves. We have drawn on their own experiences and those of other patients and carers across the UK and the world, who are continually communicating with us via support groups, online forums and one-to-one discussion. We also refer to published qualitative and quantitative data both from our own website (stories [1] and surveys [3,4]) and others – see full references at the end of this submission.

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Please read the patient stories https://vhl-uk-ireland.org/stories/ [1], and watch this short video (1.5 minutes) which explains exactly what it is like to live with VHL [2]: https://www.youtube.com/watch?v=5Y89uNyoeLc VHL is a truly devastating, rare, life-long (life-limiting and for some life-threatening), incurable disease, that can cause multiple, malignant, and benign tumours/cysts throughout the body. People may read the description of "VHL benign tumours" and do not realise that these are likely to be on multiple organs, simultaneously, repetitive, and life-threatening. VHL does not go into remission, tumours do not regress naturally and there is no cure. High risk, invasive surgery is the only way to remove these tumours and it is not uncommon to lose sight, eyes, part/whole kidneys/pancreas (plus adjoining organs during the Whipple's procedure), adrenal gland/s, suffer neurological issues such as paralysis after spine or brain surgery and require lifelong medical intervention such as dialysis or Pancreatic Enzyme Replacement Therapy (PERT - post Whipple's). The tumours themselves, or surgery to remove them, mean patients can suffer from a wide range of debilitating symptoms, depending on the site of tumours, including constant pain, loss of balance and motor skills, loss of vision, breathlessness, coughing, headaches, confusion, severe nausea, and fatigue. See survey 1 [3] for more detail. VHL is variable and unpredictable; some patients may only develop a few or a single tumour in a single organ but as tumours can often recur in the same organ, and occur concurrently across multiple organs, some patients may develop multiple tumours across multiple organs throughout their life. The effect is that the worst affected patients may have over 40 or more operations in their lifetime, and some of these might on occasion be only a matter of months apart. Survey 2 [4] showed 77% of respondents had tumours in 3-5 areas of their body, at that time. Survey 1 [3] showed 41% had 5 or more surgeries in their life so far. The effects of surgery (or multiple surgeries) can be lifelong and devastating for both patient and carer, both physically and mentally. Metastatic tumours (usually adrenals, or kidneys/pancreas if not surgically addressed before the advised 2-3cm cut off) are sometimes treated with standard chemotherapies but it is widely reported these are ineffective and have 'high grade' side effects severely impacting quality of life.

	As well as the obvious physical impacts described above, the unpredictability of VHL creates a heightened sense of stress and sadness in both VHL patients and their carers. 67% of VHL patients reported experiencing anxiety or depression in relation to the recovery from treatment of a VHL tumour [3]. It was noted that carers often experience higher anxiety levels than the patients themselves regarding surgeries and monitoring. It can sometimes be even more difficult watching a loved one suffer than experiencing the condition. Carers suggest that this is due to the feeling of helplessness and the lack of treatment options. Specialist psychological support for VHL patients and carers is nonexistent, with them replying on standard GP routes and peer to peer support. Day to day life is regularly impacted by VHL for patient and carer. Even in a 'good year', just for surveillance, patients require multiple scans and consultations. Survey 1 [3] showed 80% of patients had 5 or more appointments to manage their VHL per year, 30% had 10 or more. These could be at different hospitals at different times as consultants do not always work collaboratively across the varied specialities, leading to regular anxiety, repetitive discussions, and lots of time out of their daily lives for both patients and carers. 80% of carers reported that they attend all appointments with the person(s) that they care for [3]. Some patients with more complex cases have reported up to 40 hospital appointments in a "good year". With restricted mobility issues, this can also mean overnight stays in hotels close to the hospital. Both surveys [3,4] demonstrate that VHL affects the quality of life of a patient and their carers lives, with monitoring and effects from surgery/treatment impacting things like education, work, career, social and leisure activities, travel, relationships, family planning, mental health, future life plans and finances (e.g. no life insurance, expensive travel insurance). An important consideration is the unique, multi-familial nature of the disease. Patients can be carers and carers patients. Multiple family members can be affected by symptoms or treatment at the same time and often live with the heavy burden that some in the family have already been severely affected or have died because of VHL. High levels of guilt are often felt due to the passing on of the gene (or not if perhaps one sibling has it and the other doesn't) or burden/impact of navigating it as either patient, carer or both.

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	There are no other alternative treatments licensed for the treatment of VHL, and there is nothing else in the pipeline. The most typical treatment is surgical with the profound physical symptoms of both the tumours and the surgeries listed above in Q6. Metastatic disease in VHL is almost always terminal and attempts to use conventional chemotherapy or immunotherapies are largely reported as ineffective for VHL as they do not treat the underlying cause and are often used as a 'last resort'. The complex, repetitive, time consuming and familial nature of managing VHL places a heavy burden on the NHS and on the wellbeing and quality of life of the patient and carer for their whole lives.
8. Is there an unmet need for patients with this condition?	There is an enormous unmet need and potential opportunity. • There are no other alternative treatments licensed for the treatment of VHL, and there is nothing else in the pipeline. • VHL patients do not see natural tumour regression. Stability of any patient's tumours would be considered 'successful' ordinarily, so being able to reduce them without surgical intervention can be considered remarkable in this setting. Survey 2 [4] shows that ALL patient respondents using belzutifan (Welireg) experienced some form of stability, slowdown of growth, reduction in size or disappearance of at least one tumour (but often multiple). • Standard chemotherapies/immunotherapy impact greatly on patient/carer quality of life – sometimes tried for metastatic RCC or Endocrine cancers, which are largely reported as ineffective for VHL as they do not treat the underlying cause. • Avoiding surgeries and ineffective treatments could bring about huge improvements in the physical wellbeing of patients and the mental wellbeing and quality of life of both patients and carers.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Belzutifan is a paradigm shift for the treatment of VHL in stopping or reversing tumour growth and preventing invasive/life changing surgeries, metastatic cancer (where treatments that are usually attempted are mostly 'useless') and the need for dialysis. VHL tumours do not regress naturally and are currently treated one at a time. Belzutifan treats multiple tumours, at multiple sites, at the same time. Belzutifan has been described as a "game-changer" by the VHL Alliance in the USA and medical professionals globally. Survey 2 [4] shows ALL respondents surveyed experienced some form of **stability** (61% on one or more tumour/s), **slowdown of growth** (66% on one or more tumour/s), **reduction in size** (66% on one or more tumour/s), and in some cases **disappearance** of some tumours (27% at least one tumour). Belzutifan will therefore limit the need for multiple, high risk, invasive, life limiting (and sometimes life threatening) surgeries. Survey 2 [4] shows 80% of patients already using belzutifan (Welireg) believe they have avoided an imminent surgery. Dr. Othon Iliopoulos, MD, PhD demonstrates strong evidence for 'procedure reduction' in his presentation (@48mins) https://www.youtube.com/watch?v=ptqV5RdYdWw.[5] **Less surgery** will mean reduced incidence of complications such as stroke, paralysis, loss of motor functions, or death (brain and spine), reduction in the occurrence of kidney disease, dialysis and/ or RCC (data shows belzutifan is an extremely effective treatment for RCC targeting the underlying VHL itself), avoidance of the life changing Whipple's procedure and the risk of malabsorption, and lifelong PERT, and a reduction in the need for invasive procedures for the eye/s, where laser treatment is not possible/successful, which may even prevent loss of sight in one or both eyes and eye removal. Less surgery will also result in less exposure to added risks such as multiple anaesthesia or infection. In general, if a tumor is deemed inoperable, or if the risks associated with surgery are very high, belzutifan will provide **an option**, where currently there is none and used at the correct time, **could reduce the incidence of metastasis** where we know treatment options are largely ineffective. As well as the obvious physical advantages above, survey 2 [4] shows improvements in **quality of life** for both patients and carers for current users of belzutifan. 68% of patients said it had improved their own

	quality of life and 88% carers reported theirs had improved. The summary shows expected positive outcomes to social/ leisure activities (70% patients/88% carers), independence (68% patients/75% carers), relationships, work/career, education, and ability to travel. Survey 2 [4] also demonstrates that belzutifan will have a positive impact on the **mental health** of both patient (84%) and carer (88%), reducing worry about VHL (82% patients/75% carers) and being more able to plan for the future (82% patients/88% carers). The current practice is to treat each tumor individually, whereas belzutifan is a treatment that has shown a **benefit for a wide range of tumours simultaneously**. For example, if the "target" organ is the kidney, but the patient also has tumours in eye and brain which might otherwise be operable, the drug has been shown to shrink ALL these tumours at the same time-thus, multiple surgeries are therefore avoided or substantially delayed. It may also prevent new tumours in the same and different organs from developing. This could bring patients and carers significant **psychological relief.** Belzutifan is a tablet that can be **taken at home.** This is an advantage and an improvement in quality of life for those patients and carers who may need to travel long distances to regional cancer centres or take time off work. For some, the very act of travel can be severely debilitating due to the symptoms caused by either the tumour location or symptoms derived from previous surgery or other interventions. In survey 2 [4], 91% of belzutifan (Welireg) users said they preferred the drug to surgery, 100% of carers said they preferred the drug to surgery for their patient.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Survey 2 [4] shows that the number of patients who experienced side effects was high (75%). However, NO patients surveyed stopped treatment permanently and 52% continued using it regardless. Of those who took some action, the most common response was to reduce dosage, either temporarily or permanently (30%). This supports what we know anecdotally that side effects are low grade and patients report an improvement within 6 months as the body learns to tolerate the drug (hypoxia is the main side-affect and is considered an "on target" effect). Belzutifan side effects are significantly more tolerable than those of standard chemotherapies/immunotherapy – sometimes tried for metastatic RCC or Endocrine cancers, which are largely reported as ineffective for VHL as they do not treat the underlying cause.

	Not knowing the long-term effect of belzutifan (Welireg) worries only 34% of patients. 63% carers of VHL patients worry about the long-term effects. Compared to the alternatives (ineffective non VHL specific therapies or high-risk life changing/limiting surgeries), patients may be willing to take the risks of side effects, as 93% of belzutifan (Welireg) users said they preferred the drug to surgery, 100% of carers said they preferred the drug to surgery for their patient. "I suffered in the first few months from fatigue and higher sensitivity to heat but both have subsided and are not noticeable now that my body has adjusted." "I will gladly trade the fatigue and headaches for keeping my pancreas."
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Some late stage VHL patients are more seriously ill through the effects of multiple surgeries or metastatic tumours and are at the point where further surgery is not advisable or attempts to use standard chemo/immune therapies are futile. These patients may benefit greatly from this technology to increase their quality of life and lifespan. Furthermore, we understand that trials are taking place which seek to broaden the scope of this treatment to patients with clear cell RCC that is not VHL-related.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	All patients must manage multiple appointments, with multiple NHS departments, whilst juggling home and work-life and ensuring that nothing is missed. This can also be for multiple members of the same family. Sadly, some patients who do not follow-up and chase consultants, do not advocate for themselves, nor have someone to advocate for them, can sometimes be overlooked or lost in the system and therefore may miss the required surveillance and treatment. These patients may need support to ensure that they receive the drug if they are eligible. Patients from deprived areas, with language, learning or cultural barriers, or those with disabilities may be

	at a disadvantage. On the flip side, patients who can fund the treatment privately would be at a great advantage. Whilst some NHS Trusts do have VHL-related multi-disciplinary teams, unfortunately this is not consistent across the UK. It is also known that some patients are managed by medical teams who are lacking in VHL expertise. There may be concerns that prescription of the drug may not be evenly distributed in the UK. Patients took part in the trial of belzutifan in the belief that if it were successful, it would be made available to other VHL patients worldwide. This may affect future uptake of such trials in such a small group. Patients unable to access belzutifan because of where they live, would create global inequalities, likely distort global mortality rates and put those without access at a major disadvantage to others.
Other issues	
13. Are there any other issues that you would like the committee to consider?	Some current patients who are approaching surgery are seriously considering delaying, in the hope that belzutifan will be approved in time for them to avoid surgery altogether. As a charity we advise patients to speak with their medical teams about this. Patients are anticipating approval as they are seeing it being widely prescribed for VHL patients in other parts of the world (USA/Canada via social media) which has a profound psychological impact on them. We are aware of some patients moving continents to gain access to belzutifan. We believe there has been some uncertainty about the trial data and lack of 'control data' for belzutifan. In our view, ALL patients who have not used belzutifan (Welireg) can be viewed as the 'control group' when considering trial results, none of whom have ever reported slowed growth, shrinkage, or disappearance of their tumours without surgery as there is no natural regression. We believe that standard economic modelling tools would be ineffective for assessing this technology, due to the highly complex, unpredictable, multifaceted, familial nature of VHL. We believe however, that the use of belzutifan will provide reduced costs for surgical interventions and the resulting lifetime consequences. The comparator section of the scope is incorrect for CNS, stating “standard of care **with** belzutifan”, as opposed to pNET and RCC, which state “standard of care **without** belzutifan”.
Key messages	

14. In up to 5 bullet points, please summarise the key messages of your submission:

- VHL is a truly devastating, rare, life-long (life-limiting and for some life-threatening), incurable disease - it does not go into remission and there is currently no cure. Even if successfully removed, tumours can often recur in the same organ, and occur concurrently across multiple organs and so it is common for patients to develop multiple tumours across multiple organs throughout their life.
- VHL can affect multiple family members, with many patients being carers and carers being patients. A great number live with the burden of knowing a relative has already been severely impacted, or died, due to VHL.
- Belzutifan is a paradigm shift for the treatment of VHL in stopping or reversing tumour growth and preventing invasive/life changing surgeries, metastatic cancer (where treatments that are usually attempted are mostly 'useless') and the need for dialysis. VHL tumours do not regress naturally and are currently treated one at a time. Belzutifan treats multiple tumours, at multiple sites, at the same time.
- As well the physical improvements, mental health and quality of life/life expectancy are drastically improved for many VHL patients and carers as a direct result of using belzutifan. This is clearly shown in our surveys [3,4].
- We believe that standard economic modelling tools would be ineffective for assessing this technology, due to the highly complex, unpredictable, multifaceted, familial nature of VHL. We believe however, that the use of belzutifan will provide reduced costs for surgical interventions and the resulting lifetime consequences.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic(s) above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

References

1. **PATIENT STORIES:** https://vhl-uk-ireland.org/stories/

2. **'LIVING WITH VHL' VIDEO:** https://www.youtube.com/watch?v=5Y89uNyoeLc

3. **SURVEY 1: VHL UK/Ireland Patient/Carer Survey (RCC ONLY) - June 2022**
   Result Summary - VHL UK SURVEY SUMMARY June 2022 (vhl-uk-ireland.org)
   As part of the ongoing NICE appraisal for belzutifan (Welireg) in England, in June 2022, a patient/carer survey was submitted to the VHL UK/Ireland community to help gather information for the charity's submission. NOTE! The submission was purely for RCC Renal Cell Carcinoma at this stage, but the scope has now been widened to include CNS hemangioblastomas and pancreatic pNETS. Unfortunately, we were unable to widen the survey to incorporate the wider marketing approval due to VHL related illnesses of our team. (One trustee had a Whipple's procedure of the pancreas, quickly followed by brain surgery; another has lost an eye to multiple tumours; and a third has had brain surgery following pregnancy).

4. **SURVEY 2: VHL UK/Ireland Patient/Carer 'BELZUTIFAN experience' Survey - June 2023**
   Result Summary - VHL UK/Ireland Patient/Carer belzutifan (Welireg) Survey June 2023- Summary (vhl-uk-ireland.org)
   In June 2023, we conducted a short survey amongst VHL patients who have experienced using belzutifan(Welireg).

5. '**VHL 101'** - Dr. Othon Iliopoulos, MD, PhD https://www.youtube.com/watch?v=ptqV5RdYdWw. (August 2023)

Please note we have used data from VHL patients across the world (both inside and outside the UK).

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

Produced by Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Susan O'Meara, Reviews Manager, KSR Ltd, United Kingdom (UK)
Venetia Qendri, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands (NL)
Eline Krijkamp, Health Economist, ESHPM, EUR, NL
Jiongyu Chen, Heath Economist/Systematic Reviewer, KSR Ltd, UK
Xiaoyu Tian, Health Economist/Systematic Reviewer, KSR Ltd, UK
Mubarak Patel, Systematic Reviewer, KSR Ltd, UK
Rachel Croft, Information Specialist, KSR Ltd, UK
Lisa Stirk, Senior Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to Susan O'Meara, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 21/07/2023

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number NIHR 135685.

Declared competing interests of the authors None.

Acknowledgements: None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

No academic in confidence data are included in the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

O'Meara S, Qendri V, Krijkamp E, Chen J, Tian X, Patel M, Croft R, Stirk L, Armstrong N, Corro Ramos I, Wolff R. Belzutifan for treating tumours associated with von Hippel-Lindau disease [ID3932]: a Single Technology Assessment. Kleijnen Systematic Reviews Ltd, 2023.

Contributions of authors

Susan O'Meara acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Venetia Qendri and Eline Krijkamp acted as health economists, critiqued the company's economic evaluation and contributed to the writing of the report. Nigel Armstrong acted as health economist and systematic reviewer on this assessment, critiqued the company's clinical effectiveness evidence and economic evaluation and contributed to the writing of the report. Jiongyu Chen, Xiaoyu Tian and Mubarak Patel acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Rachel Croft and Lisa Stirk critiqued the search methods in the submission and contributed to the writing of the report. Isaac Corro Ramos acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Robert Wolff critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

.

Abbreviations

AACR	American Association for Cancer Research
AAO	American Academy of Ophthalmology
AdViSHE	Assessment of the Validation Status of Health-Economic decision models
AE	Adverse event
AiC	Academic in confidence
AIC	Akaike Information Criterion
ALT	Alanine aminotransferase
AMCP	Academy of Managed Care Pharmacy Annual Meeting
APaT	All participants as treated
ARVO	Association for Research in Vision and Ophthalmology
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
ATT	Average treatment effect of the treated
BIC	Bayesian Information Criterion
BID	Twice a day
BMI	Body mass index
BNF	British National Formulary
BOR	Best overall response
BSA	Body surface area
CADTH	Canadian Agency for Drugs and Technologies in Health
CC	complication and comorbidity
CE	Cost effectiveness
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CiC	Commercial in confidence
cM0	Cancer that has not spread to a different part of the body but cancer cells have been found in the blood, bone marrow or lymph nodes distal to main tumour
cM1	Cancer that has spread to another part of the body
CNS	Central nervous system
COVID-19	Coronavirus Disease of 2019
CPI	Consumer Price Index
CR	Complete response
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
cys	Cystadenoma
DAE	Discontinuation due to AE
DARE	Database of Abstracts of Reviews of Effects
DCR	Disease control rate
DFS	Disease-free survival
DP	Decision problem
DOR	Duration of response
DSA	Deterministic sensitivity analysis
EAG	Evidence Assessment Group
ECC	European Cancer Congress
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDSS	Expanded Disability Status Scale
eGFR	Estimated glomerular filtration rate
eMIT	Electronic market information tool

EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	EQ-5D-3 Levels
EQ-5D-5L	EQ-5D-5 Levels
ESHPM	Erasmus School of Health Policy and Management
ESMO	European Society for Medical Oncology
ESRD	End stage renal disease
ESS	Effective sample size
EU	European Union
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FE	Fixing errors
FOLFIRI	Folic acid, fluorouracil, and irinotecan
FOLFOX	Folic acid, fluorouracil, and oxaliplatin
FV	Fixing violations
G	histological cancer grade
GB	Great Britain
GBP	Great British Pound
GI	Gastrointestinal
GP	General Practitioner
GU	Genitourinary
Hb	Haemangioblastoma
HIF	Hypoxia inducible factor
HIF-2α	hypoxia inducible factor 2 alpha
HR	Hazard ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICH	International Council for Harmonisation
iDBC	Institute for Medical Technology Assessment Disease Burden Calculator
iMTA	Institute for Medical Technology Assessment
Inc.	Incremental
IPD	Individual participant (or patient) data
IRC	Independent Review Committee
iRECIST	Immune Response Evaluation Criteria in Solid Tumours
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITC	Indirect treatment comparison
IV	Intravenous
JBI	Joanna Briggs Institute
Kg	Kilogrammes
KM	Kaplan-Meier
KPS	Karnofsky Performance Status
KSR	Kleijnen Systematic Reviews Ltd
LGR	Linear growth rate
LYs	Life years
LYG	Life years gained
M1	Metastatic disease that has spread to other parts of the body
MA	Marketing authorisation
MAIC	Matching-adjusted indirect comparison
Max	Maximum
MD	Doctor of Medicine
MDT	Multidisciplinary team
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical subject headings
Mg	Milligrams

MHRA	Medicines and Healthcare Products Regulatory Agency
Min	Minimum
MJ	Matters of judgement
mRECIST	Modified Response Evaluation Criteria in Solid Tumours
MRI	Magnetic resonance imaging
MSD	Merck Sharp & Dohme
MU	Million units
N	Number of participants
NA	Not applicable
NE	Not estimable/not evaluable
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NCI UOB	National Cancer Institute Urological Oncology Branch
NCT	National Clinical Trial (registry number)
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NL	The Netherlands
NMA	Network meta-analysis
NR	Not reported
NSCLC	Non-small-cell lung cancer
ONS	Office of National Statistics
ORR	Overall response rate
OS	Overall survival
PAS	Patient Access Scheme
PD	Progressive/progressed disease
PFS	Progression-free survival
PICOTS	Population, Interventions, Comparators, Outcomes, Timeframe, Study design
PK	Pharmacokinetics
pM1	cancer measuring >0.2mm that has spread to another part of the body
pNET	Pancreatic neuroendocrine tumour
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
Q	Quartile
Q#W	Once every # weeks
QALY	Quality-adjusted life year
QD	Once a day
QoL	Quality of life
QTc	Corrected QT interval (on ECG)
QW	Once weekly
RCC	Renal cell carcinoma
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
RoB	Risk of bias
ROBINS-I	Risk Of Bias In Non-randomised Studies
RR	Response rate
RW	Real world
SA	Severity adjusted

SAE	Serious adverse event
SC	Subcutaneous
ScHARR	School of Health and Related Research
SD	Standard deviation
SD	Stable disease
SE	Standard error
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoC	Standard of care
STA	Single Technology Appraisal
TA	Technology assessment
TECH-VER	Technical Verification (checklist)
TNM	Cancer staging system based on tumour size and spread (T), spread to nearby lymph nodes (N) and presence of metastasis (M)
ToT	Time on Treatment
TRAE	Treatment-related adverse event
TSD	Technical Support Document
TTR	Time to response
TTS	Time to surgery
UK	United Kingdom
ULN	Upper limit of normal
UMC	University Medical Centre
UOB	Urological Oncology Branch
US	United States
USA	United States of America
VEGF	Vascular endothelial growth factor
VHL	Von Hippel-Lindau
WTP	Willingness-to-pay
X	Label to signify a cancer that cannot be measured (in main tumour, nearby lymph nodes or metastasis)

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Assessment Group (EAG) as being potentially important for decision making. If possible, it also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem (DP), Section 1.4 issues related to the clinical effectiveness evidence, and Section 1.5 issues related to the cost effectiveness evidence. Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Subsequent sections of the main EAG report provide further detail: 2 (decision problem), 3 (clinical effectiveness evidence), 4 (cost effectiveness methods), 5 (cost effectiveness results), 6 (EAG's additional analyses) and 7 (end of life criteria).

All issues identified represent the EAG's view, and not those of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1.1: Summary of key issues

ID3932	Summary of issue	Report Sections
1	DP: implication of differences between intervention and comparator populations given interpretation of the MA that SoC for most patients is immediate surgery	2.1, 2.2, 2.3, 4.2.4, 4.2.6
2	DP: misalignment between the DP and MK-6482-004 study populations; and between the latter and the UK target population	2.1, 2.3, 3.2.2
3	Clinical effectiveness SLR: potential risk of study selection bias resulting in possible omission of relevant comparator studies	3.1.1, 3.1.2, 3.2.1
4	Problems with nature of comparators; lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator	3.3
5	Limitations in the ITC hinder the assessment of the effectiveness of Belzutifan compared to SoC	3.4
6	There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses	2.1, 4.2.2, 4.2.3, 4.2.4, 4.2.6.5 & 4.2.6.7
7	The comparator data might not be representative for the UK	4.2.6.4
8	Data to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable	4.2.6 & 5.3.2
9	There is uncertainty in the derivation of the transition probabilities in the SoC arm	4.2.6.2
10	There is uncertainty in the implementation of time on treatment and treatment effect waning	4.2.6.8
11	There is uncertainty in the derivation and implementation of HRQoL in the model	4.2.8.1
12	Cost-effectiveness analyses should be based on subgroup-specific parameters (including QALY severity weighting)	4.2.3, 4.2.10, 5.1, 5.2 & 6.2

ID3932	Summary of issue	Report Sections
DP = decision problem; HRQoL = health-related quality of life; ITC = indirect treatment comparison; MA = marketing authority; QALY = quality-adjusted life year; SLR = systematic literature review; SoC = standard of care; UK = United Kingdom		

The EAG was unable to define a new base-case and thus to select preferred assumptions. The EAG considered that the majority of the uncertainties identified in the company submission (CS) cannot be resolved with the current evidence. The EAG believes that any alternative base-case scenario that could have been presented, would still be subjected to too many uncertainties and its results would thus be unreliable. Therefore, the EAG is afraid that a new base-case could give the wrong impression that it would be appropriate for the current decision problem.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Decreasing the risk of surgery and, therefore, surgery-related complications.
- Decreasing the risk of metastatic disease.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to standard of care (SoC).
- Decreasing costs associated to surgery and surgery-related complications.
- Decreasing costs associated to treating metastatic disease.

The modelling assumptions that have the greatest effect on the ICER are:

- Utility in the non-metastatic health states.
- The proportion to receive immediate surgery in the SoC arm.
- The removal of treatment effect waning.
- Changes in perioperative mortality risks and in risks of risks of short- and long-term complications following surgery.
- The distribution chosen to model time to surgery for the Belzutifan arm.
- The distribution chosen to model Belzutifan time on treatment.

1.3 The decision problem: summary of the EAG's key issues

The EAG noted some areas of potential misalignment between: intervention and comparator populations as described in the DP (Table 1.2); the DP and study populations (Table 1.3); and the study and United Kingdom (UK) target populations (Table 1.3).

Table 1.2: Key issue 1: Decision problem: implication of differences between intervention and comparator populations given interpretation of the MA that SoC for most patients is immediate surgery

Report Section	2.1, 2.2, 2.3, 4.2.4, 4.2.6
Description of issue and why the EAG has identified it as important	The population in the DP is essentially the same as that in the MA, which states that patients :*" require therapy*" and *"localised procedures are unsuitable or undesirable".* The company clarify what is meant by localised procedures: "There are no medical treatment options approved or funded in the UK at the point in which Belzutifan is indicated. Localised procedures are used, though they should be considered 'last resort' interventions." In the

Report Section	**2.1, 2.2, 2.3, 4.2.4, 4.2.6**
	CEA they are represented by immediate surgery for most (90% of VHL RCC and VHL pNET and 50% of VHL CNS Hb) patients. With Belzutifan there is no immediate surgery, which the EAG considered could leave patients suffering harm if they are as implied requiring immediate surgery. The company argued that no immediate surgery is required because patients: *"…are receiving a treatment shown to reduce these two types of harms [risk of metastatic disease due to tumour growth (for RCC and pNET) or symptomatic burden (in all cohorts but particularly in CNS Hb)."* However, the more immediate harm does seem to be substantial, as stated by the company: *"In current UK clinical practice, patients undergo surgery where organ function will be* ***significantly impaired or completely cease****, or where there will be* ***significant risk to neurological function for CNS lesions****, as they are the only treatment option available to* ***keep patients alive…"*** (p. 36, CS). Note also that not all patients will respond to Belzutifan, and that response takes time (in the Belzutifan trial for VHL RCC, 63.9% and median of 11 months). The EAG therefore consider that the populations implied by immediate surgery for the comparator and no immediate surgery for the interventions appear to be incompatible.
What alternative approach has the EAG suggested?	The EAG have stated that the intervention and comparator populations should be identical. Therefore, the percentage of patients receiving immediate surgery required to reduce any immediate severe harm should be the same for both comparator and intervention. Therefore, as modelled by the company, no patients eligible for Belzutifan would receive such surgery then no such immediate surgery should be administered as part of the comparator. Given that the EAG cannot separate surgery that is purely prophylactic from that required to reduce immediate severe harm, the EAG have assumed no immediate surgery for the comparator. Effectively, this would mean that the comparator becomes active surveillance. Note also that the model is still informed by rates of surgery under active surveillance that are much higher than those for Belzutifan.
What is the expected effect on the cost effectiveness estimates?	The ICER will increase considerably if there is greater parity between comparator and intervention in the rates of surgery.
What additional evidence or analyses might help to resolve this key issue?	The company need to clarify whether patients eligible for Belzutifan need immediate surgery. If it is the case that no immediate surgery is required with Belzutifan then, if the company still believes that some immediate surgery would be SoC for patients who would be eligible for Belzutifan, the company need to provide evidence as to the percentage who would receive surgery where the harm that would occur without that surgery would be so insubstantial as to be able to risk forgoing it whilst waiting for a possible Belzutifan response.
CEA = cost-effectiveness analysis; CNS Hb = central nervous system haemangioblastoma; CS = company submission; DP = decision problem; EAG = Evidence Assessment Group; ICER = incremental cost effectiveness ratio; MA = marketing authorisation; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SoC = standard of care; UK = United Kingdom; VHL = Von Hippel-Lindau	

Table 1.3: Key issue 2: Decision problem: misalignment between the DP and MK-6482-004 study populations; and between the latter and the UK target population

Report section	2.1, 2.3, 3.2.2
Description of issue and why the EAG has identified it as important	The MK-6482-004 study population is narrower than that of the DP in terms of tumour type (must have ≥1 RCC as opposed to must have ≥1 among RCC or CNS haemangioblastoma or pNET respectively). In addition, it is likely that at least some patients recruited to the MK-6482-004 study had less severe disease compared to those in the DP population. This presents challenges in generalising the findings of the MK-6482-004 study to the target UK population. The lack of data on patient subgroups defined according to tumour type or combination of tumour types meant that it was not feasible to judge the comparability between the MK-6482-004 study and UK target populations.
What alternative approach has the EAG suggested?	None.
What is the expected effect on the cost effectiveness estimates?	None if the DP population is adapted to better align with the MK-6482 study, notwithstanding any change due to addressing Key Issue 1.
What additional evidence or analyses might help to resolve this key issue?	The DP population could be narrowed to only include patients who all have an RCC. It could also be narrowed to identify patients at a similar severity to that in the MK-6482 study, including with regards to need for surgery (see Key Issue 1).

CNS = central nervous system; DP = decision problem; EAG = Evidence Assessment Group; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; UK = United Kingdom

1.4 The clinical effectiveness evidence: summary of the EAG's key issues

The company reported the methods and results of a systematic literature review (SLR) to identify clinical effectiveness evidence relating to Belzutifan. The EAG noted the potential for study selection bias, particularly in relation to studies providing comparator data. In addition, there was an absence of data on overall survival (OS), an outcome specified in the NICE Final Scope. The EAG also noted limitations in terms of the description of sources of comparator data and the conduct of the indirect treatment comparison (ITC).

Table 1.4: Key issue 3: Clinical effectiveness SLR: potential risk of study selection bias resulting in possible omission of relevant comparator studies

Report Section	3.1.1, 3.1.2, 3.2.1
Description of issue and why the EAG has identified it as important	The clinical effectiveness SLR identified 26 records that were initially included in the review. However, only one of these was included in the submission (the MK-6482-004 study) as it was the only one to investigate the clinical effectiveness of Belzutifan. It is not clear whether the remaining 25 records could have contributed comparator data. Related to this, it was not clear whether the search strategy and study selection criteria were designed to identify all relevant interventional, non-interventional and natural history studies.
What alternative approach has the EAG suggested?	Provide greater transparency in relation to the existing SLR methods so that the approaches used to identify comparator data can be fully evaluated. Alternatively, conduct another SLR

Report Section	3.1.1, 3.1.2, 3.2.1
	designed to retrieve all relevant interventional, non-interventional and natural history studies.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	See above.
EAG = Evidence Assessment Group; SLR = systematic literature review	

Table 1.5: Key issue 4: Problems with nature of comparators; lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator

Report Section	3.3
Description of issue and why the EAG has identified it as important	The VHL Natural History Study was used as the only source of comparator data in the clinical effectiveness section. It was unclear why this was the only study considered, given the use by the company of two other datasets: the pre-treatment phase of MK-6482-004 to inform rates of pre-surgery-> surgery; and the Optum Clinformatics Data Mart Claims Study, which was used in the CEA to adjust these rates on the basis that this dataset would align better with UK clinical practice than the VHL Natural History Study, which was US-based. In their response to the CL, the company indicated the superiority of the pre-treatment phase of MK-6482-004 as a source of comparator data and it is unclear why this data source was not used to estimate all TTS outcomes with full reporting in the clinical effectiveness section. It is also unclear why the Optum Clinformatics Data Mart Claims Study was not used in the ITC given its potentially greater applicability to UK clinical practice. The EAG noted limitations to all three studies in terms of using them for comparator data versus the data from MK-6482-004.
What alternative approach has the EAG suggested?	A comparison of all three data sources should have been presented as part of the clinical effectiveness evidence with the potential for all three to contribute to the ITC assessed in terms of all relevant time to event outcomes.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	As above.
CEA = cost-effectiveness analysis; CL = clarification letter; EAG = Evidence Assessment Group; ITC = indirect treatment comparison; TTS = time to surgery; UK = United Kingdom; US – United States; VHL = Von Hippel-Lindau	

Table 1.6: Key issue 5: Limitations in the ITC hinder the assessment of the effectiveness of Belzutifan compared to SoC

Report Section	**3.4**
Description of issue and why the EAG has identified it as important	The EAG considers that the ITC performed by the company is limited in that the company have not provided adequate justification for the following aspects: • method of adjustment for confounding, which did not include the use of IPD from both the MK-6482-004 trial and the VHL Natural History study, but only for the latter • choice of confounding characteristics, prognostic or treatment effect modifying, for which no objective evidence was provided • choice of outcomes for which adjustment has been performed, which only included cause-specific hazards of pre-surgery → 1st surgery non-RCC VHL-related surgeries with therapeutic intent
What alternative approach has the EAG suggested?	• Method of adjustment for confounding could include the use of IPD from both the MK-6482-004 trial and the VHL Natural History study • Choice of confounding characteristics, prognostic or treatment effect modifying, should include objective evidence • Choice of outcomes for which adjustment has been performed should include cause-specific hazards for pre-surgery->metastatic disease and pre-surgery->death for all three cohorts
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	As stated above.
EAG = Evidence Assessment Group; IPD = individual participant data; ITC = indirect treatment comparison; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau	

1.5 The cost effectiveness evidence: summary of the EAG's key issues

A full summary of the cost effectiveness (CE) evidence review conclusions can be found in Section 6.4 of this report. The company's CE results are presented in Section 5, the EAG's summary and detailed critique are in Section 4, and the EAG's exploratory results are presented in Section 6. The key issues in the CE evidence are discussed in Tables 1.7 to .1.13.

Table 1.7: Key issue 6: There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses

Report Section	**2.1, 4.2.2, 4.2.3, 4.2.4, 4.2.6.5 & 4.2.6.7**
Description of issue and why the EAG has identified it as important	There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses. This mismatch could be broadly categorised in two types of issues:

Report Section	2.1, 4.2.2, 4.2.3, 4.2.4, 4.2.6.5 & 4.2.6.7
	a) Type of primary tumour. The model distinguishes three cohorts by tumour type, which is not possible with the current evidence. b) Severity of the patient population. The population included in the economic analyses was as adult patients with VHL disease who require treatment for VHL-associated RCC, VHL-associated CNS Hb, or VHL associated pNET for whom surgery is unsuitable or undesirable. The EAG identified the following issues: - No evidence seems to be available for patients "for whom surgery is unsuitable or undesirable". - Surgery rates observed in the MK-6482-004 trial (Belzutifan) might underestimate the surgery rates for the population in the decision problem. - The percentages of immediate surgery in the SoC arm seem arbitrary and have a major impact on the model results. - The potential "harm" for Belzutifan patients for not having immediate surgery is not captured in the model. The potential "benefit" for SoC patients for having immediate surgery is not captured in the model. - Doubling the perioperative mortality risk in the three cohorts seems arbitrary and has a major impact on the model results. - Increasing risks of short- and long-term complications following surgery seems arbitrary and has an impact on the model results.
What alternative approach has the EAG suggested?	Not all uncertainties mentioned above could be explored by the EAG due to lack of data. Alternative transition probability from pre-surgery to surgery in the Belzutifan arm. Alternative percentages of patients undergoing immediate surgery in the SoC arm. Alternative assumptions on perioperative mortality risks, and other surgery-related complications.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Clinical effectiveness data on the decision problem population, allowing classification by primary tumour. Evidence supporting the definition of SoC (immediate surgery, risks associated to surgery, etc.). The potential "harm" for Belzutifan patients for not having immediate surgery and the potential "benefit" for SoC patients for having immediate surgery should be captured in the model.

CNS Hb = central nervous system haemangioblastoma; EAG = Evidence Assessment Group; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Table 1.8: Key issue 7: The comparator data might not be representative for the UK

Report Section	4.2.6.4
Description of issue and why the EAG has identified it as important	Patients in the VHL Natural History Study may have received an elevated SoC compared to UK practice. This issue was addressed by the company by making: - Adjustments on transition probabilities based on Optum Clinformatics Data Mart database may bring additional uncertainty rather than reducing it. There is uncertainty surrounding this adjustment. - There are inherent uncertainties imposed by using data from a different clinical practice (the US versus the UK) that cannot be resolved.
What alternative approach has the EAG suggested?	Not all uncertainties mentioned above could be explored by the EAG due to lack of data. Omit the adjustment in the risk of surgery and metastasis based on Optum Clinformatics Data Mart data.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Data from UK clinical practice.
EAG = Evidence Assessment Group; SoC = standard of care; UK = United Kingdom; US = United States; VHL = Von Hippel-Lindau	

Table 1.9: Key issue 8: Data to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable

Report Section	4.2.6 & 5.3.2
Description of issue and why the EAG has identified it as important	Data in the MK-6482-004 trial (Belzutifan) is extremely immature for the three cohorts, especially for the VHL pNET cohort and to a lower extent for the VHL CNS Hb cohort.
What alternative approach has the EAG suggested?	Assume alternative fitted parametric models for time to surgery, metastases, or death in all cohorts in both arms, but non-quantifiable uncertainties remain.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Long-term survival data for all three cohorts.
CNS Hb = central nervous system haemangioblastoma; EAG = Evidence Assessment Group; pNET = pancreatic neuroendocrine tumour; VHL = Von Hippel-Lindau	

Table 1.10: Key issue 9: There is uncertainty in the derivation of the transition probabilities in the SoC arm

Report Section	4.2.6.2
Description of issue and why the EAG has identified it as important	Pre-treatment data from the MK-6482-004 trial were used to inform transitions from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm. This approach may

Report Section	4.2.6.2
	be subject to biases due to the different data sources used to define transitions within the same cohort (pre-treatment period of the MK-6482-004 trial and the VHL Natural History Study).
What alternative approach has the EAG suggested?	Not all uncertainties could be explored by the EAG due to lack of data. Alternative transition probability from pre-surgery to surgery in the SoC arm.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Clinical effectiveness data on the decision problem population, allowing classification by primary tumour (SoC arm).
EAG = Evidence Assessment Group; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau	

Table 1.11: Key issue 10: There is uncertainty in the implementation of ToT and treatment effect waning

Report Section	4.2.6.8
Description of issue and why the EAG has identified it as important	Due to data immaturity, there is uncertainty regarding the choice of the parametric distribution to model ToT and the duration of Belzutifan residual benefit.
What alternative approach has the EAG suggested?	Explore different parametric models used to define ToT and alternative options for the duration of the residual treatment benefit.
What is the expected effect on the cost effectiveness estimates?	Unknown
What additional evidence or analyses might help to resolve this key issue?	Belzutifan long-term data.
EAG = Evidence Assessment Group; ToT = time on treatment	

Table 1.12: Key issue 11: There is uncertainty in the derivation and implementation of HRQoL in the model

Report Section	4.2.8.1
Description of issue and why the EAG has identified it as important	There is a mismatch between the decision problem and the evidence used to inform HRQoL (VHL RW QoL Disease Burden Study) in terms of population: - In the VHL RW QoL Disease Burden Study, HRQoL is reported by VHL patients but it is not limited to patients who require therapy and "for whom localised procedures are unsuitable or undesirable" and response status is self-assessed by patients.[1] - Utilities in KEYNOTE-564 and Kiebert et al. 2001 were used as proxies for patients with VHL disease to inform some utilities.[2, 3]

Report Section	4.2.8.1
	A potential bias attributable to assuming an immediate HRQoL benefit for Belzutifan patients.
What alternative approach has the EAG suggested?	Not all uncertainties could be fully explored by the EAG due to lack of data. Scenario analyses with reduction in health state utility values to represent utility in patients "for whom localised procedures are unsuitable or undesirable". Incorporate the effect of Belzutifan time to response in the QALY calculations.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Additional evidence about UK-specific VHL disease patient population, including those "for whom localised procedures are unsuitable or undesirable". If response categories are linked to HRQoL, patients need to be classified in response categories based on a clinical diagnosis by a physician (not self-reported). Additional evidence about HRQoL of patients in partial response is required.
EAG = Evidence Assessment Group; HRQoL = health-related quality of life; QALY = quality adjusted life year; QoL = quality of life; RW = real world; VHL = Von Hippel-Lindau	

Table 1.13: Key issue 12: CEAs should be based on subgroup-specific parameters (including QALY severity weighting)

Report Section	4.2.3, 4.2.10, 5.1, 5.2 & 6.2
Description of issue and why the EAG has identified it as important	Model is built to estimate Belzutifan cost effectiveness compared to SoC in three different subgroups of patients. However, subgroup-specific parameters were not used in the model.
What alternative approach has the EAG suggested?	Estimate input parameters using subgroup-specific data. This includes using subgroup-specific severity adjusted QALY weights. The latter was used by the EAG in its scenarios.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence needed, but input parameters that change per cohort should be re-estimated.
CEA = cost effectiveness analysis; EAG = Evidence Assessment Group; QALY = quality adjusted life year; SoC = standard of care	

1.6 Other key issues: summary of the EAG's view

No other key issues were identified by the EAG.

1.7 Summary of the EAG's view

Based on all key issues described above, the EAG was unable to define a new base-case. The EAG considered that the majority of the uncertainties identified in the CS cannot be resolved with the current evidence. The EAG believes that any alternative base-case scenario that could have been presented,

would still be subjected to too many uncertainties and its results would thus be unreliable. Therefore, the EAG is afraid that a new base-case could give the wrong impression that it would be appropriate for the current decision problem. Instead, additional scenario analyses were explored by the EAG in order to assess the impact of some alternative assumptions on the current CE results. The scenario analyses conducted by the EAG were mostly explorative given the lack of other sources of evidence and many of the alternative assumptions explored were arbitrarily selected. Results nevertheless indicated that the ICER in the current model was sensitive to several assumptions. As expected, the proportion of patients receiving immediate surgery in the SoC arm had a major impact on the results. Also, by comparing SoC with and without immediate surgery the EAG's idea that the severity of the decision problem population has not been appropriately captured by the company's model was reinforced. Results were also sensitive to changes in perioperative mortality risks and in risks of short- and long-term complications following surgery. This seemed to be arbitrarily defined by the company and had a major impact on the model results. Results indicated that the distribution chosen to model time to surgery for the Belzutifan arm also had major impact on the ICER, illustrating in this way the uncertainty associated to the immaturity of the data. Special care needs to be taken when assessing the results for the pancreatic neuroendocrine tumour (pNET) cohort since these might be lacking face validity. Alternative assumptions on time on treatment, duration of the Belzutifan residual treatment effect or the choice of utilities, should also be considered as sources of relevant uncertainty. It is notable that in all scenarios explored by the EAG, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained. Only in a couple of them, and for the Von Hippel-Lindau central nervous system haemangioblastoma (VHL CNS Hb) cohort only, the ICER was close to £30,000.

The EAG acknowledges the difficulty of representing the population in the DP with the current evidence, which is mostly derived from the MK-6482-004 trial for Belzutifan. Considering the above, the current model structure and the available data, the EAG is unable to change the model in a straightforward manner to account for patients requiring immediate surgery in the Belzutifan arm as assumed in SoC. In its current form, the company's model might be considered appropriate to reflect the company's initially sought marketing authorisation VHL-associated renal cell carcinoma (RCC) only and for the population recruited into MK-6482-004 trial but cannot provide reliable estimates of the CE of Belzutifan compared to SoC in the population defined in the DP given the model's inability to properly represent/capture "those patients who require therapy and for whom localised procedures are unsuitable or undesirable".

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the DP (as presented by the company)

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
Population	Adults who require therapy for renal cell carcinoma (RCC), central nervous system (CNS) haemangioblastomas (Hb), or pancreatic neuroendocrine tumours (pNETs) caused by von Hippel-Lindau (VHL) disease, for whom localised procedures are unsuitable or undesirable.	Adult patients with VHL disease who require therapy for VHL associated RCC, CNS Hb or pNETs, and for whom localised procedures are unsuitable or undesirable.	NA	Whilst the population definitions in the NICE Final Scope and the company's decision problem (DP) appear similar, several aspects of the DP population were unclear from the company submission (CS), and this triggered several clarification questions from the EAG. In their clarification response,[4] the company confirmed that the DP population excludes patients with advanced or metastatic disease. Related to this, the EAG noted that evidence of metastatic disease on screening imaging was an exclusion criterion for the MK-6482-004 study.[5] However, contraindication in terms of metastatic or advanced disease was not mentioned in the marketing authorisation (MA) details.[6] There is further discussion of this issue under "*Comparators*" (below).

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
				Elsewhere in their clarification response,[4] the company confirmed a discrepancy between the DP/MA population[6] (adults with ≥1 tumour among VHL-associated RCC or CNS hemangioblastomas or pNETs, for whom localised procedures are unsuitable or undesirable) and the MK-6482-004 study population (≥1 measurable VHL-associated RCC). The company acknowledged that the study population was narrower than that for the DP and MA. Further, the company confirmed a potential mismatch between the DP/MA[6] and study populations in terms of disease severity since localised procedures were stated as "*unsuitable or undesirable*" for the former whilst this aspect was not mentioned for the latter.[5] In the same vein, the DP/MA population was specified as "*adults who require therapy*" whilst the study excluded patients with an immediate need for surgical intervention for the tumour.[5] Within their response, the company acknowledged the

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
				implication that some study patients may have had less severe disease relative to those in the DP/MA population.[4] The EAG also asked the company whether eligibility for the MK-6482-004 study comprising "*sufficient organ function*" also constituted a misalignment between the DP/MA and the study. The company replied that there was no misalignment between the DP and study populations however, the rationale underpinning this statement was not clear.[4] The company stated that whilst tumour types other than RCC, CNS haemangioblastomas and pNETs (i.e., non pNET pancreatic lesions; retinal haemangioblastomas; adrenal lesions; endolymphatic sac tumours; and epididymal cystadenomas) were relevant to the DP, they were not mentioned therein because eligibility for Belzutifan is determined by presence of an RCC, CNS haemangioblastoma or pNET and

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
				the MA does not mention tumour types other than these three.[4, 6]
Intervention	Belzutifan	Belzutifan	NA	The intervention is in line with the NICE Final Scope
Comparator(s)	RCC: • Standard of care (SoC) without Belzutifan • For advanced or metastatic disease, monotherapy or combination therapy with immunotherapies or kinase inhibitors CNS hemangioblastomas: • SoC with Belzutifan (confirmed with NICE 14.06.2023 that this should read as "*SoC without Belzutifan*") pNETs: • SoC without Belzutifan • For unresectable or metastatic disease, monotherapy with lutetium (177Lu) oxodotreotide or combination therapy with everolimus and sunitinib	For VHL associated RCC, pNET, and CNS hemangioblastomas: • Current SoC ***without*** Belzutifan. There are no medical treatment options approved or funded in the UK at the point in which Belzutifan is indicated. Localised procedures are used, though they should be considered "last resort" interventions.	The relevant comparators are: • Primary tumour RCC or pNET: surgery resulting in loss of organ function • Primary tumour CNS hemangioblastoma: surgery with risk of problematic brain injury, or do nothing and risk problematic brain injury No treatments for advanced or metastatic disease are relevant as comparators as these would be used after treatment with Belzutifan. The purpose of Belzutifan is to **prevent** tumours reaching the advanced or metastatic stage. Treatments for metastatic disease are included as subsequent treatments in the economic model.	The EAG requested clarification on comparators both generally and in relation to patient subgroups based on tumour type and combinations of different tumour types. In their response to clarification,[4] the company confirmed that the comparator comprises surgery or active surveillance across all tumour types. The company did not provide the proportions of patients receiving the two intervention types as standard of care (SoC) in the UK per subgroup based on tumour type and combination of different tumour types, as requested by the EAG. However, the company did provide further details of "*localised procedures*" as applied to the above subgroups. When asked for further information about comparators for patient subgroups with more

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
				than one tumour type, the company replied that treatment decisions were driven by the primary tumour.[4]
Outcomes	The outcome measures to be considered include: • overall survival (OS) • progression-free survival (PFS) • response rates (RRs) • tumour size reduction • reduction in number of surgical interventions • adverse effects of treatment • health-related quality of life (HRQoL)	The following outcomes were collected as part of the MK-6482-004 study: • RRs • reduction in number of surgical interventions • adverse effects of treatment • PFS • tumour size reduction	OS was not a designated predefined outcome in the MK-6482-004 trial. HRQoL data were also not collected as part of the MK-6482-004 study. OS and HRQoL are considered in the cost-effectiveness analyses, derived from sources other than the MK-6482-004 study.	The choice of outcomes appears to be driven by what was available in the MK-6482-004 study and does not fully address the NICE Final Scope.
Economic analysis	• The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year (QALY). • The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the	No comment from the company	No comment from the company	The CEAs partly complied with the NICE reference case. Deviations from the NICE reference case related to the source of data for measurement and valuations of changes in HRQoL. HRQoL data were collected from three different sources, namely the VHL RW QoL Disease Burden Study,[1] KEYNOTE-564 and Kiebert et al. 2001.[2] These three studies reported utility values of patients from a population different to the

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
	technologies being compared. • Costs will be considered from an NHS and Personal Social Services perspective. • The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.			one defined in the decision problem, and only the VHL RW QoL Disease Burden Study was conducted for the VHL disease. Furthermore, it is unclear if the populations in these studies are representative for the UK population since the company was unable to provide a comparison to the UK patient characteristics.
Subgroups to be considered	None specified.	No comment from the company	No comment from the company	The EAG noted the potential for patient subgroups based on tumour type or combination of tumour types. and asked the company to provide the proportions of patients in the DP and UK clinical practice for the following: RCC only; CNS only; pNET only; RCC+CNS; RCC+pNET; CNS+pNET; and RCC+CNS+pNET. The company did not provide this information, referring to "*a lack of published data on the proportion of eligible patients with VHL disease in UK clinical practice specifically with RCC only, CNS only, pNET only, CNS+RCC, pNET+RCC, CNS+pNET, and all three.*"[4] The company indicated that subgroup analyses were not

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
				conducted as part of the economic evaluation (see Section B.3.12 of the CS).[5] However, cost-effectiveness results are presented for three different subgroups of patients. These subgroups are defined as patients with RCC, patients with CNS Hb and patients with pNETs. On page 255 of the CS, the company also mentioned that ", *these VHL cohorts should not be strictly considered as discrete subgroups perfectly distinct from one another as clinical trial subgroups often are. Nor should any VHL cohort be considered perfectly representative of a patient group afflicted with a specific VHL-associated primary manifestation. For example, a patient with pNET as the primary VHL tumour manifestation may have a diverse mix of VHL-associated sequelae that may overlap with other VHL cohorts over the course of their live*".[5] This is problematic for the economic analyses given that the comparator differs per subgroup and cost-effectiveness results have been presented per subgroup.

	Final Scope issued by NICE	DP addressed in the CS	Rationale if different from the final NICE scope	EAG comment
Other considerations including issues related to equity or equality	Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	No comment from the company	No comment from the company	The company did not include any statements within the DP but comments relating to equality and equity were provided elsewhere in the CS, highlighting potential issues around the inherited nature of the disease (meaning that some families would be disproportionately affected), the impact of the disease on younger people and inequities relating to service provision.[5]

Based on Table 2 of the CS.[5]

CEA = cost-effectiveness analysis; CNS = central nervous system; CS = company submission; DP = decision problem; EAG = Evidence Assessment Group; Hb = haemangioblastoma; HRQoL = health-related quality of life; MA = marketing authorisation; N/A = not applicable; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; OS = overall survival; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; QoL = quality of life; RCC = renal cell carcinoma; RR = response rate; RW = real world; SoC = standard of care; UK = United Kingdom; VHL = Von Hippel-Lindau (disease)

2.1 Population

The population defined in the National Institute for Health and Care Excellence (NICE) Final Scope is: "*Adults who require therapy for renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNETs) caused by von Hippel-Lindau (VHL) disease, for whom localised procedures are unsuitable or undesirable.*"[7] The company's definition of the population within the decision problem (DP) is essentially the same as that in the NICE Final Scope and therefore no further justification is presented (see Table 2.1 above).[5] However, several aspects of the DP population were unclear from the company submission (CS)[5] including aspects of tumour site and stage and the company were asked to respond to several questions during the clarification phase accordingly, as detailed below.

Belzutifan has a Great Britain (GB) marketing authorisation (MA) that was first granted on 31 May 2022 with Medicines and Healthcare products Regulatory Agency (MHRA) number PLGB 53095/0087. The summary of product characteristics (SmPC) for Belzutifan (Welireg®) states that the drug "*is indicated for the treatment of adult patients with VHL disease who require therapy for VHL-associated RCC, CNS haemangioblastomas or pNETs, and for whom localised procedures are unsuitable or undesirable.*"[6] The Evidence Assessment Group (EAG) notes that this is essentially equivalent to the scope. The company discusses the circumstances under which "*localised procedures*" could be considered as "*unsuitable or undesirable*" under the "*Treatment pathway*" section of the CS (p.35 of Document B)[5] as follows:

- *"In VHL-associated RCC, when the localised procedure would render the patient renal replacement therapy-dependent.*
- *In VHL-associated pNET, when the localised procedure would lead to loss of pancreatic function leading to lifelong pancreatogenic diabetes and being immune-compromised such that the patient will require lifelong insulin therapy, antibiotic therapy and/or pancreatic enzyme insufficiency impacting digestion.*
- *In VHL-associated CNS hemangioblastoma, when the localised procedure could lead to severe neurological or neuromuscular deficits equating to severe permanent disability. This most often arises with tumours located in the brainstem where they are difficult to access or operate on without damaging important nearby tissues, potentially leading to significant morbidity and death."*[5]

EAG comment:

Several clarification questions related to the description of the population and referred to the apparent misalignment between the DP and the included MK-6482-004 study, the single study identified by the company as providing clinical effectiveness information on Belzutifan.[5]

Whilst the DP and MA population is defined as adult patients who require therapy for Von Hippel-Lindau (VHL)-associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas (Hb) or pancreatic neuroendocrine tumour (pNETs), and for whom localised procedures are unsuitable or undesirable, the EAG noted the patient eligibility for the MK-6482-004 study as "*Patients with VHL disease who have at least one measurable RCC tumour*" (Table 4 of Document B).[5] Given this discrepancy, the EAG asked whether the intended DP population had to have all three tumour types (RCC, CNS haemangioblastoma and pNET) or at least one of them (clarification question A10a). Related to this, the EAG also asked whether the DP population had to have an RCC tumour and if so, whether this had to be the primary tumour (clarification question A10c). The company replied that the DP/MA population comprises patients with at least one of the aforementioned tumour types. They therefore acknowledged that the MK-6482-004 study population was narrower than that

for the DP and MA given that eligible patients had to have an RCC. The company also state that (in their response to clarification questions A.10b and A.26): *"The primary tumour is the tumour that is driving treatment decisions."*, and use this as a motivation for three economic models, one for each of the three main tumour types, despite all of the CNS Hb and pNET patients in the MK-6842-004 also having an RCC tumour.[4]

Table 12 of Section B.2.4 and Section B.2.7 ("*Other tumours*") mention participants recruited to the MK-6482-004 study having tumours at sites other than those mentioned above: i.e., non pNET pancreatic lesions; retinal Hb; adrenal lesions; endolymphatic sac tumours; and epididymal cystadenomas.[5] The EAG asked the company to clarify whether participants who had these tumour types in addition to RCC, CNS Hb and/or pNETs were relevant to the DP (clarification question A11). The company replied that whilst these "*Other tumours*" were relevant to the DP, they were not mentioned therein because eligibility for Belzutifan is determined by presence of an RCC, CNS haemangioblastoma or pNET and the MA does not mention tumour types other than these three.[4, 6]

The DP population was described as those who "*require therapy*" for RCC, CNS Hb or pNETs whereas a participant exclusion criterion for MK-6482-004 was "*patients who had an immediate need for surgical intervention for tumour treatment*".[5] The company were asked to clarify whether 'not requiring immediate surgery' was the same as 'not requiring therapy' (clarification question A12a). In addition, the DP population were described as those "*for whom localised procedures are unsuitable or undesirable*"[5] whereas this aspect was not mentioned for MK-6482-004. The EAG asked the company whether 'therapy' and 'immediate' surgery were synonymous, and whether these misalignments meant that the patients recruited to MK-6482-004 were at an earlier and less severe disease stage compared to the DP population (clarification questions A12b and A23). Within their response, the company clarified that the two terms were not equivalent, and acknowledged the implication that some participants in the MK-6482-004 may have had less severe disease relative to those in the DP/MA population.[4]

Several points in the CS mentioned adequate organ function as a participant inclusion criterion for the MK-6482-004 study (e.g., p.45 of Document B provides a specific definition)[5] but it was unclear how this related to the DP population. The EAG asked the company to clarify whether the statement "*Patients must have sufficient organ function (as described in the MK-6482-004 study participant eligibility criteria) to be eligible to receive Belzutifan"* (p.37 of Document B)[5] represented a further misalignment with the DP population (clarification question A12c). The company replied that there was no misalignment between the DP and study populations however, the rationale underpinning this statement was not clear.

There is further consideration of the potential misalignment between the DP and study populations in Section 3.2.

In Table 2 of Document B, it is stated that (bold emphasis is as it appears in the source document): "*No treatments for advanced or metastatic disease are relevant as comparators as these would be used after treatment with Belzutifan. The purpose of Belzutifan is to* ***prevent*** *tumours reaching the advanced or metastatic stage.*"[5] The EAG asked whether the population in the DP should be re-expressed as excluding those with advanced or metastatic stage disease (clarification questions 9a and 9b). In their response, the company confirmed that the DP population excludes patients with advanced or metastatic disease.[4] Related to this, the EAG noted that evidence of metastatic disease on screening imaging was a participant exclusion criterion for the MK-6482-004 study.[5] However, contraindication for Belzutifan to treat metastatic or advanced disease was not mentioned in the MA details.[6] There is further discussion of this issue in Section 2.3.

In summary, the EAG notes misalignment between the DP and MK-6482-004 study populations in that the study population is narrower than that of the DP in terms of tumour type (must have ≥1 RCC as opposed to must have ≥1 among RCC or CNS Hb or pNET respectively). In addition, it is likely that at least some patients recruited to the MK-6482-004 study had less severe disease compared to those in the DP population. This presents challenges in generalising the findings of the MK-6482-004 study to the target population and in terms of baseline imbalance for treatment comparison within the evidence synthesis of the submission (discussed further in Section 3.4). This is therefore a key issue.

2.2 Intervention

The DP intervention (Belzutifan) is in line with that of the NICE Final Scope.[7]

The recommended dose of Belzutifan is 120 mg (three 40 mg tablets) administered orally once daily, with or without food with tablets swallowed whole. The SmPC information recommends that treatment should continue until disease progression or unacceptable toxicity occurs.[6]

Dose modification may be required because of adverse events (AEs) such as anaemia, hypoxia and other AEs. For Grade 2 (hypoxia) or Grade 3 AEs (all types), this may involve withholding Belzutifan until resolution of signs and symptoms and then reinstating the drug at a reduced dose. Permanent withdrawal of Belzutifan is recommended for Grade 4 (life-threatening) hypoxia.[6]

EAG comment:

The DP intervention is in line with the NICE Final Scope.

2.3 Comparators

The comparator definition within the NICE Final Scope[7] was "*Standard of care (SoC) without Belzutifan*" for RCC and pNETs. The scope also specified comparators for advanced or metastatic disease in RCC ("*…monotherapy or combination therapy with immunotherapies or kinase inhibitors*") and unresectable or metastatic disease in pNETs ("*…monotherapy with lutetium (177Lu) oxodotreotide or combination therapy with everolimus and sunitinib*"). The comparator definition for CNS Hb was described as "*SoC with belzuifan*". However, NICE confirmed with the EAG on 14 June 2023 that this should read as "*SoC without Belzutifan*". There was no mention of comparators for CNS Hb in terms of advanced, unresectable or metastatic disease.[7]

In the DP,[5] the company defined the comparator for all three tumour types as "*Current SoC without Belzutifan*" and also stated that: "*There are no medical treatment options approved or funded in the UK at the point in which Belzutifan is indicated. Localised procedures are used, though they should be considered 'last resort' interventions.*" Table 1 of the CS defines "*localised procedures*" as all non-systemic (i.e., non-pharmacological) interventions including radiotherapy, radiofrequency ablation, thermo-ablation, cryoablation, microwave ablation, irreversible electroporation, any other image-guided ablation and all surgical procedures. In their DP table rationale statements, the company mentioned that surgery was the relevant comparator for all three tumour types whilst outlining the potential risks of such interventions. The company stated furthermore that treatments for metastatic or advanced disease were irrelevant as comparators because these would be used after treatment with Belzutifan, the purpose of the latter being to prevent progression to metastatic or advanced disease.[5] This has been discussed earlier (Section 2.1). On p.131 of the CS, it is stated that *"In routine clinical practice, the decision point for a patient meeting the criteria of Belzutifan eligibility would have three options: 1) surgery that is unsuitable or undesirable because it results in loss of organ function, 2) active surveillance to monitor a tumour that is above 3cm (RCC) or 2cm (pNETs) and therefore there is an increased risk of metastatic disease and/or other symptoms of tumour burden (particularly in CNS Hb tumours), or 3) Belzutifan".*[5] Therefore, the EAG requested clarification as to the proportions of

patients who would be eligible for each of (1) and (2) by tumour type combination, given the DP population, which includes all combinations. The company responded that there is in effect a dilemma for patients: *"…facing a choice of having surgery that will have significant and severe consequences or not having surgery and increasing the risk of the cancer becoming advanced/metastatic (RCC/pNET) or increasingly symptomatically burdensome (CNS Hb)."*[4]The company clarified that, because the DP population is those who have progressed to a stage where surgery is required, the term 'active surveillance' is misleading and only refers to a small proportion of patients where *"…are no reasonable treatment options even after active surveillance has revealed the presence of a tumour that should be resected."*[4] This fits with how the comparator is expressed in the economic model (see Section 4) in that, for VHL RCC and VHL pNET cohorts, immediate surgery is assumed for 90% of patients. For VHL CNS Hb, immediate surgery is assumed for 50% of patients, but the outcomes associated with surgery is assumed for 100% of the cohort, explained by: *"…tumour burden creating neurological disability for the remaining 50%"*.[5] Note that there is no immediate surgery if Belzutifan is administered.

Taking the population, intervention, and comparator DP descriptions into account, the provided information is confusing. On the one hand, it is stated that, although localised procedures (which include surgery) are deemed unsuitable or undesirable for the DP population, the comparator description indicates that surgery must be delivered immediately. On the other hand, despite immediate surgery being required, if Belzutifan is given, there is no immediate surgery, which the EAG put to the company in the clarification letter, could lead to patients suffering harm, given that not all patients will respond to Belzutifan (only 63.9% have a clinically meaningful response with RCC and take a median of 11 months for that to occur – see Section 3.2.6). In their response to clarification question B4e, the company stated that: *"No Belzutifan-treated patients receive immediate surgery given they have recourse to an effective therapy which provides an alternative to surgery. Therefore, no Belzutifan-treated patient is assumed to be left to suffer the harm of not receiving surgery as they are receiving a treatment shown to reduce these two types of harms…risk of metastatic disease due to tumour growth (for RCC and pNET) or symptomatic burden (in all cohorts but particularly in CNS Hb)."*[4] However, the company state that the immediate surgery, described as 'localised procedure', is not just intended to prevent long term harm, but that which is more immediate: *"In current UK clinical practice, patients undergo surgery where organ function will be significantly impaired or completely cease, or where there will be significant risk to neurological function for CNS lesions, as they are the only treatment option available to keep patients alive, or prevent symptomatic disease progressing to the point where the severe sequelae of such procedures are on-balance preferable, or prevent the patient developing advanced or metastatic disease."* (p.36 of the CS).[5] The EAG is concerned that this implies differences in the characteristics of populations between the intervention and comparator groups. This suggests that surgery in the DP population is essential immediately, which is compatible with the comparator description, but not with that of the intervention, unless significant short-term harm is assumed for probably a large proportion of those patients, whilst waiting to see if Belzutifan works. The implication is that there is mismatch between the population of patients eligible for Belzutifan, which ought to be the DP population, and the current interpretation of the DP population and thus the comparator. This is therefore a key issue. Given that the population needs to be those eligible for Belzutifan and it seems implausible that it could be given to patients who need immediate surgery to reduce or avoid substantial immediate harms, including 'keeping patients alive', the most appropriate way of negating the mismatch is to remove that kind of immediate surgery from the comparator. Theoretically, there might be a small proportion of patients who need immediate surgery that could be forgone without any substantial harm, but the EAG has been presented with no evidence for this.

The EAG noted other aspects of standard of care (SoC) mentioned elsewhere in the CS that do not feature in the DP statements.[5] For example, active surveillance at clinical genetic centres or endocrinology services is mentioned in Section B.1.3 (pages 27 and 33). Specific recommendations about methods of surveillance per tumour type are described on p.33 of the CS[5] which cites a report of a national audit of VHL disease in the United Kingdom (UK).[8] According to the company, the aim of active surveillance is to ascertain whether disease-related thresholds have been observed to indicate requirement for therapy (listed in Table 1 of the CS): RCC tumours greater than 3 cm in diameter; pNETs greater than 2 cm in diameter; or CNS Hb causing symptoms that require intervention. Section B.1.3 (pages 34 to 35) also mentions another aspect of treatment in relation to patients for whom surgery is deemed unsuitable or undesirable namely, "*best alternative care*". This is described as "*a highly varied sequence of interventions*" (p.34 of the CS) but no further details of this are provided.[5]

EAG comment:

The EAG asked whether the comparator consists of surgery (that may be considered unsuitable or undesirable because of resulting in loss of organ function) and active surveillance (clarification question A13a). The company confirmed that this was correct: "*…surgery or active surveillance does comprise our comparator arm*". Further, the company described the surgical component of the comparator as "*surgery with poor outcomes for the majority of patients with a small proportion receiving active surveillance*" (response to clarification question A13c).[4] This would seem to imply that most patients receive surgery whilst proportionately fewer receive active surveillance. Related to this, the EAG asked the company to provide estimates for the proportions of patients receiving surgery versus active surveillance in subgroups based on tumour type or combination of tumour types (clarification question A13d). The company did not provide the requested information in their response but stated that: "*For the patient population relevant to this appraisal, we would like to distinguish between patients for whom active surveillance is the most appropriate care and patients who continue to be monitored/scanned but for whom there are no (good) treatment options. For some VHL patients there are no reasonable treatment options even after active surveillance has revealed the presence of a tumour that should be resected. Those that do not have surgery progress to advanced and metastatic cancer (which is not the objective of general active surveillance). General active surveillance is a precursor to a treatment intervention however, the active surveillance termed in the statement highlighted in the initial question…describes surveillance for patients who require therapy but have no reasonable treatment options. We apologise for the confusion in the term 'active surveillance' in our CS and would like to clarify that here this refers to patients who have run out of treatment options, not those who are well-managed with regular monitoring.*"[4] This suggests that the role of active surveillance in the management of patients with VHL-associated RCC, CNS Hb or pNETs is unclear.

The EAG asked for information about comparators for patient subgroups based on tumour type or combination of different tumour types (clarification question A13b). The company replied that: "*It is dependent upon the individual's presentation and their personal circumstances. The treatment rationale we model is that there is a dominant/primary tumour that is driving treatment decisions.*" Further, company also stated that "*Nothing in the clinical trial nor in our clinician engagements suggests that patients routinely have surgeries that deal with two tumour sites at one time (e.g., CNS and renal resections in one surgery). It is plausible that a patient may have multiple tumours in one organ resected in a single surgical procedure. Therefore, while clinicians are dealing with the dominant tumour, for some patients, the other tumours continue to grow with currently available SoC procedures.*"[4]

The EAG requested clarification of the term "*localised procedures*" in relation to each patient subgroup based on tumour type/combination of tumour types (clarification question A14). The company replied

that the following would be used in clinical practice whilst mentioning (as before) that a dominant or primary tumour would inform treatment decisions:[4]

- "*For RCC these would be radiofrequency ablation, thermo-ablation, cryoablation, microwave ablation, irreversible electroporation, partial nephrectomy.*
- *For CNS hemangioblastomas these would be radiotherapy, stereotactic radiosurgery and neurosurgery.*
- *For pNETs these would be endoscopic ultrasound-guided radiofrequency ablation, enucleation, radiotherapy, Whipple's procedure, partial pancreatectomy, and radical pancreatectomy, pancreatectomy with full splenectomy.*"[4]

The EAG noted that no supporting references were cited in the CS for the above lists of localised procedures per patient subgroup.

In summary, the overall picture of the nature of relevant comparators including surgery, other localised procedures, active surveillance and best alternative care is unclear, including for the overall population and in patient subgroups based on tumour type/combination of tumour types. Specifically, the role of surgery as part of SoC was unclear and the EAG is concerned that there could be clinical differences between the intervention and comparator populations i.e., patients for whom surgery is deemed suitable (the comparator) are likely to be fitter than those for whom Belzutifan is indicated (the intervention population for whom localised procedures including surgery are unsuitable/undesirable). This has been highlighted as a key issue.

The company stated that treatment for patients in subgroups based on tumour type/combination of tumour types would be informed by the "*primary tumour*".[4] However, there was no firm definition of this term in the CS[5] or the clarification response[4] beyond statements such as: "*The primary tumour is the tumour that is driving treatment decisions*" (company's response to clarification question A10b)[4] and "*While tumours may be multi-system, i.e., multiple locations at the same time, patients often have a 'primary tumour' that drives treatment decisions (unlike other oncology therapy areas, the primary tumour we refer to in VHL for this submission is not necessarily the first tumour.)*" (p.27 of the CS).[5] The EAG noted a lack of supporting references for the arguments made.

2.4 Outcomes

The following outcomes were listed in the NICE Final Scope:[7]

- Overall survival (OS)
- Progression-free survival (PFS)
- Response rates (RR)
- Tumour size reduction
- Reduction in number of surgical interventions
- Adverse effects of treatment
- Health-related quality of life (HRQoL)

The company listed a subset of the above outcomes in the DP, focusing on those reported in the MK-6482-004 study:[5]

- PFS
- RR
- Tumour size reduction
- Reduction in number of surgical interventions
- Adverse effects of treatment

The company stated that OS was not a predefined outcome in the MK-6482-004 study and that this and HRQoL were considered in the CEA, with data derived from sources other than the aforementioned study.[5]

EAG comment:

The EAG note the discrepancy between the outcomes listed in the NICE Final Scope and the DP table. The derivation of HRQoL data from other sources is discussed further in Section 4.

2.5 Subgroups

No subgroups were specified within the NICE Final Scope[7] and the company did not provide any statements relating to this as part of the DP.[5]

EAG comment:

The EAG noted the potential for patient subgroups based on tumour type or combination of tumour types. The company were asked (clarification question A10d) to provide the proportions of patients in the DP and UK clinical practice for each of the following subgroups defined according to tumour type/combination of tumour types:

- RCC only
- CNS only
- pNET only
- RCC + CNS
- RCC + pNET
- CNS + pNET
- RCC + CNS + pNET

In their response, the company outlined the proportion of patients in subgroups of the MK-6482-004 study, all of whom had to have RCC to be included. The following proportions from MK-6482-004 were cited: 50 (82%) of participants had RCC + CNS Hb, 22 (36%) had RCC + pNETs, and 17 (28%) had RCC + CNS Hb + pNETs at baseline as confirmed by independent review committee (IRC) assessment. However, the company were not able to confirm whether this aligned with the proportion of patients in subgroups as defined within the DP or within UK clinical practice, and mentioned a lack of published data hindering estimation for the latter.[4]

EAG comment:

The proportions of patients in subgroups based on single tumour types or combination of tumour types in the DP and UK clinical practice populations remains unclear.

2.6 Other relevant factors

The company did not make any comments on considerations such as equity or equality within the DP table, but they did include statements relating to this elsewhere in the CS[5] as follows:

- "*VHL disease affects males and females and all ethnic groups equally*" (p.33 of the CS)[5]
- "*No equity or equality considerations are anticipated, although the inherited nature of the disease means some families are disproportionately impacted over multiple generations. There are inequities in the type of service available to VHL patients depending on which centre leads their care.*" (p.21 of the CS)[5]

- "*It should be noted that the onset of RCC, pNETs, and/or CNS hemangioblastomas can affect patients with VHL disease when they are very young*." (p.45 of the CS).[5]

EAG comment: None.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

In the CS (Section B.2.2), the company reported that: "*A SLR was performed to identify all relevant published and unpublished randomised controlled trials (RCTs) and non-randomised clinical trials (non-RCTs) relating to Belzutifan as per the final scope.*" The following Sections provide a summary and critique of the clinical effectiveness systematic literature review (SLR) based on the CS[5] and the company's response to clarification questions.[4]

3.1.1 Searches

The following paragraphs contain summaries and critiques of the searches related to the clinical effectiveness SLR presented in the CS.[5] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[9, 10] The CS[5] was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[11] The EAG has presented only the major limitations of each search strategy in the report.

Appendix D of the CS details the SLR undertaken to identify relevant clinical evidence for the efficacy and safety of Belzutifan for treating tumours associated with VHL disease.[5] The searches were conducted in June 2022.

A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
MEDLINE	Ovid	1/1/46-14/6/22	15/6/22
Embase	Ovid	1/1/74-14/6/22	15/6/22
CENTRAL	Ovid	May 2022	15/6/22
Conferences			
ASCO	Internet	Past 2 years	Not stated
ESMO	Internet	Past 2 years	Not stated
Trials registries			
ClinicalTrials.gov	Internet	Not stated	Not stated

Based on details in Section D.1.1 of Appendix D of the CS.[5]
ASCO = American Society of Clinical Oncology; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; ESMO = European Society for Medical Oncology.

EAG comment:

- Searches were undertaken in June 2022 to identify relevant clinical evidence for the efficacy and safety of Belzutifan for treating tumours associated with VHL disease. The CS, Appendix D and the Company's response to clarification provided sufficient details for the EAG to appraise the literature searches.[4, 5]

- A good range of databases, conferences, grey literature resources and trials registers were searched. Reference checking was conducted, and strategies utilised study design filters recommended by the Scottish Intercollegiate Guidelines Network (SIGN).
- Database searches were not limited to date or by language of publication.
- Searches were well structured, transparent and reproducible, and a good range of subject indexing terms (MeSH/EMTREE) and free text was used.
- Exact dates of searches for conferences and trial registries were not provided in the CS and were not included in the company's response to clarification.
- The EAG noted that there were no free-text search terms for natural history studies in the clinical effectiveness database searches, however data from the unpublished VHL Natural History Study were included in the CS. Although relevant subject indexing terms were included in the search strategies, additional free-text terms may have helped identify other similar studies.

3.1.2 Inclusion criteria

This and the following Sections summarise and critique the further clinical effectiveness SLR methods, as described in the CS.[5] The eligibility criteria for the SLR are presented in Table 3.2 below.

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Inclusion criteria	Exclusion criteria
Population	• Patients with von Hippel-Lindau (VHL) disease who require therapy for associated: o Renal cell carcinoma (RCC) o Central nervous system (CNS) hemangioblastomas o Pancreatic neuroendocrine tumours (pNETs)	• Patients with metastatic disease • Stage IV (M1) • Studies including a mixed population of VHL and non-VHL patients that did not report outcomes separately for the VHL subgroup*
Interventions	Any intervention studied in the population of interest including, but not limited to: • MK-6482 • Surgery	None stated.
Comparators	With respect to comparative evidence (e.g., RCTs), the following comparators are of interest: • Placebo or best supportive care • Any intervention of interest With respect to non-comparative evidence (e.g., single-arm trials, observational cohorts), no comparator treatment is required.	None stated.
Outcomes	• Overall response rate (ORR) • Duration of response (DOR) • Time to response (TTR) • Time to surgery (TTS)	• Studies reporting only surgical-related complications as an outcome* • Studies reporting only surgical-related death as an outcome*

	Inclusion criteria	Exclusion criteria
	• Progression-free survival (PFS) • Overall survival (OS) • Drug-related adverse events (AEs) • Grade 3-5 AEs (all, drug-related) • Discontinuation due to AE (DAEs) • Serious AEs (SAEs)	
Study design	• RCTs • Controlled clinical trials • Non-randomised clinical trials, including single-arm prospective interventional trials • Prospective and retrospective cohort studies	• Case-control studies • Case series/case reports • Cross-sectional studies • Systematic reviews and meta-analyses
Language restrictions	Studies published in English only	Studies published in languages other than English

Based on Table 108 of Appendix D of the CS.[5]
*Based on details in Section D.1.1 of Appendix D of the CS.[5]
AEs = adverse events; CNS = central nervous system; CS = company submission; DAE = discontinuation due to AE; DOR = duration of response; M1 = not defined in the CS but usually denotes metastatic disease that has spread to other parts of the body;[12] NICE = National Institute for Health and Clinical Excellence; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; RCT = randomised controlled trial; SAE = serious AE; TTR = time to response; TTS = time to surgery; VHL = Von Hippel-Lindau

EAG comment:

- The eligible population for the SLR is broadly in line with that described in the NICE Final Scope[7] and the DP.[5] However, discrepancies were noted in terms of interventions, comparators and outcomes. Surgery is mentioned as an eligible intervention for the SLR but is described as a comparator in the DP[5] and is not mentioned at all in the NICE Final Scope.[7] "*Best supportive care*" is stated as a comparator for the SLR (Table 3.2) but both the NICE Final Scope[7] and the DP mention "*standard of care*" (Table 2.1) whilst "*best alternative care*" is mentioned within the "*Treatment Pathway*" section of the CS (pages 34 and 35 of Section B.1.3 of the CS).[5] None of these comparators are clearly defined and it is uncertain whether they amount to the same regimen of care. There is also a lack of overlap between the SLR-eligible outcomes and those listed in the NICE Final Scope[7] and the DP.[5] The outcomes that are common to all three lists are PFS and adverse effects. In addition, both the SLR and the NICE Final Scope[7] list OS as an outcome but this is not included in the DP. Both the NICE Final Scope[7] and the DP[5] list response rates, reduction in number of surgical interventions and tumour size reduction but these do not appear within the SLR eligibility criteria. The SLR outcomes include several that do not appear within the NICE Final Scope or the DP (overall response rate [ORR], duration of response [DOR], time to response [TTR] and time to surgery [TTS]). These areas of mismatch may mean that the clinical effectiveness evidence generated by the SLR does not fully address either the NICE Final Scope[7] or the DP.[5]
- The EAG noted that the SLR excluded case series which is the design of the VHL Natural History study, used to provide comparator data for the matching-adjusted indirect comparison (MAIC) (further details in Section 3.4). The company was asked to clarify whether the VHL Natural History study was retrieved as part of the clinical effectiveness SLR (clarification question A.15a) and

whether the SLR was designed to identify all relevant natural history (or non-intervention studies (clarification question A.15b). The company responded that, "*The VHL Natural History Study was not retrieved as part of the SLR*"; and further that, "*…the VHL Natural History Study is a (currently) unpublished study specifically commissioned by MSD to address the lack of available relevant comparator data in the published literature. The VHL Natural History Study, not (yet) being a published study, was not (and could not be) identified via a systematic literature review.*"[4] The company declined to conduct another SLR designed to retrieve all relevant interventional, non-interventional or natural history studies (as per clarification question A15.c) stating that this was, "*Not applicable…..the SLR of clinical effectiveness evidence was designed to include such studies.*"[4] The EAG remains unclear as to whether relevant comparator studies could have been omitted from the submission. Details of the comparator studies are discussed further in Section 3.3.

- The SLR study selection criteria limited inclusion to English language publications. Methodological guidelines recommend the inclusion of non-English studies in systematic reviews in order to minimise the risk of language bias.[13-15]. The company was asked to provide the number of relevant studies omitted from the review because of being published in non-English languages and to consider the impact of the exclusion of such studies on clinical effectiveness estimates (clarification question A.19). In their response, the company stated that, "*A total of two non-English studies were identified during full text review and were excluded*" and further, "*Due to the very limited number of studies that were excluded for non-English language (i.e., two), we anticipate that the impact of this is minimal.*"[4] However, the company did not provide specific details of the excluded studies and therefore it was not possible for the EAG to determine the impact of these omissions on clinical effectiveness estimates. This means that the potential impact of language bias cannot be discounted.

3.1.3 Critique of study selection and data extraction

The company provided the following description of the study selection process in Section D1.1 of Appendix D of the CS: "*Two reviewers, working independently, reviewed all abstracts and proceedings identified by the searches according to the PICOTS criteria…with the exception of outcomes criteria, which was only applied to the full-text selection. Studies identified as eligible during abstract screening were then screened again by the same two reviewers by viewing the full-text versions of the study. Studies remaining eligible for inclusion after reviewing the full-text articles were then moved to data extraction. In each selection phase, the independent reviewers reconciled differences between them. A third reviewer was included to reach consensus on any discrepancies that were insolvable between the two reviewers. The process of study identification and selection was summarised with a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.*"[5] No details of the data extraction process or plan were provided in the CS.

EAG comment:

The screening process is satisfactory and has followed recommended good practice in systematic reviews.[16] The company was asked to provide details of the data extraction process, i.e., the number of reviewers involved and methods for resolving disagreements (clarification question A.20a). The company responded that, "*Two reviewers working independently reviewed eligible studies for the final list of selected eligible studies. Any discrepancies observed between the two reviewers were resolved by involving a third reviewer and coming to a consensus.*"[4] Unfortunately, this does not clearly answer the question posed by the EAG and therefore the approach used for data extraction remains unclear.

In clarification question A.20b, the EAG asked the company to provide details of the *a priori* plan for data extraction (i.e., details of the types of data to be extracted). The company stated the following by way of response:[4]

"*The following data extraction variables were captured a priori:*

The following information was extracted regarding study characteristics: study title, first author, study identifiers (e.g. cohort name, NCT number), study design, study duration (year of initiation/completion), phase, masking, number of patients/subjects enrolled and number completed, study duration, initiation, and completion dates, follow-up duration, inclusion/exclusion criteria, outcomes reported, study quality items, data source, eligibility period for observational data, and date of data cut-off.

The following information was extracted regarding interventions: treatment name, treatment dose, method of administration, frequency of administration, planned treatment duration, observed treatment duration, and concomitant/background therapies.

The following information was extracted regarding baseline/study inception patient/subject characteristics: age, age of VHL diagnosis, age of VHL associated RCC diagnosis, any other VHL associated tumours, age at start of treatment, gender, race and ethnicity, region/country, method of VHL status diagnosis, performance status (ECOG, KPS), and VHL type (1, 2A, 2B, 2C).

The following information was extracted regarding outcomes: tumour response proportion (objective, complete, partial, stable disease, and progressive disease); including response criteria used (e.g. RECIST 1.1, iRECIST or mRECIST), duration of response, time to response, time to surgery, progression-free survival, overall survival, drug-related adverse events, grade 3-5 AEs (all, drug-related), discontinuation due to AEs, serious AEs, and deaths."[4]

Whilst the above details represent a reasonable approach for data extraction, information about the format is not mentioned (e.g., Excel spreadsheet, Access database) and the actual data extracted are not provided. It is evident from the text and tables in Section B.2 of the CS[5] that some of the above types of data were extracted for the MK-6482-004 study. However, the CS (Appendix Q)[5] included less detail on the main source of comparator data (the VHL Natural History study), particularly in relation to outcomes. This is discussed further in Section 3.3.

3.1.4 Quality assessment

According to the CS (Appendix D),[5] quality assessment of the MK-6482-004 study was conducted using the Cochrane Risk of Bias 2 tool.[17] This tool assesses six risk of bias (RoB) domains: 1) sequence generation, 2) allocation concealment, 3) blinding of participants, personnel and outcomes assessors, 4) incomplete outcome data, 5) selective outcome reporting, and 6) other sources of bias. Appendix D included details for deriving the overall RoB for MK-6482-004 based on "*key domains*", however the latter were not defined in the context of the company's SLR:

- low RoB (low RoB for all key domains)
- unclear RoB (unclear RoB for one or more key domains)
- high-RoB (high-RoB for one or more key domains)

The CS did not include details of the quality assessment process (i.e., the number of reviewers involved and the approach for resolving disagreements).[5]

EAG comment:

- The Cochrane Risk of Bias 2 tool is designed to assess the methodological quality of randomised controlled trials (RCTs) and is not suitable for assessing non-comparative studies such as MK-

6482-004. The company were asked to provide a RoB assessment for all included intervention and comparator studies, using checklists suitable for the respective study designs (clarification question A.20c). In response, the company presented an assessment of MK-6482-004 using the Risk Of Bias In Non-randomised Studies (ROBINS-I) tool.[18] However, the company presented a brief tabulation of the assessment, referring only to the overall RoB rating for each top-level domain and did not include any information relating to signalling questions or rationale for the judgements made. The ROBINS-I tool is intended primarily for use with non-randomised, comparative studies. Given the study design of MK-6482-004 (a single-arm trial), the EAG suggest that an example of a more appropriate tool is the Joanna Briggs Institute (JBI) critical appraisal checklist for quasi-experimental studies.[19] The results of the company's and the EAG's RoB assessment are provided in Section 3.2.5 of this report.

- The company were asked to describe the process of RoB assessment (clarification question A.20a). The company clarified that: *"Two reviewers working independently reviewed eligible studies for the final list of selected eligible studies. Any discrepancies observed between the two reviewers were resolved by involving a third reviewer and coming to a consensus.*" The EAG remains unclear about the process used for of RoB assessment, therefore the potential for bias in the review process cannot be discounted.

3.1.5 Evidence synthesis

In Section B.2.8 of the CS, the company stated that, "*The MK-6482-004 study is the only study that reports clinical effectiveness data on the treatment effect of Belzutifan in the relevant indication, therefore no meta-analysis possible. Information on the effectiveness of the comparator (standard of care in UK clinical practice) was derived from data collected in a retrospective non-interventional study conducted in the United States (the VHL Natural History Study).*" Section B.2.9 of the CS reports the methods and results of the MAIC using data from the two aforementioned studies.[5] Further details of the comparator study and the MAIC are provided in Sections 3.3 and 3.4 respectively, of this report.

3.2 Critique of trials of the technology of interest, their analysis and interpretation

3.2.1 Study retrieval

In Appendix D of the CS,[5] the company reported that, "*A total of 824 citations were identified by searching the bibliographic databases and conferences. After excluding 160 duplicates, a total of 664 citations were screened. This resulted in identification of 245 citations eligible for full-text review. Of the 245 full-text articles screened, 219 were excluded: 52 due to a population that was not of interest; five due to study design that was not of interest; 147 for no outcomes of interest and 15 for other reasons. Thus, a total of 26 citations representing 26 unique studies were initially included in this review.*

However, only one of the 26 studies identified investigated the efficacy of Belzutifan, specifically the MK-6482-004 study as reported in the Jonasch et al. 2021 publication,[20] and so only that one study has been included for the purposes of this submission."

EAG comment:

The company were asked to explain why the other 25 citations met the inclusion criteria for the SLR but were then excluded from the results of the review (clarification question A.17a). The company responded by stating that, *"The SLR for clinical effectiveness evidence was performed primarily in order to identify all relevant published and unpublished randomised controlled trials (RCTs) and non-randomised clinical trials (non-RCTs) relating to Belzutifan as per the final scope. Consequently, for this purpose (as stated in section B.2.2 of the company submission), only studies that reported relevant*

data on Belzutifan are of interest and any studies that did not do so were excluded at the true final stage."[4]

The EAG noted that the purpose of the SLR was explained differently between Section B.2.2 and Appendix D of the CS.[5] The respective explanations are: "*A SLR was performed to identify all relevant published and unpublished randomised controlled trials (RCTs) and non-randomised clinical trials (non-RCTs) relating to Belzutifan as per the final scope*" and "*To identify and select relevant studies, a systematic literature review (SLR) was carried out in accordance with NICE guidance, according to a protocol developed a priori, to identify relevant studies that investigated Belzutifan and any relevant comparator treatments for the indication of interest for this appraisal*". However, this does not justify the exclusion of studies that could have provided relevant comparator data and the EAG remains unclear as to whether such studies were omitted from the SLR. Therefore, the risk of study selection bias cannot be discounted and this (along with other problems with the review process) represents a key issue.

3.2.2 Details of the included trial

Details outlined in Section B.2.3 of the CS[5] indicated that the study labelled MK-6482-004 was a single arm trial, evaluating the efficacy and safety of 120 mg (three 40 mg tablets) of Belzutifan administered orally once daily in adults with VHL disease, who have at least one measurable RCC tumour. Outcomes of the study included: ORR, TTR, PFS, TTS and AEs. The study was conducted at 11 centres in the United States (US), Denmark, France, and the UK. One patient received treatment in the UK.[5] Table summarises details of the MK-6482-004 objectives, study design and participant eligibility criteria.

Table 3.3: Details of MK-6482-004

Study	**MK-6482-004**
Study design	Open-label, multicentre, single-arm, non-randomised, interventional, Phase 2 study. Participants were evaluated with imaging every 12 weeks.
Population	Patients with Von Hippel-Lindau (VHL) disease who have at least one measurable renal cell carcinoma (RCC) tumour
Intervention(s)	Belzutifan 120 mg once daily
Comparator(s)	None
Reported outcomes specified in the decision problem	Overall response rate (ORR) Duration of response (DOR) Time to response (TTR) Progression-free survival (PFS) Time to surgery (TTS) Adverse events (AE)
Eligibility criteria	
Key inclusion criteria	Male and female participants of at least 18 years of age were eligible for enrolment in this study. Key inclusion criteria were as follows: • Had a diagnosis of VHL disease based on a germline VHL alteration. • Had at least 1 measurable solid RCC tumour and no RCC tumour greater than 3.0 cm that requires immediate surgical intervention. The diagnosis of RCC could be radiologic (histologic diagnosis not required). Participants could have VHL disease-associated tumours in other organ systems. • Had an Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 1.

Study	MK-6482-004
	Had adequate organ function as defined below: • Absolute neutrophil count ≥1,000/μL, haemoglobin level ≥10 g/dL and platelet count ≥100,000/μL without transfusion or growth factor support within 2 weeks prior to obtaining the haematology values at screening. • Serum creatinine level ≤2.0 x upper limit of normal (ULN). • Aspartate amino transferase (AST) and alanine transaminase (ALT) <2.5 x ULN, total bilirubin <1.5 x ULN (<3 x ULN in patients with Gilbert's disease), and alkaline phosphatase ≤2.5 x ULN.
Key exclusion criteria	Participants were excluded from the study if they met any of the following criteria: • Had any systemic anticancer therapy included anti-VEGF (vascular endothelial growth factor) therapy or a systemic investigational anticancer agent). • Had a surgical procedure for VHL disease or any major surgical procedure completed within 4 weeks prior to study enrolment. • Had received prior treatment with PT2977 or another hypoxia inducible factor (HIF)-2α inhibitor. • Had radiotherapy within 4 weeks prior to study enrolment. • Had an immediate need for surgical intervention for tumour treatment. • Had evidence of metastatic disease on screening imaging. • Had malabsorption due to prior gastrointestinal (GI) surgery or GI disease. • Had any major cardiovascular event within 6 months prior to study drug administration including but not limited to myocardial infarction, unstable angina, cerebrovascular accident, transient ischemic event, pulmonary embolism, clinically significant ventricular arrhythmias (e.g., sustained ventricular tachycardia, ventricular fibrillation, torsades de pointes) or New York Heart Association Class III or IV heart failure.
Settings and locations	Patients were enrolled at 11 centres in the United States, Denmark, France, and the United Kingdom (UK). One patient received treatment in the UK.
Trial drugs	Trial drug: Belzutifan Dose strength: Not mentioned Dose and regimen: 120 mg administered once a day, unless unacceptable adverse events or disease progression occurred. Route of administration: Orally
Study Objectives	
Primary Objectives	To evaluate the efficacy of MK-6482 (Belzutifan) for the treatment of VHL disease-associated RCC as measured by ORR per response evaluation criteria in solid tumours (RECIST) 1.1 (described in Appendix N). Primary endpoint: ORR: the proportion of participants who have achieved a complete response (CR) or partial response (PR). Relevant VHL disease-associated tumour(s) for the objective and endpoint: RCC.
Secondary Objectives	To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by DOR per RECIST 1.1.

Study	MK-6482-004
	Secondary endpoint: DOR: the time from first documented evidence of CR or PR until either disease progression or death due to any cause, whichever occurs first. To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by time to response rate TTR per RECIST 1.1. Secondary endpoint: TTR: the time from the start of study intervention to the first documentation of a response, calculated for participants with a best confirmed response of CR or PR. To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by PFS per RECIST 1.1. Secondary endpoint: PFS: the time from the start of study intervention to the first documented disease progression or death due to any cause, whichever occurs first. To evaluate the efficacy of MK-6482 for the treatment of VHL disease-associated RCC as measured by TTS. Secondary endpoint: TTS: the time from the start of study intervention to the date of surgery. Relevant VHL disease-associated tumour(s) for the above objective and endpoint: RCC for all the above secondary objectives. To evaluate efficacy of MK-6482 for the treatment of VHL disease-associated non-RCC tumours (retinal and central nervous system [CNS] hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)*. Secondary endpoint: ORR, DOR, TTR, PFS, and TTS Relevant VHL disease-associated tumour(s) for the objective and endpoint: Non-RCC tumours (retinal and CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)*.
Exploratory Objectives	Listed as secondary outcomes: To evaluate the safety and tolerability of MK- 6482 for the treatment of VHL disease- associated RCC. Endpoint: Adverse events (AEs) and study intervention discontinuation due to AEs. To assess the pharmacokinetics of MK- 6482.

Study	MK-6482-004
	Endpoint: Plasma concentrations of MK-6482 and its metabolite(s). Relevant VHL disease-associated tumour(s) for these two objectives and endpoints: All patients.

Based on Tables 4, 5, 6 and 7 of the CS.[5]
AE = adverse event; ALT = alanine transaminase; AST = Aspartate amino transferase; CNS = central nervous system; CR = complete response; CS = company submission; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; GB = Great Britain; GI = gastrointestinal; HIF = hypoxia inducible factor; ORR = overall response rate; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumours; PR = partial response; RCC = renal cell carcinoma; RECIST = Response Evaluation Criteria in Solid Tumours; TTR = time to response; TTS = time to surgery; UK = United Kingdom; ULN = upper limit of normal; VEGF = vascular endothelial growth factor; VHL = Von Hippel-Lindau
*As stated in the footnote of Table 7 of the CS:[5] "*Subgroup data on tumours other than VHL-associated RCC, CNS hemangioblastomas, or pNETs are presented in detail in this document as they are not included in the GB marketing authorisation for Belzutifan in this indication or the scope of this appraisal.*"

EAG comment:

- The company were asked to clarify whether the MK-6482-004 study population was representative of the UK population, and to compare study and UK patient characteristics for all three tumour subgroups (clarification question A.21). The company responded that, "*the baseline characteristics of this study have been presented to UK clinical experts with experience treating patients with VHL disease. Experts broadly agreed that these were representative of/applicable to the patients in the UK who would be treated with Belzutifan in accordance with the marketing authorisation.*"[4] Details relating to the number of clinical experts consulted and characteristics such as their role, expertise and conflicts of interests were not provided with the CS[5] and were not available from documentation provided later on.[21] Later in their response to the same clarification question (A.21) the company stated that, "*It should also be noted that the MK-6482-004 study focused on/included patients who had to have VHL disease-associated RCC, and so the patient population included in the study may not be fully representative of the general VHL disease population in the UK*".[4] Given the equivocal nature of statements provided by the company within their response, the EAG remains uncertain about the comparability between the UK and study populations. In responding to the EAG's request to compare study and UK patient characteristics for all tumour subgroups, the company failed to provide any information, suggesting that this was due to the rarity of VHL disease in the UK resulting in a lack of available published data. This is therefore a key issue.
- The population in the DP is defined as, "*Adult patients with VHL disease who require therapy for VHL associated RCC, CNS hemangioblastomas or pNETs, and for whom localised procedures are unsuitable or undesirable.*"[5] However, the MK-6482-004 trial did not specifically require participants to be considered unsuitable or undesirable for localised procedures. The company was asked to discuss the implications of this discrepancy (clarification question A.23) and responded as follows:

  "*As described previously…we consider the MHRA indication wording to be both broader and narrower than the trial population. We consider it is broader as patients with a e.g. CNS hemangioblastoma (with or without other tumours) but no RCC could be treated with Belzutifan. Similarly, patients with pNETs (with or without other tumours) but no RCC tumour could be treatment with Belzutifan.*

We consider it narrower as it is only for patients who "require therapy" but also for whom localised therapies are "unsuitable or undesirable". As such, these patients are more unwell, have a more severe presentation that requires an intervention than some in the study."[4]

This indicates that the study population is not entirely representative of the DP population.

3.2.3 Statistical analysis of the included studies

Details of the statistical analysis and definitions of study groups in the MK-6482-004 study as presented in the CS,[5] are summarised in Table .

Table 3.4: Summary of statistical analysis and study groups for MK-6482-004

	MK-6482-004
Study design and overview	The MK-6482-004 study was a single-arm open-label Phase 2 study that evaluated the efficacy and safety of Belzutifan in patients with von Hippel-Lindau (VHL) disease who have at least 1 measurable renal cell carcinoma (RCC) tumour (as defined by RECIST 1.1).
Treatment assignment and stratification	This was an open-label single-group trial and so had no assignment, randomisation, or stratification.
Study hypotheses	No formal hypothesis testing. For the purposes of sample size determination only, null hypotheses and alternative hypotheses were formulated (described later in the "Sample size and power" section of this table)

Study objectives	**Specific to VHL RCC tumours:**
	• Primary objective:
	  o To evaluate the efficacy of Belzutifan for the treatment of VHL disease associated RCC as measured by overall response rate (ORR) per Response Evaluation Criteria in Solid Tumours Version 1.1 (RECIST 1.1)
	• Secondary objectives:
	  o To evaluate efficacy of Belzutifan for the treatment of VHL disease associated-RCC measured as follows:
	  ▪ Duration of response (DOR)
	  ▪ Time to response (TTR)
	  ▪ Progression-free survival (PFS)
	  ▪ Time to Surgery (TTS)
	Specific to VHL non-RCC tumours:
	• Secondary objectives:
	  o To evaluate efficacy of Belzutifan for the treatment of VHL disease associated non-RCC tumours (retinal and CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas)
	Applies to all patients in the study:
	• Secondary objectives:
	  o To evaluate safety and tolerability of Belzutifan
	  o To assess the pharmacokinetics (PK) of Belzutifan
	• Exploratory objective:
	  o To evaluate changes in pharmacodynamic markers (e.g., serum erythropoietin)

	MK-6482-004
Analysis populations	• All Patients: All patients who have signed the written informed consent form. This population will be used for the summary of patient disposition and data listings. • Efficacy Analysis Set: The All Participants as Treated (APaT) population will be used for the analyses of efficacy. The APaT population consists of all allocated patients who received at least one dose of Belzutifan. • Safety Analysis Set: The APaT population will be used for the analysis of safety data in this study. The APaT population consists of all allocated patients who received at least one dose of Belzutifan. • Pharmacokinetic Analysis Set: The Pharmacokinetic Analysis Set will include all patients who received at least 1 dose of Belzutifan and have at least one post-dose pharmacokinetic sample collection. • Pharmacodynamic Analysis Set: The Pharmacodynamic Analysis Set will include all patients who received at least one dose of the study drug and have evaluable pharmacodynamics data above the limit of quantification.
Primary endpoint	**Specific to VHL RCC tumours:** • Overall response rate (ORR) in VHL disease associated RCC tumours, defined as proportion of patients with a best confirmed response of Complete Response (CR) or Partial Response (PR) as determined by RECIST 1.1
Key secondary endpoint	**Specific to VHL RCC tumours:** • Secondary endpoints: o Duration of response (DOR) in VHL disease associated RCC tumours, defined as the interval from the first documentation of response, as determined by RECIST 1.1, to the earlier of the first documentation of disease progression or death from any cause, and calculated for patients with a best confirmed response of CR or PR. o Time to response (TTR) in VHL disease associated RCC tumours, defined as the interval from the start of study treatment to the first documentation of a response, as determined by RECIST 1.1, and calculated for patients with a best confirmed response of CR or PR. o Progression-free survival (PFS) in VHL disease-associated RCC tumours, defined as the interval from the start of study treatment until the earlier of the first documentation of disease progression determined by RECIST 1.1 or death from any cause. o Time to surgery (TTS) for VHL disease associated RCC tumours, defined as the interval from the start of study treatment to the date of surgery. **Specific to VHL non-RCC tumours:** • Secondary endpoints: o ORR, DOR, TTR, PFS, and TTS for non-RCC tumours associated with VHL disease in individual organ systems (retinal lesions, CNS hemangioblastomas, pancreatic, adrenal, endolymphatic sac tumour and epididymal cystadenomas). **Applies to all patients in the study:** • Exploratory endpoints: o Changes in pharmacodynamic markers

	MK-6482-004
	• Safety endpoints: o Physical examinations o Vital sign measurements (including pulse oximetry) o 12-lead electrocardiograms (ECG) with QTc interval determination o Clinical laboratory measurements o Concomitant medications o Incidence, intensity, and relationship of AEs and serious adverse events (SAEs) o Effects on fertility in males (semen analysis, and measurement of testosterone, follicle-stimulating hormone, luteinizing hormone, and inhibin B levels)
Sample size and power	This study will enrol approximately 50 patients. Even though no formal hypothesis testing will be performed for this study, the required sample size for this study is based on the following assumptions. The null hypothesis is that the ORR is 15% (P0 = 0.15). The alternative hypothesis is that the ORR is 30% (P1 = 0.3). A sample size of 50 patients will provide greater than 80% power to reject the null under the alternative hypothesis using a one-sided test at a 0.05 level of significance.
Interim and final analyses	Periodic review of the trial data will be performed. Any analysis for the study will only take place after all patients have had the opportunity to complete at least two imaging assessments on study or have discontinued study therapy by the time of analysis data cut-off. The final analyses for the study will utilize a data cut-off date which will be at least 36 weeks after enrolment of the last patient.
Data management, patient withdrawals	Patients who discontinue from study treatment would complete the safety follow-up and long-term follow-up assessments according to the Schedule of Events. During the safety follow-up visit the patient would be evaluated for continuation or resolution of any AEs/SAEs. Patients who discontinue study treatment for any reason would undergo long term follow-up every 6 months for up to 3 years following enrolment of last patient into the study.

Based on Table 11 and Appendix M of the CS.[5]
AE = adverse events; APaT = all participants as treated; CNS = central nervous system; CR = complete response; CS = company submission; DOR = duration of response; ECG = electrocardiogram; ORR = overall response rate; PFS = progression-free survival; PK = pharmacokinetics; PR = partial response; QTc not defined in the CS but this usually refers to the corrected QT interval on an ECG reading;[22] RCC = renal cell carcinoma; RECIST = Response Evaluation Criteria in Solid Tumours; SAE = serious adverse events; TTR = time to response; TTS = time to surgery; VHL = Von Hippel-Lindau

3.2.4 Baseline characteristics

Details in the CS indicated that a total of 61 participants (32 males and 29 females) were included in the MK-6482-004 study.[5] A summary of participants' baseline characteristics is presented in Table .

Table 3.5: Study demographic and baseline characteristics (safety analysis set) – all patients

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
Participants in population	**32**		**29**		**61**	
Age (Years)						
Participants with data	32		29		61	
Mean	38.8		43.3		41.0	

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
SD	12.7		14.1		13.5	
Median	36.0		44.0		41.0	
Range	22.0 to 65.0		19.0 to 66.0		19.0 to 66.0	
Ethnicity						
Hispanic or Latino	3	(9.4)	3	(10.3)	6	(9.8)
Not Hispanic or Latino	28	(87.5)	26	(89.7)	54	(88.5)
Unknown	1	(3.1)	0	(0.0)	1	(1.6)
Race						
Asian	1	(3.1)	0	(0.0)	1	(1.6)
Black or African American	1	(3.1)	1	(3.4)	2	(3.3)
Native Hawaiian or Other Pacific Islander	1	(3.1)	0	(0.0)	1	(1.6)
White	28	(87.5)	27	(93.1)	55	(90.2)
Unknown	1	(3.1)	1	(3.4)	2	(3.3)
Weight (kg)						
Participants with data	32		29		61	
Mean	86.7		72.1		79.7	
SD	21.4		23.4		23.4	
Median	81.5		65.0		74.4	
Range	63.0 to 147.6		47.7 to 147.0		47.7 to 147.6	
Height (cm)						
Participants with data	32		27		59	
Mean	176.6		161.1		169.5	
SD	8.7		6.7		11.0	
Median	175.5		160.1		169.0	
Range	159.5 to 195.0		148.0 to 174.0		148.0 to 195.0	
BMI (kg/m^2)						
Participants with data	32		27		59	
Mean	27.7		27.8		27.8	
SD	6.0		8.8		7.4	
Median	27.0		24.5		26.3	
Range	18.4 to 42.7		17.2 to 52.0		17.2 to 52.0	
ECOG Performance Status						
0	24	(75.0)	26	(89.7)	50	(82.0)
1	8	(25.0)	2	(6.9)	10	(16.4)
2	0	(0.0)	1	(3.4)	1	(1.6)

Based on Table 9 of the CS.[5]

	Male		Female		Total	
	n	(%)	n	(%)	n	(%)
The following table footnotes were included in the CS:[5] • Database Cut-off Date: 01APR2022 • Number of participants: Safety Population • Note: Baseline is defined as the last available measurement prior to the first dose administered. BMI = body mass index; CS = company submission; ECOG = Eastern Cooperative Oncology Group; SD = standard deviation						

The CS stated that all patients included in the MK-6482-004 study had at least one concurrent non-RCC tumour at baseline.[5] Baseline variables relating to VHL disease, tumour characteristics and prior surgery are shown in Table .

Table 3.6: Study demographic and baseline characteristics (safety analysis set) – all patients – additional data

	Belzutifan (N=61)
Age at time of VHL diagnosis (years)	
N	61
Mean	31.3 (14.29)
Median	32.0
Min, Max	4, 66
VHL Subtype, n (%)	
Type 1	51 (83.6)
Type 2A	2 (3.3)
Type 2B	6 (9.8)
Type 2C	0
Missing	2 (3.3)
VHL-associated Non-RCC tumours, n (%)	
Pancreatic lesions	32 (52.5)
- Pancreatic lesions of which were pNETs	22 (36.1)
Adrenal lesions (Pheochromocytomas)	3 (4.9)
CNS hemangioblastoma [3]	51 (83.6)
Endolymphatic sac tumours	1 (1.6)
Epididymal cystadenomas	10 (16.4)
Retinal lesions	17 (27.9)
Other	2 (3.3)
Time from Original Diagnosis of VHL associated RCC to First Dose (months) [1]	
n	45
Mean (SD)	103.43 (96.231)
Median	77.60
Q1, Q3	24.54, 136.97

	Belzutifan (N=61)
Min, Max	0.5, 389.4
Time from Last Surgery to First Dose (months)	
n	59
Mean (SD)	37.01 (38.493)
Median	23.49
Q1, Q3	9.66, 41.13
Min, Max	0.6, 137.6
Number of Prior Surgeries per Subject	
n	59
Mean (SD)	5.5 (3.34)
Median	5.0
Min, Max	1, 15
Age at time of VHL associated RCC diagnosis (years) [2]	
n	45
Mean (SD)	33.8 (13.06)
Median	32.0
Min, Max	15, 62
Histology, n (%)	
Renal Cell Carcinoma of Clear Cell Subtype	43 (70.5)
Other	2 (3.3)
Not Done	16 (26.2)
Histological Grade	
GX - Grade cannot be assessed	2 (3.3)
G1 - Nucleoli absent or inconspicuous and basophilic at 400x magnification	10 (16.4)
G2 - Nucleoli conspicuous and eosinophilic at 400x magnification, visible but not prominent at 100x magnification	23 (37.7)
G3 - Nucleoli conspicuous and eosinophilic at 100x magnification	8 (13.1)
G4 - Marked nuclear pleomorphism and/or multinucleate giant cells and/or rhabdoid and/or sarcomatoid differentiation	0
Missing	2 (3.3)
TNM Stage T	
TX	1 (1.6)
T0	0
T1	5 (8.2)
T1a*	48 (78.7)
T1b	2 (3.3)
T2	0
T2a	0

	Belzutifan (N=61)
T2b	1 (1.6)
T3	0
T3a	0
T3b	0
T3c	0
T4	0
TNM Stage N	
NX	NX 13 (21.3)
N0	N0 46 (75.4)
N1	N1 0
TNM Stage M	
cM0	59 (96.7)
cM1	0
pM1	0

Based on Table 10 of the CS.[5]
The following table footnotes were included in the CS:[5]

- [1] (First Dose Date-Date of first positive biopsy+1)/30.4375
- [2] (Date of VHL associated RCC diagnosis-Birthdate+1)/365.25
- [3] The number patients with CNS hemangioblastomas shown in this table is according to investigator assessment, study results are reported later on in this document in terms of the number of patients with CNS hemangioblastoma according to independent review committee determination where this was found to be n=50.
- *T1a means that the tumour is less than 4cm across and is completely inside the kidney.
- Date of Data Cut-off: 01APR 2022

cM0 not defined in the CS but usually means that there is no sign of cancer spread to a different part of the body on physical examination or scans but cancer cells (detected by laboratory tests) are present in the blood, bone marrow or lymph nodes distal to the main tumour;[23] cM1 not defined in the CS but usually means that the cancer has spread to another part of the body;[23] CNS = central nervous system; CS = company submission; max = maximum; min = minimum; G = histological cancer grade; N = number of participants; pM1 not defined in the CS but usually means that cancer measuring >0.2mm has spread to another part of the body;[23] pNET = pancreatic neuroendocrine tumour; Q = quartile; RCC = renal cell carcinoma; SD = standard deviation; TNM not defined in the CS but usually refers to a cancer staging system whereby "T" describes the size of the tumour and any spread of cancer into nearby tissue, "N" describes spread of cancer to nearby lymph nodes and "M" describes metastasis (spread of cancer to other parts of the body);[12] VHL = Von Hippel-Lindau; X (in the context of TNM staging) not defined in the CS but usually means that cancer in the main tumour or nearby lymph nodes/ or metastasis cannot be measured;[12]

EAG comment:

- Baseline characteristics relating to demographics, anthropometry, performance status, disease/tumour characteristics and prior surgical treatment are shown in Table and 3.6. However, the presentation in the CS did not include the number of each type of VHL-associated tumour per patient; therefore, the EAG requested these details (clarification questions A.27a and b). The company replied that they were, "*not in a position to provide detailed individual patient data (i.e. the "per patient" data) from the MK-6482-004 study. Data relevant to this appraisal and the decision problem with regard to number and types of tumours patients in the MK-6482-004 study had at baseline are already provided in the company submission in Appendix P…*".[4] However,

although tables in Appendix P of the CS listed the numbers of RCC tumour, pNETs and CNS Hb for the overall study sample, the individual participant data were not provided.[5]

- The company were also asked to a clarify a discrepancy in the number of pancreatic lesions as shown in Table 10 of the CS (n=32)[5] and the published paper of MK-6482-004 (Jonasch *et al*., 2021) (n=61)[20] (clarification question A.27c). The company stated that, "*The number of patients with pancreatic lesions reported in the Jonasch et al. NEJM (2021) study of N=61 refers to the patients with pancreatic lesions as determined via investigator assessment, whereas the n=32 reported in Table 11 of Document B is the number of patients with pancreatic lesions according to independent review committee determination which is a smaller number*."[4] The EAG acknowledge that the difference may have arisen as a result of different assessment methods used to determine the presence of pancreatic lesions however, remain concerned at the size of the discrepancy. The EAG did also note that the number of pNETs (n=22) was consistent across both documents.[5, 20]

3.2.5 Risk of bias assessment

The company provided two inappropriate risk of bias assessments, one in the CS[5] (using the Cochrane RoB 2 tool[17]) and the other as part of their response to clarification question A.20c[4] (the ROBINS-I checklist[18]). These results of these assessments are shown in Table and 3.8 below.

Table 3.7: Cochrane RoB 2 assessment of the MK-6482-004 study

Study	MK-6482-004
NCT number	NCT03401788
Sequence generation	High risk
Allocation concealment	High risk
Blinding of participants, personnel and outcome assessors	High risk
Incomplete outcome reporting	Low risk
Selective outcome reporting	Low risk
Other sources of bias	High risk
Based on Table 113 of the CS.[5] CS = company submission; NCT = National Clinical Trial; RoB = risk of bias	

Table 3.8: ROBINS-I assessment of the MK-6482-004 study

Study	MK-6482-004
Bias due to confounding	Low
Bias in selection of participants into the study	Low
Bias in classification of interventions	Low
Bias due to deviations from intended interventions	Low
Bias due to missing data	Low
Bias in measurement of outcomes	Low
Bias in selection of the reported result	Low
Overall risk of bias	Low
Based on Table 3 of the company's response to clarification question A.20c.[4] ROBINS-I = Risk Of Bias In Non-randomised Studies	

EAG comment:

Since both of the above tools are designed to assess the RoB in controlled studies,[17, 18] they are not appropriate for use with a single-arm study such as MK-6482-004. The EAG also noted the minimalist

presentation of details for each assessment and expected to see more information (e.g., responses to signalling questions and descriptions of the rationale for the judgement made per question) given that both tools are domain-based. Given the single-arm design of MK-6482-004, the EAG suggest that an example of a more appropriate tool is the JBI critical appraisal checklist for quasi-experimental studies where the possible responses for each checklist item include "Yes", "No", "Unclear" and "Not applicable".[19] The result of the EAG assessment, based on the Jonasch *et al*., (2021) paper[20] is presented in Table 3.9.

Table 3.9: EAG assessment of MK-6482-004 using the JBI quasi-experimental studies checklist

Critical appraisal item	Judgement (rationale)
1. Is it clear in the study what is the 'cause' and what is the 'effect' (i.e., there is no confusion about which variable comes first)?	Yes (it is clear that the outcomes were assessed after the start of the treatment period)
2. Were the participants included in any comparisons similar?	NA (non-comparative study design)
3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	NA (non-comparative study design)
4. Was there a control group?	No (non-comparative study design)
5. Were there multiple measurements of the outcome both pre and post the intervention/exposure?	Yes (pre- treatment, RCC tumours were evaluated by independent central radiology reviewers ≥2 times before screening imaging was performed, to estimate growth kinetics before treatment. Linear growth of lesions during treatment was calculated in patients who had a screening and ≥2 imaging assessments whilst receiving treatment. This is an appropriate approach given the population and indication).
6. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?	Yes (the efficacy population consisted of all included participants)
7. Were the outcomes of participants included in any comparisons measured in the same way?	NA (non-comparative study design)
8. Were outcomes measured in a reliable way?	Yes (tumour assessment of solid lesions were performed by IRC using RECIST version 1.1, for each organ system affected by VHL disease)
9. Was appropriate statistical analysis used?	Yes (the presentation of outcome data was mainly descriptive which is appropriate for the non-comparative study design)

The critical appraisal is based on the Jonasch *et al*., 2021 paper.[20]
IRC = independent review committee; JBI = Joanna Briggs Institute; NA = not applicable; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours

The EAG's conclusion is that in light of the single-arm study design, the MK-6482-004 study is at risk of bias because of the potential for confounding of the treatment effect.

3.2.6 Efficacy results of the included studies

3.2.6.1 Patient disposition, follow-up duration, study drug exposure and summary of outcomes

The study comprised of 61 participants who received one dose of Belzutifan. As of the 1 April 2022 database cut-off date, 38 participants (62.3%) were receiving Belzutifan, 23 participants (37.7%) had discontinued Belzutifan and six participants (9.8%) had discontinued from the study entirely.[5] Table summaries the patient disposition for the MK-6482-004 study.

Table 3.10: MK-6482-004 summary of patient disposition (safety analysis set)

	Belzutifan (N=61) n (%)
Treatment Ongoing at Data Cut-off Date	38 (62.3)
Discontinued Treatment	23 (37.7)
Reason for Treatment Discontinuation	
Disease progression per RECIST 1.1 for VHL disease-associated RCC tumours	6 (9.8)
Disease progression due to symptomatic deterioration of the patient's health status	0
Adverse event that in the opinion of the investigator or medical monitor would lead to undue risk if study treatment were continued	2 (3.3)
Study drug interruption for more than three consecutive weeks due to a Grade 3-4 or intolerable toxicity that is attributed to study drug	0
Gross noncompliance with protocol	0
Pregnancy in a female patient during the study	1 (1.6)
Death*	2 (3.3)
Lost to follow-up	0
Patient decision to discontinue study drug	11 (18.0)
Sponsor discontinuation of study	0
Other	1 (1.6)
On Study at Data Cut-off Date [1]	55 (90.2)
Off Study	6 (9.8)
Reason for Study Discontinuation	
Death	2 (3.3)
Informed Consent Withdrawn	2 (3.3)
Lost To Follow-up	0
Sponsor discontinuation of study	0
Other	2 (3.3)
Completed Safety Follow-up Visit	13 (21.3)
On Long Term Follow-up Period at Data Cut-off Date	10 (16.4)

Based on Table 12 of the CS.[5]
The following table footnotes were included in the CS:[5]
- [1] Patients are still on study treatment or in long term follow-up as of the cut-off date.

	Belzutifan (N=61) n (%)
• Date of Data Cut-off: 01APR2022 • *The two deaths (suicide attempt and toxicity to various agents) were assessed a not drug-related by the investigator. CS = company submission; RCC = renal cell carcinoma; RECIST = Response Evaluation Criteria in Solid Tumours; VHL = Von Hippel-Lindau	

The median duration of follow-up among the 61 participants with RCC in the safety analysis set at the 1 April 2022 data cut-off date was 37.7 months (range: 4.2 to 46.1 months).[5] Table summarises the follow-up duration data in the MK-6482-004 study.

Table 3.11: MK-6482-004 summary of follow-up duration (safety analysis set)

Follow-up duration (months)	Belzutifan (N=61)			
Date of data cut-off	**01-Jun-2020**	**01-Dec-2020**	**15-Jul-2021**	**01-Apr-2022**
Median (Range)	[redacted]	21.8 (4.2-30.1)	[redacted]	37.7 (4.2-46.1)
Mean (SD)	[redacted]	22.4 (3.35)	[redacted]	38.1 (5.01)

Based on Table 13 of the CS.[5]
Footnote from the CS: Follow-up duration is defined as the time from first dose to the date of death or the database cut-off date if the subject is still alive.[5]
CS = company submission; SD = standard deviation

The median duration of exposure to Belzutifan was [redacted] at the 1 April 2022 database cut-off date.[5] The data on the duration of exposure to Belzutifan during the MK-6482-004 study is summarised in Table .

Table 3.12: MK-6482-004 study drug exposure (safety analysis set)

	Belzutifan (N=61)			
Date of data cut-off*	**1-Jun-20**	**1-Dec-20**	**15-Jul-21**	**1-Apr-22**
Number of patients exposed	[redacted]	61	[redacted]	[redacted]
Duration of exposure (weeks)				
N	[redacted]	61	[redacted]	[redacted]
Mean (SD)	[redacted]	92.77 (23.561)	[redacted]	[redacted]
Median	[redacted]	94.14	[redacted]	[redacted]
Min, Max	[redacted]	8.4, 130.9	[redacted]	[redacted]
Cumulative dose received (mg/subject)				
N	[redacted]	61	[redacted]	[redacted]
Mean (SD)	[redacted]	72937.7 (21453.74)	[redacted]	[redacted]
Median	[redacted]	77760.0	[redacted]	[redacted]

	Belzutifan (N=61)			
Date of data cut-off*	**1-Jun-20**	**1-Dec-20**	**15-Jul-21**	**1-Apr-22**
Min, Max	[redacted]	4680, 106680	[redacted]	[redacted]
Based on Table 14 of the CS.[5] CS = company submission; max = maximum; min – minimum; SD = standard deviation				

At the time of data cut-off (1 April 2022), there was a 63.9% (95% confidence interval (CI): 50.6%, 75.8%) ORR for all RCC patients. Subgroup analysis yielded an ORR of 44% (95% CI: 30.0%, 58.7%) for patients with CNS Hb and 90.9% (95% CI: 70.8%, 98.9%) for patients with pNETs.[5] However, it should be noted that the number of participants in each tumour subgroup were not provided. A summary of the efficacy results from the MK-6482-004 study at the 1 April 2022 data cut-off date is presented in Table . Further detail on each outcome will be provided in the following subsections.

Table 3.13: Summary of MK-6482-004 study efficacy results (1 April 2022 data cut-off)

Outcome	Summary of results
RCC (all patients, N=61)	
Overall response rate (ORR)	63.9% (95% CI: 50.6%, 75.8%)
Disease control rate (DCR)	98.4% (95% CI: 91.2%, 100.0%)
Duration of response (DOR)	Median DOR not reached (range: 5.4+ to 35.8+ months)
Time to response (TTR)	Median TTR was 11.1 months (range: 2.7 to 30.5 months) among 39 participants with response
Progression-free survival (PFS)	[redacted]
Time to surgery (TTS)	NE
Subgroup of patients with CNS hemangioblastoma (n=50)	
ORR	44.0% (95% CI: 30.0%, 58.7%)
DCR	90.0% (95% CI: 78.2%, 96.7%)
DOR	Median DOR not reached (range: 3.7+ to 38.7+ months)
TTR	[redacted]

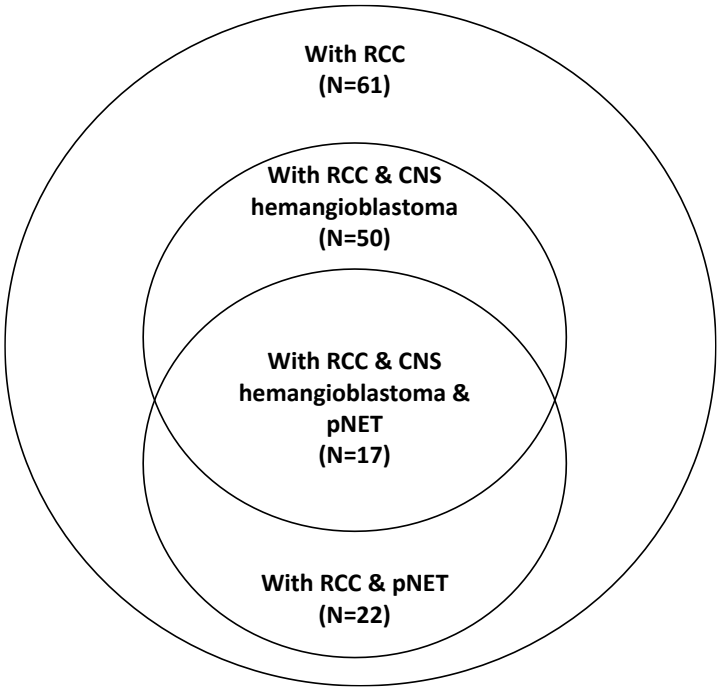
Outcome	Summary of results
PFS	[redacted]
TTS	NE
Subgroup of patients with pNET (n=22)	
ORR	90.9% (95% CI: 70.8%, 98.9%)
DCR	100% (95% CI: 84.6%, 100.0%)
DOR	Median DOR not reached (range: 11.0+ to 37.3+ months)
TTR	[redacted]
PFS	[redacted]
TTS	NE

Based on Table 15 and Section B.2.7 of the CS[5] and the company's response to clarification question A.22.[4]
95% CI = 95% confidence interval; CNS = central nervous system; CS = company submission; DCR = disease control rate; DOR = duration of response; NE = not evaluable; ORR = overall response rate; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; TTR = time to response; TTS = time to surgery

EAG comment:

In clarification question A.22, the EAG asked the company to provide the number of patients in each subgroup used to estimate efficacy results in Table above. The company did not provide the number of patients explicitly but referred the EAG back to parts of the CS (namely, Sections B.2.6 and B.2.7)[5] which did provide the number of patients in each subgroup. The EAG has added this information to Table 3.13 and noted a Venn diagram presented in Section B.2.7 of the CS which summarises the same details (see Figure 3.1).[5]

Figure 3.1: Summary of subgroups in the MK-6482-004 study



Based on Figure 9 of the CS.[5]
CNS = central nervous system; CS = company submission; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma.

As part of the same clarification question, the EAG asked the company to clarify whether the efficacy measures (such as ORR) are based on subgroup-specific tumours (e.g., in the subgroup of patients with CNS Hb does the ORR = 44% refer to CNS Hb tumours only or any tumour that a patient in the CNS subgroup might have (thus CNS Hb, RCC or pNET-associated tumours). The company replied that, "*The data on ORR, DCR, DOR, TTR, PFS, and TTS… are specific to the tumours for which they are reported i.e. the results reported under the "RCC (all patients)" subheading are specific to RCC tumours, the results reported under the "Subgroup of patients with CNS hemangioblastoma" subheading are specific to CNS hemangioblastomas, and the results reported under the "Subgroup of patients with pNET" subheading are specific to pNETs. For example, in the subgroup of patients with CNS hemangioblastoma, the ORR = 44% refers to response outcomes observed in CNS hemangioblastomas only.*"[4] The EAG is satisfied with this clarification.

3.2.6.2 Overall response rate (ORR)

Section B.2.6 of the CS included the following statements: "*The confirmed ORR among the 61 participants with RCC in the Efficacy Analysis Set was 63.9% (95% CI: 50.6, 75.8), with a rate and associated lower 95% CI >50% (i.e. even at the lowest estimate of efficacy at least half of patients experience CR or PR), this is demonstrative of the efficacy of Belzutifan in treating these tumours, as such tumours do not shrink/respond spontaneously in the absence of effective treatment.*"[5] Further details are shown in Table 3.14 and Figure 3.2 below.

Table 3.14: MK-6482-004 summary of best overall tumour response for RCC tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	**01-JUN-2020**	**01-DEC-2020**	**15-JUL-2021**	**01-APR-2022**
Best Overall Response, n (%)				
Complete Response (CR)	[redacted]	0	[redacted]	4 (6.6)
Partial Response (PR)	[redacted]	30 (49.2)	[redacted]	35 (57.4)
Stable Disease (SD)	[redacted]	30 (49.2)	[redacted]	21 (34.4)
Progressive Disease (PD)*	[redacted]	0	[redacted]	0
Not Evaluable (NE)	[redacted]	1 (1.6)	[redacted]	1 (1.6)
Ongoing with unconfirmed response, n (%)	[redacted]	4 (6.6)	[redacted]	3 (4.9)
Ongoing without a response, n (%)	[redacted]	20 (32.8)	[redacted]	7 (11.5)
Overall response rate CR + PR (ORR), n (%)	[redacted]	30 (49.2)	[redacted]	39 (63.9)
95% confidence interval	[redacted]	(36.1, 62.3)	[redacted]	(50.6, 75.8)
90% confidence interval	[redacted]	(38.0, 60.4)	[redacted]	(52.6, 74.2)
Disease Control Rate CR + PR + SD (DCR), n (%)	[redacted]	60 (98.4)	[redacted]	60 (98.4)
95% confidence interval	[redacted]	(91.2, 100.0)	[redacted]	(91.2, 100.0)
90% confidence interval	[redacted]	(92.5, 99.9)	[redacted]	(92.5, 99.9)

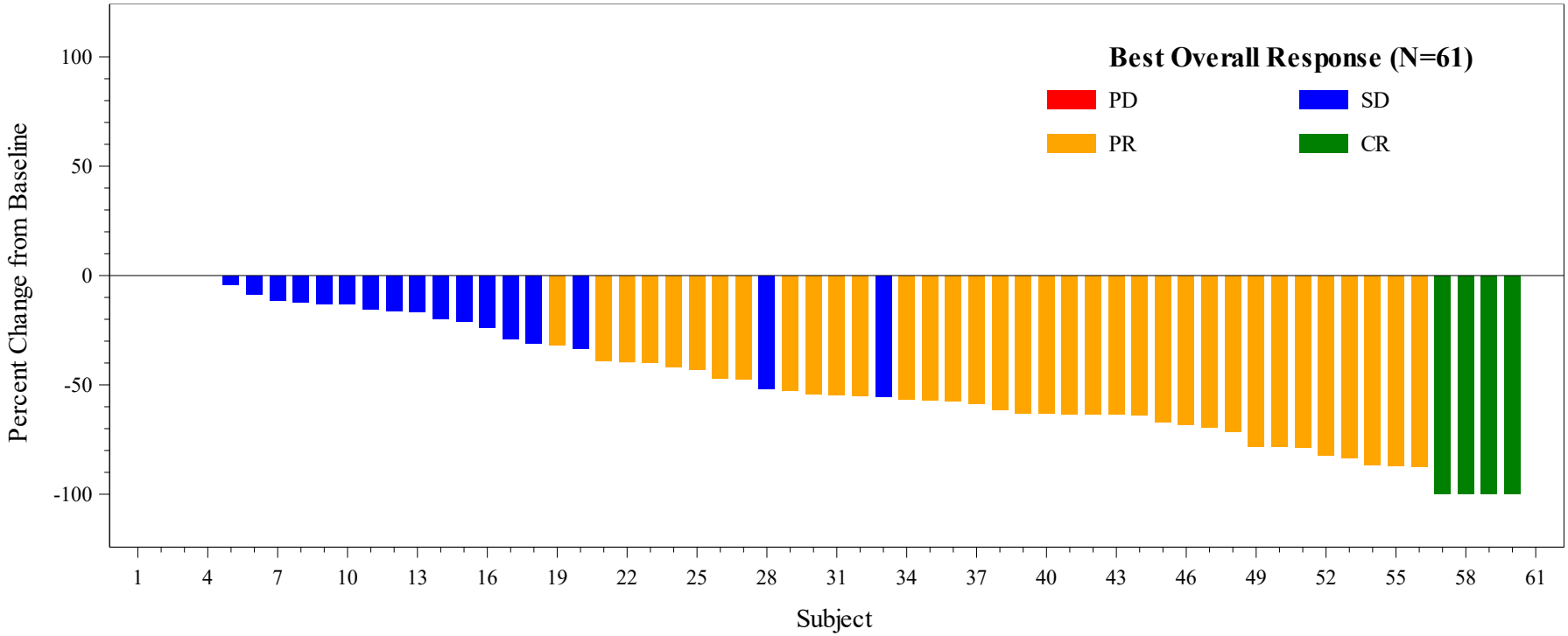
Based on Table 16 of the CS.[5]

The following table footnotes were included in the CS:[5]

- Note: 95% and 90% CIs are constructed using 2-sided Clopper-Pearson method.
- Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.

CI = confidence interval; CR = complete response; CS = company submission; DCR = disease control rate; IRC = Independent Review Committee; NE = not evaluable; ORR = overall response rate; PD = progressive disease; PR = partial response; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours; SD = stable disease

Figure 3.2: Waterfall plot - percentage change in total sum of RCC target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)



Based on Figure 3 of the CS.[5]

The following table footnotes were included in the CS:[5]

- Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
- Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).

Number (%) of patients with maximum % reduction in sum of diameters of target lesions <0 = 56 (91.8), i.e., 98.1% of patients had their tumour reduce in size at some point during follow-up in their RCC target lesions.

- Date of Data Cut-off: 01APR2022

CR = complete response; CS = company submission; PD = progressive disease; PR = partial response; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours; SD = stable disease

The company further explained (in Section B.2.6 of the CS)[5] that, "*Of the four patients who experienced a complete response in their target RCC tumour by the 01-APR-2022 data cut-off date, their target RCC tumour* [redacted] *from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with Belzutifan persist. For the 35 patients who had experienced a partial response by the 01-APR-2022 data cut-off date, the change in their target RCC tumour size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date are shown in* [Figure 3.3 below].

[redacted]

[redacted]

[redacted]

[redacted]"[5]

Figure 3.3: Spider plot - percentage change in total sum of RCC target lesion diameters from date of partial response (RECIST 1.1) – independent review committee (efficacy analysis set)

Based on Figure 5 of the CS.[5]
Footnote from the CS: Date of Data Cut-off: 01APR2022.[5]
CS = company submission; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours

3.2.6.3 Duration of response (DOR)

In Section B.2.6 of the CS[5] the company explained that, "*In the 39 patients for whom CR or PR was recorded…the median DOR was not reached as of the 01-APR-2022 database cut-off date (50% of the patients who had CR or PR need to have subsequently had disease progression or death in order for median DOR to be calculated, but only 5 such patients [12.8%] had progressed or died by the 01-APR-2022 data cutoff date). Bearing in mind that at the 01-APR-2022 data cut-off date the median length of follow-up is 37.7 months and the median time-to-response is 11.1 months…the fact that only 12.8% of patients who had CR or PR have subsequently had disease progression or death at this data cut-off date is indicative of a durable response. The range of DOR was 5.4+ to 25.8+ months*".[5] This information is summarised in Table 3.15 and Figure 3.4.

Table 3.15: MK-6482-004 summary of DOR for RCC tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with Confirmed Response, n (%)	39 (63.9)
Responders who Progressed or Died (%)	5 (12.8)
Duration of Response (Months) 95% CI	
n	39
Mean [1]	23.5
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (19.3, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	5.4+, 35.8+
Number (%) of Patients with Extended Response Duration [2]	
≥6 Months	36 (100.0)
≥12 Months	35 (100.0)
≥18 Months	29 (93.5)
≥24 Months	22 (86.6)
≥30 Months	10 (86.6)
≥36 Months	0 (NR)

Based on Table 17 of the CS.[5]
The following table footnotes were included in the CS:[5]

- Duration of Response is analysed using the Kaplan-Meier estimator. Median, first and third quartiles of Duration of Response is reported along with 95% Brookmeyer-Crowley confidence intervals.
- [1] Arithmetic mean.
- [2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
- + indicates there was no progressive disease by the time of last disease assessment.
- Date of Data Cut-off: 01APR2022

CI = 95% confidence interval; CS = company submission; DOR = duration of response; IRC = Independent Review Committee; max = maximum; min = minimum; n = number of participants; NE = not estimable; Q = quartile; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours

Figure 3.4: MK-6482-004 Kaplan-Meier (KM) plot of duration of response for RCC tumours (RECIST 1.1) – IRC (efficacy analysis set)

Based on figure 6 of the CS.[5]

The following table footnotes were included in the CS:[5]

This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.

- Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
- Date of Data Cut-off: 01APR2022

95% CI = 95% confidence interval; CS = company submission; IRC = Independent Review Committee; KM = Kaplan-Meier; NE = not estimable; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours

3.2.6.4 Time to response (TTR)

In Section B.2.6 of the CS[5] the company reported that, "*The median TTR was 11.1 months (range: 2.7 to 30.5 months) among 39 participants with a confirmed best overall response (BOR) of CR or PR*".[5] The relevant details are shown in Table .

Table 3.16: MK-6482-004 summary of TTR for RCC tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with Confirmed Response, n (%)	39 (63.9)
Time to Response (Months)	
n	39
Mean (SD)	12.4 (8.08)
Median	11.1
Min, Max	2.7, 30.5

Based on Table 18 of the CS.[5]
Footnote from the CS: Date of Data Cut-off: 01APR2022.[5]
CS = company submission; IRC = Independent Review Committee; max = maximum; min = minimum; n = number of participants; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours; SD = standard deviation; TTR = time to response

3.2.6.5 Progression-free survival (PFS)

The company reported the following: "*The median (95% CI) PFS was* [redacted] *months. The PFS rate at Month 36 was* [redacted]".[5] Further details are presented in (Table and Figure 3.5 below.

Table 3.17: MK-6482-004 summary of PFS for RCC tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Subjects with Events, n (%)	[redacted]
Progression Disease	[redacted]
Death	[redacted]
Censored Subjects, n (%)	[redacted]
New Anticancer Therapy Initiated	[redacted]
No Baseline or Post-Baseline Tumour Assessment	[redacted]
Death or Progression after More than One Missed Assessments	[redacted]
No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1]	
Median (95% CI)	[redacted]
Q1 (95% CI)	[redacted]
Q3 (95% CI)	[redacted]

	Belzutifan (N=61)
Progression-Free Survival Rate (%) (95% CI) [number at risk] at	
Month 6	[redacted]
Month 12	[redacted]
Month 18	[redacted]
Month 24	[redacted]
Month 30	[redacted]
Month 36	[redacted]
Month 42	[redacted]
Month 48	[redacted]

Based on Table 19 of the CS.[5]

The following table footnotes were included in the CS:[5]

- [1] PFS are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley CIs.
- Date of Data Cut-off: 01APR2022

95% CI = 95% confidence interval; CS = company submission; IRC = Independent Review Committee; n = number of participants; NE = not estimable; PFS = progression free survival; Q = quartile; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours

Figure 3.5: Kaplan-Meier plot of PFS for RCC tumours (RECIST 1.1) – independent review committee (efficacy analysis set)

Based on Figure 7 of the CS.[5]

The following table footnotes were included in the CS:[5]

- PFS is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalized Brookmeyer-Crowley method. Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
- Date of Data Cut-off: 01APR2022

95% CI = 95% confidence interval; CS = company submission; NE = not estimable; PFS = progression-free survival; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours.

EAG comment:

Data on OS were not presented in the CS[5] and the EAG asked the company to provide these, even if the data were immature (clarification question A.24). The company replied as follows: "*As described in section B.2.6 Table 13 of the company submission, only two patients had died by the time of the latest 01-APR-2022 data cut-off date of the MK-6482-004 study. Performing an analysis of overall survival based on only two deaths (one due to suicide and one due to acute fentanyl toxicity, as stated in the section B.2.10 of the company submission) in a study with only 61 participants would be highly inappropriate and clearly will not yield results that could be meaningfully or appropriately used for decision-making on the clinical effectiveness of Belzutifan. Therefore, this analysis has not been performed.*"[4] This amounts to a discrepancy between the study data and the outcomes described in the NICE Final Scope[7] and therefore means that the CS cannot address this part of the Scope. The EAG suggests that there is persisting uncertainty relating to alignment between the (as yet) low number of deaths observed in the MK-6482-004 study and the number in the UK target population. The possibility of a misalignment is highlighted by the company's mention of the immaturity of the OS data in their response to clarification question B.2h.[4]

3.2.6.6 Time to surgery (TTS)

In Section B.2.6 of the CS, the company stated the following: "*At the 01-APR-2022 database cut-off date, 7 patients (11.5%) had undergone surgery, the median time to surgery is not evaluable*."[5] Table shows further details.

Table 3.18: Summary of TTS for RCC tumours (efficacy analysis set)

	Belzutifan (N=61)
Number of Subjects Undergo Surgeries, n (%)	7 (11.5)
Time to Surgery (Months) 95% CI	
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (39.2, NE)
Q3 (95% CI)	NE (NE, NE)

Based on Table 20 of the CS.[5]
The following table footnotes were included in the CS:[5]

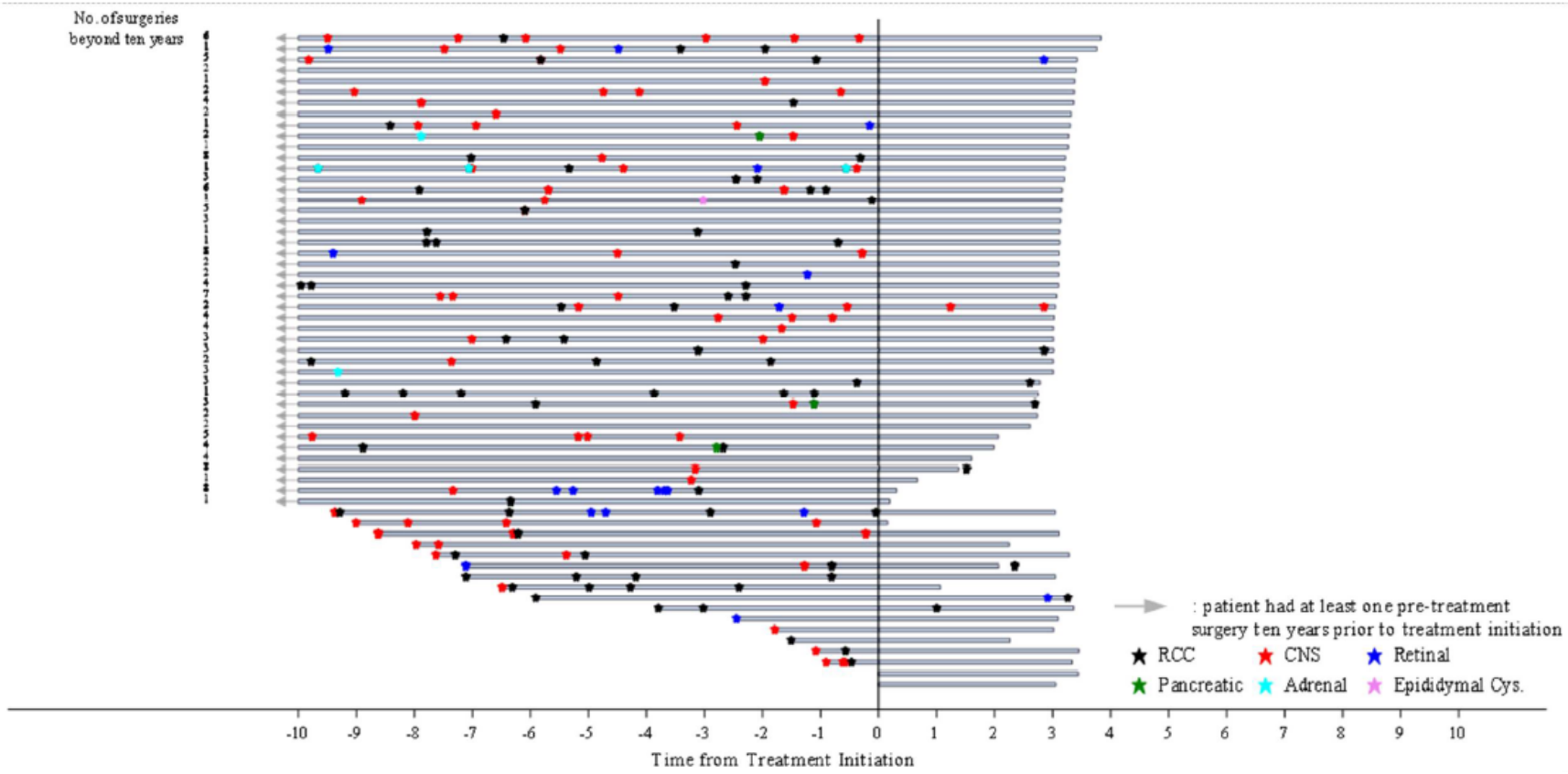
- The Q1, median, Q3, and 95% CI are obtained from Kaplan-Meier estimates.
- Surgery includes any procedure, excluding radiation, which leads to reduction of RCC tumour size.
- Date of Data Cut-off: 01APR2022.

95% CI = 95% confidence interval; CS = company submission; n = number of participants; NE = not estimable; Q = quartile; RCC = renal cell carcinoma; TTS = time to surgery

3.2.6.7 Rate of surgeries

In Section B.2.6 of the CS,[5] the company outlined the following: "*A comparison of the VHL disease-associated tumour-related surgeries patients underwent before and after initiation of treatment with Belzutifan is shown in Figure 8 [Figure 3.6 below] (note that only pre-treatment surgeries less than 10 years prior to treatment initiation are presented). From this it can be seen that that the frequency of VHL disease-associated surgeries in the time period after initiation of treatment with Belzutifan is lower than observed in the time period before, which is indicative of a potentially practice-changing favourable effect of Belzutifan treatment on subsequent rate of VHL disease-associated surgeries.*"[5] A representation of the data is provided in Figure 3.6 below.

Figure 3.6: Distribution of all surgeries pre- and post-treatment initiation over time for individual patients - safety analysis set



Based figure 8 of the CS.[5]

The following table footnotes were included in the CS:[5]

- Horizontal bars represent each patient.
- Only pre-treatment surgeries less than 10 years prior to treatment initiation are presented.
- Length of the bars on the right side of the y-axis represents duration of treatment at time of data cut-off.
- Surgery is defined as a tumour reduction procedure excluding radiation.
- Date of Data Cut-off: 01-APR-2022.

CNS = central nervous system; CS = company submission; cys = cystadenoma; RCC = renal cell carcinoma

3.2.7 Sub-grouping

In Section B.2.7 of the CS, the company provided the following explanation of numbers in subgroups defined according to tumour type and combination of tumour types:[5]

- "*Of the total 61 patients with RCC (results for this population reported in section B.2.6):*
  - *50 of the 61 patients with RCC also had CNS hemangioblastomas (results for this population reported later in this section B.2.7. Please note that 50 of the 61 patients with RCC also had CNS hemangioblastomas according to IRC, this number differs slightly to the number of patients with RCC also had CNS hemangioblastomas reported in the baseline characteristics table in Table 10 as being 51 as the 51 is the number according to only investigator assessment and not IRC).*
    - *17 of the 50 patients with RCC and CNS hemangioblastomas also had pNETs (results for this population not reported separately)*
    - *33 of the 50 patients with RCC and CNS hemangioblastomas did not also have pNETs (results for this population not reported separately)*
  - *22 of the 61 patients with RCC also had pNETs (results for this population reported later in this section B.2.7)*
    - *17 of the 22 patients with RCC and pNETs also had CNS hemangioblastomas (results for this population not reported separately)*
    - *5 of the 22 patients with RCC and pNETs did not also have CNS hemangioblastomas (results for this population not reported separately)*"[5]

The company also represented the above information in a Venn diagram (shown previously, Figure 3.1).

EAG comment:

The EAG asked the company to provide baseline and outcome data for the subgroups of patients with: RCC and CNS Hb and pNETs (n=17); RCC and CNS Hb but not pNETs (n=33); and RCC and pNETs but not CNS Hb (n=5) (clarification question A.25). The company declined to provide the requested data, whilst commenting as follows:[4]

- "*Such data would be of extremely limited value to decision making due to the small sample sizes involved (as low as n=5 as described in the question).*
- *These highly specific subgroups do not reflect the GB marketing authorisation or NICE decision problem for this indication which are for a patient population that is not restricted to these subgroups or excludes these subgroups.*
- *There is no evidence or scientific rationale to expect that the treatment effect of Belzutifan in any one tumour is affected by the presence of any other tumours in the same patient at the same time, so such subgroup data would not provide any additional information useful for the assessment of the clinical effectiveness of treatment with Belzutifan.*
- *The cost-effectiveness analysis in the company submission is based on the treatment of effect of Belzutifan on individual tumours, so such subgroup data will not provide any additional information useful for assessing the cost-effectiveness of treatment with Belzutifan in this indication.*"[4]

The rationale for the first comment is not clear since some subgroups were larger than n=5, i.e., n=17 and n=33, with the latter comprising over 50% of the total study patient cohort. The absence of these data as requested by the EAG hinders the understanding and comparison of differences between subgroups before and after treatment. In addition, the company's later argument about declining to provide information on subgroups that are not mentioned within the SmPC for Belzutifan,[6] does not

seem consistent with information in the CS (Section B.2.7) which outlines results for "*Other tumours*" (pages 95 to 96 of the CS).[5] Other tumours as described in the CS comprised pancreatic lesions (both pNETs and non-pNETs), retinal Hb, adrenal lesions, endolymphatic sac tumours and epididymal cystadenomas,[5] none of which are mentioned in the SmPC for Belzutifan.[6]

The following subsections summarise the results for the subgroups of patients with VHL RCC who also had CNS Hb or pNETs or other tumours.

3.2.7.1 Central nervous system (CNS) hemangioblastomas (Hb)

Overall response rate

The confirmed ORR among the 50 participants with CNS Hb at baseline per the IRC assessment was 44.0% (95% CI: 30.0 – 58.7). Four of the patients (8.0%) achieved a best overall response (BOR) of CR and 18 patients (36.0%) achieved a BOR of partial response (PR).[5] Further details are shown in Table 3.19 and Figure 3.7.

The company also provided the following statements:

"*Of the four patients in whom a complete response was reported in their target CNS hemangioblastoma by the 01-APR-2022 data cut-off date, their target tumour* [redacted] *from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with Belzutifan persist. For the 18 patients in whom a partial response was reported by the 01-APR-20223 data cut-off date, the change in their target CNS hemangioblastoma size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date… It can be seen that* [redacted]."[5] These data are represented in Figure 3.8.

Table 3.19: MK-6482-004 summary of best overall tumour response for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Patients with VHL Disease Associated CNS Hb at Baseline, N1 (N1/N%)	[redacted]	50 (82.0)	[redacted]	50 (82.0)
Best Overall Response, n (n/N1%)				
Complete Response (CR)	[redacted]	3 (6.0)	[redacted]	4 (8.0)
Partial Response (PR)	[redacted]	12 (24.0)	[redacted]	18 (36.0)
Stable Disease (SD)	[redacted]	31 (62.0)	[redacted]	23 (46.0)
Progressive Disease (PD)	[redacted]	2 (4.0)	[redacted]	3 (6.0)
Not Evaluable (NE)	[redacted]	2 (4.0)	[redacted]	2 (4.0)
Ongoing with unconfirmed response, n (n/N1%)	[redacted]	2 (4.0)	[redacted]	1 (2.0)
Ongoing without a response, n (n/N1%)	[redacted]	28 (56.0)	[redacted]	13 (26.0)
Overall response rate CR + PR (ORR), n (n/N1%)	[redacted]	15* (30.0)	[redacted]	22 (44.0)
95% Confidence interval	[redacted]	(17.9, 44.6)	[redacted]	(30.0, 58.7)
90% Confidence interval	[redacted]	(19.5, 42.4)	[redacted]	(32.0, 56.6)
Disease Control Rate CR + PR + SD (DCR), n (n/N1%)	[redacted]	46 (92.0)	[redacted]	45 (90.0)
95% Confidence interval	[redacted]	(80.8, 97.8)	[redacted]	(78.2, 96.7)
90% Confidence interval	[redacted]	(82.6, 97.2)	[redacted]	(80.1, 96.0)

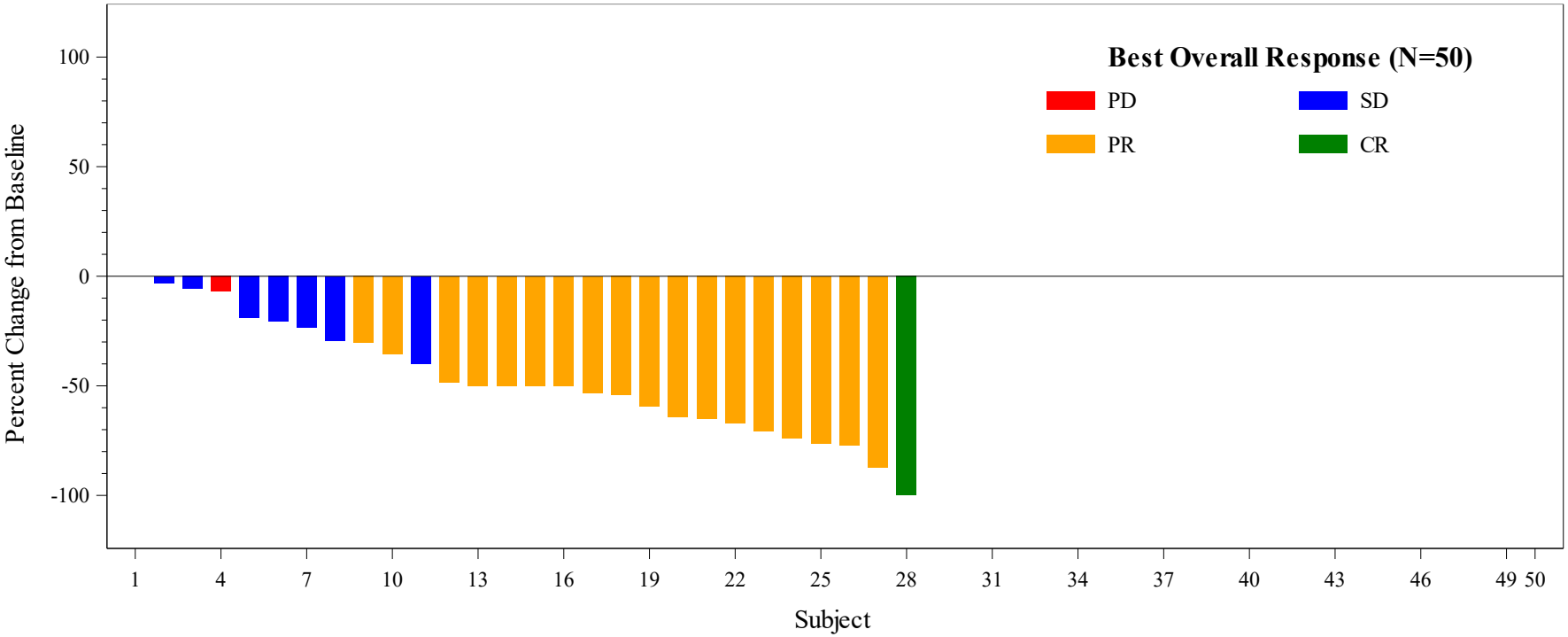
Based on Table 21 of the CS.[5]

The following table footnotes were included in the CS:[5]

- Note: 95% and 90% confidence intervals are constructed using 2-sided Clopper-Pearson method.
- Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.
- Patients evaluable at baseline per IRC are included.
- *There were changes in assessments of imaging data that resulted in an overall decrease in the number of participants with responses for CNS hemangioblastoma (from 16 to 15 participants with confirmed response) and an overall decrease in the number of PFS events for pancreatic neoplasms (1 less PFS event) and CNS hemangioblastomas (1 less PFS event) since submission of initial application, at the 01-DEC-2020 data cut-off date.

CNS = central nervous system; CR = complete response; CS = company submission; DCR = disease control rate; Hb – haemangioblastoma; IRC Independent Review Committee; NE = not evaluable; ORR = overall response rate; PD = progressive disease; PFS = progression free survival; PR = partial response; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours; SD = stable disease; VHL = Von Hippel-Lindau

Figure 3.7: Waterfall plot - percentage change in total sum of CNS Hb target lesions diameters from baseline to post-baseline maximum % reduction (RECIST 1.1) – independent review committee (efficacy analysis set)



Based on Figure 10 of the CS.[5]

The following footnotes were included in the CS:[5]

- Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
- Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).

Number (%) of patients with maximum % reduction in sum of diameters of target lesions < 0 = 27 (54.0), i.e., 54.0% of patients had their tumour reduce in size at some point during follow-up in their CNS hemangioblastoma target lesions.

- Date of Data Cut-off: 01APR2022

CNS = central nervous system; CR = complete response; CS = company submission; PD = progressive disease; PR = partial response; RECIST = response evaluation criteria in solid tumours; SD = stable disease.

Figure 3.8: Spider plot - Percentage change in total sum of CNS Hb target lesion diameters from date of PR (RECIST 1.1) – independent review committee (efficacy analysis set)

Based on Figure 11 of the CS.[5]
Footnote from the CS: Database cut-off date: 01APR2022.[5]
Hb = haemangioblastoma; CNS = central nervous system; CS = company submission; PR = partial response; RECIST = response evaluation criteria in solid tumours

Duration of response

The company stated that: "*The median DOR was not reached as of the 01-APR-2022 database cut-off date. The range of DOR was 3.7+ to 38.7+ months, 12 patients achieved a DOR ≥30 months.*"[5] Further details are shown in Table and **Error! Reference source not found.**.

Table 3.20: MK-6482-004 summary of DOR for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1/N	50 (82.0)
Patients with Confirmed Response, n (n/N1%)	22 (44.0)
Responders who Progressed or Died (%)	4 (18.2)
Duration of Response (Month) 95% CI	
Mean [1]	23.9
Median (95% CI)	NE (30.9, NE)
Q1 (95% CI)	31.3 (5.5, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	3.7+, 38.7+
Number (%) of Patients with Extended Response Duration [2]	
≥6 Months	19 (95.2)
≥12 Months	16 (90.2)
≥18 Months	14 (90.2)
≥24 Months	13 (90.2)
≥30 Months	12 (90.2)
≥36 Months	2 (72.2)
≥42 Months	0 (0)

Based on Table 22 of the CS.[5]

The following footnotes were included in the CS:[5]

- DOR is analysed using the Kaplan-Meier estimator. Median, first and third quartiles of DOR are reported along with 95% Brookmeyer-Crowley CIs.
- [1] Arithmetic mean.
- [2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
- + indicates there was no progressive disease by the time of last disease assessment.
- Date of Data Cut-off: 01APR2022

CI = confidence interval; CNS = central nervous system; CS = company submission; DOR = duration of response; Hb = haemangioblastoma; IRC = Independent Review Committee; NE = not estimable; Q = quartile; RECIST = response evaluation criteria in solid tumours; VHL = Von Hippel-Lindau

Figure 3.9: MK-6482-004 Kaplan-Meier plot of DOR for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

Based on Figure 12 of the CS.[5]
The following footnotes were included in the CS:[5]
This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.

- Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
- Date of Data Cut-off: 01APR2022

CNS = central nervous system; CS = company submission; DOR = duration of response; Hb = haemangioblastoma; IRC = Independent Review Committee; NE = not estimable; RECIST = response evaluation criteria in solid tumours

Time to response (TTR)

The company summarised the data for this outcome in the CS as follows: "*The median TTR was [redacted] participants with a confirmed BOR of CR or PR*".[5] The details are summarised in Table).

Table 3.21: MK-6482-004 summary of TTR for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1 (N1/N%)	[redacted]
Patients with Confirmed Response, n (n/N1%)	[redacted]
Time to Response (months)	
n	[redacted]
Mean (SD)	[redacted]
Median	[redacted]
Min, Max	[redacted]

Based on Table 23 of the CS.[5]
Footnote from the CS: Date of Data Cut-off: 01APR2022.[5]
CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; IRC = Independent Review Committee; max = maximum; min = minimum; n = number of patients; RECIST = response evaluation criteria in solid tumours; SD = standard deviation; TTR = time to response; VHL = Von Hippel-Lindau.

Progression-free survival (PFS)

The company stated that: "*The median PFS for patients with CNS hemangioblastoma [redacted] at the 01-APR-2022 database cut-off date, [redacted] patients had a PFS event*".[5] Further details are shown in Table and **Error! Reference source not found.**.

Table 3.22: MK-6482-004 summary of PFS for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated CNS Hemangioblastomas at Baseline, N1/N	[redacted]
Subjects with Events, n (n/N1 %)	[redacted]
Progression Disease	[redacted]
Death	[redacted]
Censored Subjects, n (n/N1 %)	[redacted]
New Anticancer Therapy Initiated	[redacted]
No Baseline or Post-Baseline Tumour Assessment	[redacted]
Death or Progression after More than One Missed Assessments	[redacted]
No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1]	
Median (95% CI)	[redacted]

	Belzutifan (N=61)
Q1 (95% CI)	[redacted]
Q3 (95% CI)	[redacted]

Based on Table 24 of the CS.[5]

The following footnotes were included in the CS:[5]

- [1] Progression-Free Survival are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley confidence intervals.
- Date of Data Cut-off: 01APR2022

CI = confidence interval; CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; IRC = Independent Review Committee; NE = not estimable; PFS = progression free survival; Q = quartile; RECIST = response evaluation criteria in solid tumours; VHL = Von Hippel-Lindau

Figure 3.10: Kaplan-Meier plot of PFF for CNS Hb (RECIST 1.1) – IRC (efficacy analysis set)

Based on Figure 13 of the CS.[5]

The following footnotes were included in the CS:[5]

- Progression-Free Survival is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method. Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.
- Date of Data Cut-off: 01APR2022

CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; IRC = Independent Review Committee; NE = not estimable; PFS = progression-free survival; RECIST = response evaluation criteria in solid tumours.

Time to surgery (TTS)

The company's statement regarding this outcome was as follows: "*At the 01-APR-2022 database cut-off date, only one patient with CNS hemangioblastoma had undergone surgery. Consequently, the median time to surgery is not evaluable for this subgroup.*"[5]

Rate of surgeries

The following brief description was provided in the CS for rate of surgeries: "*A comparison of the VHL disease-associated tumour-related surgeries these patients underwent before and after initiation of treatment with Belzutifan is shown in Figure 14 [**Error! Reference source not found.** below] (note that o nly the blue bars indicate patients with CNS hemangioblastomas at baseline per IRC).*"[5]

Figure 3.11: Distribution of all surgeries pre- and post-treatment initiation over time for individual patients for individual patients with baseline CNS Hb per IRC - safety analysis set

Based on Figure 14 of the CS.[5]

The following footnotes were included in the CS:[5]

- Horizontal bars represent each patient.
- Blue bars indicate patients with CNS Hb at baseline per IRC.
- Only pre-treatment surgeries less than 10 years prior to treatment initiation are presented.
- Length of the bars on the right side of the y-axis represents duration of treatment at time of data cut-off.
- Surgery is defined as a tumour reduction procedure excluding radiation.
- Date of Data Cut-off: 01-APR-2022.

CNS = central nervous system; CS = company submission; cys. = cystadenoma; Hb = haemangioblastoma; IRC = Independent Review Committee; RCC = renal cell carcinoma

3.2.7.2 Pancreatic neuroendocrine tumours

Overall response rate (ORR)

The company summarised data for the above outcome as follows:

"*The confirmed ORR among 22 participants with pancreatic neuroendocrine tumours at baseline per IRC assessment was 90.9% (95% CI: 70.8, 98.9). As of the 01-APR-2022 database cut-off date, seven participants (31.8%) achieved a BOR of CR and 13 participants (59.1%) achieved a BOR of PR*".[5] Further details are shown in Table and **Error! Reference source not found.**.

The company further explained:

"*Of the seven patients who experienced a complete response in their target pNET by the 01-APR-2022 data cut-off date, their target tumour* [redacted] *from the timepoint complete response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date, showing that complete responses that arise during treatment with Belzutifan persist. For the 13 patients who experienced a partial response by the 01-APR-20223 data cut-off date, the change in their target pNET size from the timepoint partial response (as per RECIST 1.1) was first recorded to the 01-APR-2022 data cut-off date are shown in Figure 16 [**Error! Reference source not found.**3 below]. It can be seen that* [redacted]."[5]

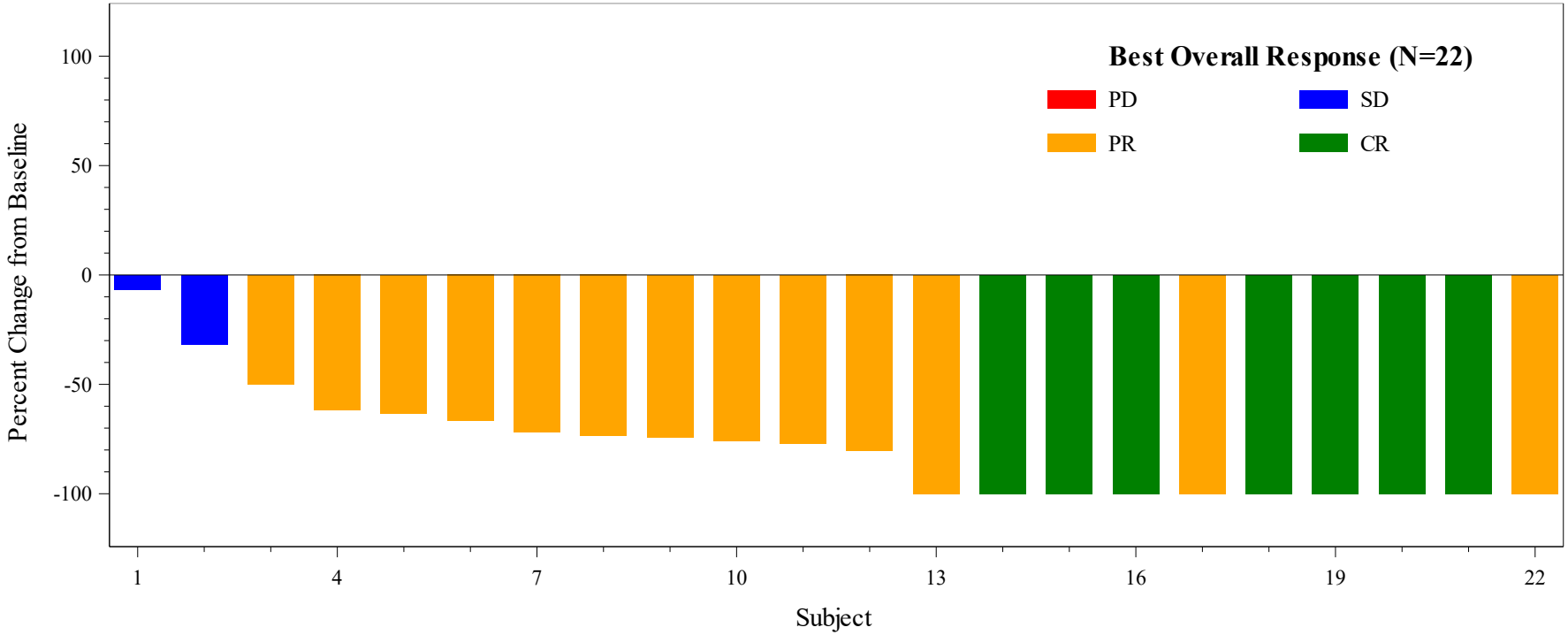
Table 3.23: MK-6482-004 summary of best overall tumour response for pancreatic neuroendocrine tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)			
Data cut-off date	01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1 (N1/N%)	[redacted]	22 (36.1)	[redacted]	22 (36.1)
Best Overall Response, n (n/N1%)				
• Complete Response (CR)	[redacted]	3 (13.6)	[redacted]	7 (31.8)
• Partial Response (PR)	[redacted]	17 (77.3)	[redacted]	13 (59.1)
• Stable Disease (SD)	[redacted]	2 (9.1)	[redacted]	2 (9.1)
• Progressive Disease (PD)	[redacted]	0	[redacted]	0
• Not Evaluable (NE)	[redacted]	0	[redacted]	0
Ongoing with unconfirmed response, n (n/N1%)	[redacted]	1 (4.5)	[redacted]	0
Ongoing without a response, n (n/N1%)	[redacted]	0	[redacted]	0
Overall response rate CR + PR (ORR), n (n/N1%)	[redacted]	20 (90.9)	[redacted]	20 (90.9)
• 95% confidence interval	[redacted]	(70.8, 98.9)	[redacted]	(70.8, 98.9)
• 90% confidence interval	[redacted]	(74.1, 98.4)	[redacted]	(74.1, 98.4)
Disease Control Rate CR + PR + SD (DCR), n (n/N1%)	[redacted]	22 (100.0)	[redacted]	22 (100.0)
• 95% confidence interval	[redacted]	(84.6, 100.0)	[redacted]	(84.6, 100.0)
• 90% confidence interval	[redacted]	(87.3, 100.0)	[redacted]	(87.3, 100.0)

Based on Table 25 of the CS.[5]
The following footnotes were included in the CS:[5]
- Note: 95% and 90% confidence intervals are constructed using 2-sided Clopper-Pearson method.
- Best overall response of RCC CR and PR should be confirmed by a second assessment at least 4 weeks after the initial response.

CR = complete response; CS = company submission; DCR = disease control rate; IRC = Independent Review Committee; NE = not evaluable; ORR = overall response rate; PD = progressive disease; PR = partial response; RCC = renal cell carcinoma; RECIST = response evaluation criteria in solid tumours; SD = stable disease; VHL = Von Hippel-Lindau

Figure 3.12: Waterfall plot - percentage change in total sum of target lesions diameters for pancreatic neuroendocrine tumours from baseline to post-baseline maximum % reduction (RECIST 1.1) – IRC (efficacy analysis set)



Based on Figure 15 of the CS.[5]

The following footnotes were included in the CS:[5]

- Subjects without either post-baseline evaluable lesion measurements or target lesions or with all post-baseline non-evaluable time-point responses appear as blank on the right of the figure.
- Note that best overall response according to RECIST 1.1 criteria is determined by more than only the change in tumour size between baseline and last measurement (see Appendix N for a description of best overall response according to RECIST 1.1).

Number (%) of patients with maximum % reduction in sum of diameters of target lesions < 0 = 22 (100.0), i.e. 100% of patients had their tumour reduce in size at some point during follow-up in their pNET target lesions.

- Date of Data Cut-off: 01APR2022

CR = complete response; CS = company submission; IRC = Independent Review Committee; PD = progressive disease; pNET = pancreatic neuroendocrine tumour; PR = partial response; RECIST = response evaluation criteria in solid tumours; SD = stable disease.

Figure 3.13: Spider plot - Percentage change in total sum of target lesion diameters for pNETs from date of partial response (RECIST 1.1) – IRC (efficacy analysis set)

Based on Figure 16 of the CS.[5]
Footnote from the CS: Database cut-off date: 01APR2022.[5]
CS = company submission; IRC = Independent Review Committee; pNET = pancreatic neuroendocrine tumour; RECIST = response evaluation criteria in solid tumours

Duration of response (DOR)

The company explained the following in relation to this outcome: "*The median DOR was not reached as of the 01-APR-2022 database cut-off date, no patients in this subgroup had progression or died by the 01-APR-2022 data cutoff date. The range of DOR was 11.0+ to 37.3+ months, 15 participants achieved a DOR ≥24 months*".[5] Further details are presented in Table 3.24 and Figure 3.14.

Table 3.24: MK-6482-004 summary of DOR for pancreatic neuroendocrine tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	22 (36.1)
Patients with Confirmed Response, n (n/N1%)	20 (90.9)
Responders who Progressed or Died (%)	0
Duration of Response (Months) 95% CI	
Mean [1]	27.4
Median (95% CI)	NE (NE, NE)
Q1 (95% CI)	NE (NE, NE)
Q3 (95% CI)	NE (NE, NE)
Min, Max	11.0+, 37.3+
Number (%) of Patients with Extended Response Duration [2]	
≥6 Months	20 (100.0)
≥12 Months	19 (100.0)
≥18 Months	19 (100.0)
≥24 Months	15 (100.0)
≥30 Months	8 (100.0)
≥36 Months	1 (100.0)
≥42 Months	0 (0)

Based on Table 26 of the CS.[5]

The following footnotes were included in the CS:[5]

- Duration of Response is analysed using the Kaplan-Meier estimator. Median, first and third quartiles of Duration of Response are reported along with 95% Brookmeyer-Crowley confidence intervals.
- [1] Arithmetic mean.
- [2] % is calculated by Kaplan-Meier method. For the patients without extended response duration at each duration threshold, they either experienced disease progression or death or their response duration had not reached that duration threshold yet.
- + indicates there was no progressive disease by the time of last disease assessment.
- Date of Data Cut-off: 01APR2022

CI = confidence interval; CS = company submission; DOR = duration of response, IRC = Independent Review Committee; max = maximum; min = minimum; NE = not estimable; Q = quartile; RECIST = response evaluation criteria in solid tumours

Figure 3.14: MK-6482-004 Kaplan-Meier plot of duration of response for pancreatic neuroendocrine tumours (RECIST 1.1) – IRC (efficacy analysis set)

Based on Figure 17 of the CS.[5]

The following footnotes were included in the CS:[5]

This figure shows the proportion of patients (1.0 = 100%) still with response (have not had tumour progression or have died) at timepoints measured from the first recording of confirmed response (at Time (Months) = 0). Note that the number at risk changes with time which affects the denominator and data point relative to the y-axis in this figure.

- Duration of Response is analysed using the Kaplan-Meier estimator and their 95% confidence intervals are estimated with the generalised Brookmeyer-Crowley method.
- Date of Data Cut-off: 01APR2022

CI = confidence interval; CS = company submission; IRC = Independent Review Committee; NE = not estimable; RECIST = response evaluation criteria in solid tumours.

Time to response (TTR)

The company provided the following brief statement in relation to the outcome of TTR: "*The median TTR was* [redacted] *participants with a confirmed BOR of CR or PR*".[5] Further details are presented in Table 3.25.

Table 3.25: MK-6482-004 summary of TTR for pancreatic neuroendocrine tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	[redacted]
Patients with Confirmed Response, n (n/N1%)	[redacted]
Time to Response (Months)	
n	[redacted]
Mean (SD)	[redacted]
Median	[redacted]
Min, Max	[redacted]

Based on Table 27 of the CS.[5]
Footnote from the CS: Date of Data Cut-off: 01APR2022.[5]
CS = company submission; IRC = Independent Review Committee; max = maximum; min = minimum; n = number of patients; RECIST = response evaluation criteria in solid tumours; SD = standard deviation; TTR = time to response; VHL = Von Hippel-Lindau

Progression-free survival (PFS)

The company stated that: "*The median PFS for patients with pancreatic neuroendocrine tumours* [redacted] *at the 01-APR-2022 database cut-off date,* [redacted] *had a PFS event.*"[5] The relevant details are shown in Table 3.26.

Table 3.26: MK-6482-004 summary of PFS for pancreatic neuroendocrine tumours (RECIST 1.1) – IRC (efficacy analysis set)

	Belzutifan (N=61)
Patients with VHL Disease Associated Pancreatic Neuroendocrine Tumours at Baseline, N1/N	[redacted]
Subjects with Events, n (n/N1 %)	[redacted]
Progression Disease	[redacted]
Death	[redacted]
Censored Subjects, n (n/N1 %)	[redacted]
New Anticancer Therapy Initiated	[redacted]
No Baseline or Post-Baseline Tumour Assessment	[redacted]
Death or Progression after More than One Missed Assessments	[redacted]
No Progression at the Time of Data Cut-Off or Before End of Treatment	[redacted]
Progression-Free Survival (Months) [1]	
Median (95% CI)	[redacted]
Q1 (95% CI)	[redacted]
Q3 (95% CI)	[redacted]

Based on Table 28 of the CS.[5]
The following footnotes were included in the CS:[5]

- [1] Progression-Free Survival are analysed using the Kaplan-Meier estimator. Median, first and third quartiles of PFS are reported along with 95% Brookmeyer-Crowley confidence intervals.

	Belzutifan (N=61)
• Date of Data Cut-off: 01APR2022 CI = confidence interval; CS = company submission; IRC = Independent Review Committee; NE = not estimable; PFS = progression-free survival; Q = quartile; RECIST = response evaluation criteria in solid tumours; VHL = Von Hippel-Lindau	

Time to surgery (TTS)

The company made the following comment: "*At the 01-APR-2022 database cut-off date, no patient with pancreatic neuroendocrine tumour had undergone surgery. Consequently, the time to surgery is not evaluable for this subgroup. Data sources used to estimate time to surgery in this population for the cost-effectiveness analyses are described in section B.3 [of the CS].*"[5]

3.2.7.3 Other tumours

The MK-6482-004 study collected data on several other tumour types in addition to RCCs, pNETs and CNS Hb. These "*other*" tumours included: pancreatic lesions (both pNETs and non-pNETs); retinal Hb; adrenal lesions and endolymphatic sac tumours; and epididymal cystadenomas. The company provided brief details of outcomes for each of these tumour types, and the statements made are reproduced below.[5]

Pancreatic lesions

"*Results from the MK-6482-004 study were collected for the pancreatic lesions subgroup, which included both pNET and non-pNET lesions, pNET lesions were defined as solid parenchymal lesions that do not communicate with the pancreatic duct, while non-pNET lesions were defined as all pancreatic lesions that were not pNET lesions.*"[5]

"Treatment with Belzutifan showed [redacted] *ORR in participants with pancreatic lesions; the ORR by IRC was* [redacted]*. The DCR for pancreatic lesions was* [redacted]*."*[5]

"The median DOR for participants with pancreatic lesions was [redacted]*, and based on Kaplan-Meier estimation,* [redacted] *of responders had an ongoing response at 30 months. The median TTR was* [redacted]*. Median PFS and TTS were* [redacted]*.* [redacted] *underwent surgery for pancreatic lesions as of the 01-APR-2022 data cut-off date.*"[5]

Retinal hemangioblastoma

"*Twelve of 17 participants in the MK-6482-004 study with baseline retinal hemangioblastomas were evaluable for response with follow-up evaluations. Of 12 participants with retinal hemangioblastoma, evaluable retinal hemangioblastomas were determined in 16 eyes at baseline per IRC assessment*".[5]

"Treatment response in retinal hemangioblastoma, per IRC, was assessed using multiple parameters such as number/size/location, degree of feeder/drainer engorgement (mild/prominent), presence of intraretinal heme, presence of preretinal heme, presence of vitreous heme, presence of lipid exudation, presence of subretinal fluid, and presence of fibrosis."[5, 24]

"An improvement of retinal hemangioblastoma was observed after treatment with Belzutifan. The response of 'Improved' was 100% (95% CI: 79.4, 100.0) in all 16 eyes and 100% (95% CI 73.5, 100.0) in all 12 participants. Median DOR was not reached. All 12 participants had an improvement for ≥12 months, and of these, 9 participants had improvement for ≥30 months. Median TTR was [redacted]*."*[5]

"Visual acuity of participants with retinal hemangioblastoma underwent ophthalmologic evaluation (by investigator assessment). Visual acuity in most participants [redacted]."[5]

Adrenal lesions and endolymphatic sac tumours

"*As of the 01-APR-2022, per investigator assessment,* [redacted] *with adrenal lesions (n=3) and endolymphatic sac tumours (n=1) had a BOR of* [redacted]*; median TTR* [redacted]*. Median PFS was* [redacted]."[5]

Epididymal cystadenomas

"*Sixteen participants had epididymal cystadenomas at baseline and were followed up by ultrasound examination. Per investigator assessment, at Week 49,* [redacted] *had improvement in lesions compared with baseline,* [redacted] *had stable lesions, and* [redacted] *had progressed.*"[5]

EAG comment:

The EAG asked the company to confirm the number of "*evaluable*" and "*non-evaluable*" participants with each of the above tumour types (as per Section B.2.7 of the CS[5]) and to define the two terms in the context of these tumour types (clarification questions A.28a and A.28b). The company replied as follows:

"*a) "Evaluable" and "non-evaluable" tumours were only relevant/a consideration for the retinal hemangioblastomas, and not relevant or a differentiation that was made for any of the other tumour types, in the MK-6482-004 study. Specifically for retinal hemangioblastomas, "not evaluable" refers to 1) Overall poor quality of images; or 2) Ancillary findings suggest a retinal hemangioblastomas may be present; however, tumour was not visible or cannot be assessed due to poor quality of images received.*"[4]

"*b) For the retinal hemangioblastomas, which was the only type of tumour where evaluable/non-evaluable in this context was considered in the MK-6482-004 study, whether a tumour was evaluable or not was determined during independent review based on the quality of imaging/scan data they received in order to make their assessment of tumour response. If the independent review committee determined that the imaging data they received for a tumour was not of sufficient quality to evaluate tumour response, the tumour was determined to be non-evaluable.*"[4]

The EAG also requested that the company explain how other tumour types relate to "primary" and "non-primary" tumours that are mentioned elsewhere in the CS[5] (clarification question A.28c). The company provided the following response:

"*Primary tumours referred to in the company submission refers specifically to the VHL-associated RCC, CNS Hb or pNET with the greatest burden on the patient. Therefore, "non-primary tumours" refer to non-RCC tumours in the VHL-RCC cohort, non-CNS Hb tumours in the VHL-CNS Hb cohort and non-pNET tumours in the VHL-pNET cohort. For example, non-RCC tumours refers to CNS Hb, pNET, retinal Hb, adrenal lesions etc. "Other tumours" refers to any tumour outside of RCC, CNS Hb or pNET regardless of which cohort a patient belongs to. Hence, all "other tumours" are "non-primary tumours", but not all "non-primary tumours" are "other tumours". In an VHL-RCC patient, a CNS Hb and an adrenal lesion are both considered non-primary tumours, but an adrenal lesion would also be considered "other tumours" in the context of the subgroup analysis.*"[4]

Finally, the EAG asked the company to outline the impact on treatment effect of the distribution of the other tumours within the study population (clarification A.28d). The company replied that:

"*There is no evidence to suggest that the simultaneous presence of any of these other tumours in a patient would affect the treatment effect of Belzutifan on RCC, CNS hemangioblastomas, and/or pNETs in the same patient.*"[4]

"*However, as data from the MK-6482-004 study suggest that treatment with Belzutifan may have a beneficial effect on these tumours, where they are present in addition to RCC, CNS hemangioblastomas, and/or pNETs, then it may be likely that a patient with RCC, CNS hemangioblastomas, and/or pNETs plus one or more of these other tumours may experience a greater overall beneficial treatment effect due to the additional beneficial effect on these other tumours (i.e. the more tumours a patient has, the greater the total beneficial effect, even though the level of effect for each tumour will not be affected).*"[4]

The EAG appreciates these clarifications.

3.2.8 Adverse events (AEs)

3.2.8.1 Overall AEs

An overall summary of AE data is presented in Table 3.27. With later data cut-off dates the number of AEs, proportion of Grade 3 and above AEs and serious AEs (SAEs) increased. In the latest data cut treatment-related AEs (TRAEs) were observed for all 61 participants. Two AEs resulted in death, one due to suicide and one due to acute fentanyl toxicity. Serious AEs were reported for 18 (29.5%) participants, and treatment-related SAEs were reported for four (6.6%) participants. Participants who discontinued the study intervention due to an AE was four (6.6%). Treatment-related AEs leading to dose reduced and interrupted were eight (13.1) and 13 (21.3) respectively.[5]

Table 3.27: Overview of MK-6482-004 AEs

Category		Belzutifan (N=61) n (%)			
Data cut-off date		01-JUN-2020	01-DEC-2020	15-JUL-2021	01-APR-2022
Number of AEs		[redacted]	945	[redacted]	1260
Subjects with any AEs		[redacted]	61 (100.0)	[redacted]	61 (100.0)
Subjects with any TRAEs		[redacted]	61 (100.0)	[redacted]	61 (100.0)
Subjects with any AEs of CTCAE Grade 3 and above		[redacted]	20 (32.8)	[redacted]	27 (44.3)
Subjects with any SAEs		[redacted]	11 (18.0)	[redacted]	18 (29.5)
Subjects with any treatment-related SAEs		[redacted]	3 (4.9)	[redacted]	4 (6.6)
Severity grade (Refer to NCI-CTCAE V 4.03)[a]	Mild (Grade 1)	[redacted]	10 (16.4)	[redacted]	8 (13.1)
	Moderate (Grade 2)	[redacted]	31 (50.8)	[redacted]	26 (42.6)
	Severe (Grade 3)	[redacted]	18 (29.5)	[redacted]	22 (36.1)
	Life Threatening (Grade 4)	[redacted]	1 (1.6)	[redacted]	3 (4.9)
	Death (Grade 5)	[redacted]	1 (1.6)	[redacted]	2 (3.3)
Related Severity grade (Refer to NCI-CTCAE V 4.03)[a]	Mild (Grade 1)	[redacted]	25 (41.0)	[redacted]	21 (34.4)
	Moderate (Grade 2)	[redacted]	27 (44.3)	[redacted]	29 (47.5)
	Severe (Grade 3)	[redacted]	9 (14.8)	[redacted]	11 (18.0)
Subjects with AEs leading to death		[redacted]	1 (1.6)	[redacted]	2 (3.3)
Subjects with AEs leading to treatment discontinued		[redacted]	2 (3.3)	[redacted]	4 (6.6)
Subjects with AEs leading to dose reduced		[redacted]	9 (14.8)	[redacted]	10 (16.4)
Subjects with TRAEs leading to dose reduced		[redacted]	7 (11.5)	[redacted]	8 (13.1)
Subjects with AEs leading to dose interrupted		[redacted]	26 (42.6)	[redacted]	26 (42.6)
Subjects with TRAEs leading to dose interrupted		[redacted]	14 (23.0)	[redacted]	13 (21.3)

Based on Table 42 of the CS.[5]
[a] Adverse events by maximum severity grade for subject level.
AEs =adverse events; CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; NCI = National Cancer Institute; SAEs = serious adverse events; TRAEs = treatment-related adverse events

3.2.8.2 Most frequently reported adverse events

The company reported that: "[redacted] *in the Safety Analysis Set (ApaT population) had at least 1 AE. The most frequently reported AEs (in >25% of participants) were* [redacted]".[5] The details for AEs with an incidence of at least 10% are summarised in Table 3.28.

Table 3.28: MK-6482-004 patients with AEs by decreasing incidence (incidence ≥10%) (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Subjects in population	[redacted]
with one or more Aes	[redacted]
with no AE	[redacted]
Anaemia	[redacted]
Fatigue	[redacted]
Headache	[redacted]
Dizziness	[redacted]
Nausea	[redacted]
Dyspnoea	[redacted]
Myalgia	[redacted]
Constipation	[redacted]
Arthralgia	[redacted]
Vision blurred	[redacted]
Abdominal pain	[redacted]
Alanine aminotransferase increased	[redacted]
Back pain	[redacted]
Diarrhoea	[redacted]
Upper respiratory tract infection	[redacted]
Weight increased	[redacted]
Hypertension	[redacted]
Insomnia	[redacted]
Oedema peripheral	[redacted]
COVID-19	[redacted]
Disturbance in attention	[redacted]
Urinary tract infection	[redacted]
Anxiety	[redacted]
Aspartate aminotransferase increased	[redacted]
Blood creatinine increased	[redacted]
Cough	[redacted]
Muscle spasms	[redacted]
Vomiting	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan n (%)
Based on Table 114 of Appendix F of the CS.[5] The following footnotes were included in the CS:[5] • Date of Data Cut-off: 01-APR-2022 • Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included. AE = adverse event; COVID-19 = Coronavirus Disease of 2019; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities	

3.2.8.3 Grade 3 to 5 AEs (including drug-related)

The company outlined that: [redacted] *reported 1 or more Grade 3 to 5 AEs. The most frequently reported Grade 3 to 5 AEs were* [redacted] *…All other Grade 3 to 5 AEs were reported for* [redacted]." In addition: "[redacted] *reported 1 or more AEs considered related to Belzutifan by the investigator. Most drug-related AEs were Grade 1 and Grade 2 in severity.* [redacted] *reported drug-related AEs with CTCAE Grade 3 and above…* [redacted] *reported a drug-related Grade 5 AE.*"[5] Table 3.29 shows the number of patients with AEs of grades 3 to 5 in order of decreasing incidence and also indicates those considered to be drug-related.

Table 3.29: MK-6482-004 patients with Grade 3 to 5 AEs listed by decreasing incidence (incidence ≥0%) and drug-related Grade 3 to 5 AEs (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan n (%)	
Date of Data Cut-off: 01-APR-2022	**Grade 3 to 5 Aes listed by decreasing incidence**	**Drug-related Aes CTCAE ≥Grade 3**
Subjects in population	[redacted]	[redacted]
with one or more Aes	[redacted]	[redacted]
with no AE	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Syncope	[redacted]	[redacted]
Anaphylactic reaction	[redacted]	[redacted]
Azoospermia	[redacted]	[redacted]
COVID-19	[redacted]	[redacted]
COVID-19 pneumonia	[redacted]	[redacted]
Cholecystectomy	[redacted]	[redacted]
Coronary artery dissection	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Embolism	[redacted]	[redacted]
Haemorrhage intracranial	[redacted]	[redacted]
Hyperglycaemia	[redacted]	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan n (%)	
Date of Data Cut-off: 01-APR-2022	**Grade 3 to 5 Aes listed by decreasing incidence**	**Drug-related Aes CTCAE ≥Grade 3**
Hypotension	[redacted]	[redacted]
Hypoxia	[redacted]	[redacted]
Musculoskeletal pain	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]
Non-small cell lung cancer	[redacted]	[redacted]
Otitis media chronic	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Retinal detachment	[redacted]	[redacted]
Retinal vein occlusion	[redacted]	[redacted]
Skin laceration	[redacted]	[redacted]
Suicide attempt	[redacted]	[redacted]
Toxicity to various agents	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]
Vitreous haemorrhage	[redacted]	[redacted]
Weight increased	[redacted]	[redacted]

Based on Tables 115 and 116 of Appendix F of the CS[5] and Table 12-4 of the CSR.[24]
The following footnotes were included in the CS:[5]

- Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
- Date of Data Cut-off: 01APR2022

AE = adverse event; COVID-19 = Coronavirus Disease of 2019; CS = company submission; CSR = clinical study report; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities.

3.2.8.4 Deaths due to AEs

The company reported the following: [redacted] *due to AE occurred during the study.* [redacted]."[5]

EAG comment:

As the company stated that the two deaths were not caused by the study drug, the EAG asked the company to explain how the investigator (the company) could be confident about this, especially considering that the company states that AEs leading to discontinuation of the study treatment included "*toxicity to various agents*" in one patient (Table 119 of the CS[5] and reproduced in Table 3.31 below). The EAG asked which agents specifically were relevant and also queried whether, if the toxicity arose from a non-study drug, this implied that other participants could have been taking additional medication that could have contaminated the overall findings (clarification question A33). The company summarised their arguments by stating that: "*…the underlying VHL-disease is a significant confounding factor for the concerned events. There is no evidence to suggest a possible causal association between the event and Belzutifan, which is in line with the guidelines of FDA and ICH on causality assessment.*" In addition: "*While participants in the MK-6482-004 study may have received*

concomitant medications (as described in section B.2.3 of the company submission) during the study period (as is the case for nearly all clinical trials), these are very unlikely to have contaminated the overall findings of the study as none of the concomitant medications taken would be ones that could have an antitumour effect (as described in section B.2.3 of the company submission)."[4] Information in Section B.2.3 of the CS indicates that non-study drugs included anti-emetics, growth factors, blood products, transfusions, antibiotics, pain medications, bisphosphonates, and replacement hormonal therapies (insulin, thyroid hormones, oestrogen/progesterone).[5] The EAG considers that there is remaining uncertainty as to whether the two deaths could have been linked to the study drug and also whether contamination could have occurred from the use of non-study drugs.

3.2.8.5 Serious adverse events

The company reported that: "[redacted] *reported 1 or more SAEs….The most frequently reported SAE were* [redacted]"[5]

"[redacted] *reported an SAE that was considered related to Belzutifan by the investigator.* [redacted] *reported in 1 participant each*".[5]

Table 3.30 shows the number of patients with SAEs in order of decreasing incidence and also indicates those considered to be drug related.

Table 3.30: MK-6482-004 patients with SAEs listed by decreasing incidence (incidence >0%) and drug-related SAEs (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n (%)	
Date of Data Cut-off: 01-APR-2022	**Any SAEs listed by decreasing incidence**	**Drug-related SAEs**
Subjects in population	[redacted]	[redacted]
with one or more Aes	[redacted]	[redacted]
with no AE	[redacted]	[redacted]
Embolism	[redacted]	[redacted]
Haemorrhage intracranial	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Anaphylactic reaction	[redacted]	[redacted]
COVID-19	[redacted]	[redacted]
COVID-19 pneumonia	[redacted]	[redacted]
Cholecystectomy	[redacted]	[redacted]
Coronary artery dissection	[redacted]	[redacted]
Cystitis	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n (%)	
Date of Data Cut-off: 01-APR-2022	**Any SAEs listed by decreasing incidence**	**Drug-related SAEs**
Hypotension	[redacted]	[redacted]
Hypoxia	[redacted]	[redacted]
Non-small cell lung cancer	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Retinal detachment	[redacted]	[redacted]
Retinal vein occlusion	[redacted]	[redacted]
Seizure	[redacted]	[redacted]
Skin laceration	[redacted]	[redacted]
Suicide attempt	[redacted]	[redacted]
Toxicity to various agents	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]
Vitreous haemorrhage	[redacted]	[redacted]

Based on Tables 117 and 118 of the CS.[5]
The following footnotes were included in the CS:[5]

- Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
- Date of Data Cut-off: 01APR2022

Aes = adverse events; COVID-19 = Coronavirus Disease of 2019; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event

3.2.8.6 Discontinuation of study treatment due to AEs

The following details were outlined in the CS: "*A total of* [redacted] *discontinued Belzutifan due to the following AEs, which occurred in* [redacted] *each:* [redacted]."[5] Further details are presented in Table 3.31.

Table 3.31: MK-6482-004 patients with AEs leading to study drug discontinuation; and drug-related AEs leading to study drug discontinuation

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n(%)	
Date of Data Cut-off: 01-APR-2022	**Any AEs**	**Drug-related AEs**
Subjects in population	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
• Dizziness	[redacted]	[redacted]
• Haemorrhage intracranial	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
• Toxicity to various agents	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]
• Suicide attempt	[redacted]	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan (N=61) n(%)	
Date of Data Cut-off: 01-APR-2022	**Any AEs**	**Drug-related AEs**
Based on Table 119 of the CS[5] and Table 14.3-33 of the CSR.[24] The following footnotes were included in the CS:[5] • Adverse events up to 30 days of last dose are included. Subject with two or more adverse events in the same system organ class (or with the same preferred term) is counted only once for that system organ class (or preferred term). Adverse Events were coded using MedDRA version 25.0. Uncoded preferred terms are presented in their verbatim terms. • Date of Data Cut-off: 01-APR-2022 AE = adverse event; CS = company submission; CSR = clinical study report; MedDRA = Medical Dictionary for Regulatory Activities.		

3.2.8.7 Interruption to study treatment due to AEs

The company reported the following details: "*A total of* [redacted] *experienced 1 or more AEs leading to interruption of Belzutifan…The most frequently reported AEs leading to interruption of Belzutifan included* [redacted]."[5]

[redacted] *reported AEs leading to interruption of Belzutifan that were considered drug-related by the investigator including* [redacted]."[5]

Table 3.32 shows the number of patients with AEs resulting in treatment interruption listed by decreasing incidence and also indicates those considered to be drug related.

Table 3.32: MK-6482-004 patients with AEs resulting in treatment interruption listed by decreasing incidence (incidence >0%) and same considered as drug-related (safety analysis set)

MedDRA System Organ Class Preferred Term	Belzutifan (%)	
Date of Data Cut-off: 01-APR-2022	**Any Aes listed by decreasing incidence**	**Drug-related Aes**
Subjects in population	[redacted]	[redacted]
with one or more adverse events	[redacted]	[redacted]
with no adverse event	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]
Influenza like illness	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
COVID-19	[redacted]	[redacted]
Haemorrhage intracranial	[redacted]	[redacted]
Syncope	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]

MedDRA System Organ Class Preferred Term	Belzutifan (%)	
Date of Data Cut-off: 01-APR-2022	**Any Aes listed by decreasing incidence**	**Drug-related Aes**
Arthralgia	[redacted]	[redacted]
COVID-19 pneumonia	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]
Cholecystectomy	[redacted]	[redacted]
Cystitis	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Dyspepsia	[redacted]	[redacted]
Embolism	[redacted]	[redacted]
Hypersensitivity	[redacted]	[redacted]
Nasal congestion	[redacted]	[redacted]
Pericardial effusion	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Retinal detachment	[redacted]	[redacted]
Retinal vein occlusion	[redacted]	[redacted]
Sensation of foreign body	[redacted]	[redacted]
Skin laceration	[redacted]	[redacted]
Tremor	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
Upper-airway cough syndrome	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]
Vertigo	[redacted]	[redacted]
Viral infection	[redacted]	[redacted]
Vision blurred	[redacted]	[redacted]

Based on Tables 120 and 121 of the CS.[5]
The following footnotes were included in the CS:[5]

- Every subject is counted a single time for each applicable row and column. A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding. Adverse events up to 30 days of last dose are included.
- Date of Data Cut-off: 01APR2022.

AE = adverse event; COVID-19 = Coronavirus Disease of 2019; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities

3.2.9 Ongoing studies

The company described the following ongoing evaluations in Section B.2.11 of the CS:

"*There is currently a phase 2 single arm study to evaluate the efficacy and safety of Belzutifan monotherapy in participants with advanced pheochromocytoma/paraganglioma, pNET or VHL disease-associated tumours (the MK-6482-015 study),[25] the primary objective of the study is to evaluate the ORR associated with treatment with Belzutifan, the estimated primary completion date for this study is 12-AUG-2026.*"[5]

"*Additionally, a condition of the MHRA marketing authorisation for Belzutifan in this indication is for MSD to set up and report on results from a prospective patient registry with the objective to further characterise efficacy and understand long term safety of Belzutifan, particularly in VHL-associated RCC and CNS hemangioblastomas. The protocol for this prospective patient registry is currently being assessed by the MHRA.*"[5]

EAG comment: None.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

One comparator study was described in the clinical effectiveness section of the CS, namely the VHL Natural History study.[5, 26] However, other sources of comparator data were used to inform the CE model: retrospective analysis of the pre-treatment phase of the MK-6482-004 study and the Optum Clinformatics Data Mart claims study. The company reported a MAIC based on a comparison of outcomes for Belzutifan versus SoC provided by the MK-6482-004 and VHL Natural History studies (the MAIC is discussed in detail in Section 3.4).[5]

3.3.1 The VHL Natural History study

In their response to clarification question A15b, the company explained that "*…the VHL Natural History Study is a (currently) unpublished study specifically commissioned by MSD to address the lack of available relevant comparator data in the published literature.*"[4] The overarching aim of the study is to increase understanding of the natural history of VHL-associated disease and the primary objective is to describe the linear growth rate (LGR) of renal solid tumours among patients with VHL disease. The research design is described as a retrospective and non-interventional study of existing medical records.[5, 26]

The data source was the United States (US) National Cancer Institute's (NCI's) Urological Oncology Branch (UOB) Hereditary Database which comprised data from patients with VHL disease receiving care at the US National Institutes of Health (NIH) Clinical Center in Bethesda, Maryland from 31 July 2004 to 30 June 2020. Patients with at least one VHL-associated RCC tumour measured during the study period and who met other eligibility criteria of the VHL Natural History study (e.g., treatment with Belzutifan and history of metastatic disease were both exclusion criteria) were identified and followed up until the date of death or most recent clinical encounter within the aforementioned assessment period.[5, 26] Further details of participant selection criteria for the VHL Natural History study are presented on pages 99 to 100 of the CS.[5] A total of 308 patients met the VHL Natural History study participant eligibility criteria. After application of additional criteria to match those of the MK-6482-004 study, a subgroup of 247 patients was identified.[5, 26]

The CS[5] did not provide details of the outcomes for the VHL Natural History study, however these were listed in the Merck Sharp & Dohme (MSD) data on file document as: RCC LGR (primary outcome); and frequency and type of tumour reduction procedures (including surgical and other procedures), time to first tumour reduction procedure (all procedures), time to first surgery, renal function (estimated glomerular filtration rate [eGFR]) and other laboratory measurements such as detection of anaemia (all secondary outcomes).[26] Patients were followed up from the patient-level index date (defined as the earliest date that a measurable RCC was detected during the study period) until either the date of death or last clinical encounter during the study period.[5]

In the CS, the company provided a series of tables showing baseline characteristics for the VHL Natural History Study (based on n=247 patients matched to MK-6482-004) alongside those for the MK-6482-004 patients for the following VHL disease-associated cohorts: all patients (all of whom had RCC);

RCC and CNS Hb; and RCC and pNETs. The equivalent of 'baseline' in the VHL Natural History Study was the "*patient-level index date", defined as the earliest date that a measurable renal solid tumour was detected during the study period (July 31, 2004 to June 30, 2020)*".[5] The baseline characteristics are discussed further in Section 3.4.

The methods and results of the MAIC are discussed in Section 3.4. Results for the VHL Natural History study are shown in the MSD data on file document (cited previously).[26]

3.3.2 The pre-treatment phase of MK-6482-004

Data collected retrospectively from the pre-treatment period of the MK-6482-004 study were used to represent outcomes in the SoC arm. Whilst some information was provided on how these data were used to estimate parameters for the economic model, no details of clinical effectiveness estimates were presented in the CS.[5] This prompted the EAG to request further information on study design, baseline characteristics, treatment description and outcomes (clarification question A.16). As part of their response, the company outlined the rationale for using the MK-6482-004 pre-treatment phase as well as types of parameters that were estimated from these data:[4, 27]

"*…..it was not feasible to identify whether patients in the VHL Natural History Study had CNS Hb or pNET tumours prevalent on the patient-level index date; thus, the criteria for identifying the CNS Hb and pNET subgroups of the MK-6482-004 trial could not be well-replicated within the VHL Natural History Study.*

Due to this limitation, the following model parameters were estimated for the CNS Hb and pNET target populations using retrospectively collected data from the period before Belzutifan initiation among patients in MK-6482-004:

- *Parameter estimates for transitions from pre-surgery → surgery under active surveillance*
- *Note: As described elsewhere, the surgery state refers specifically to tumour reduction surgeries for the primary tumour type, i.e., CNS surgery when modelling the CNS cohort or pancreatic surgery when modelling the pNET cohort.*
- *Incidence rate (events/person-week) of surgeries for non-primary VHL-related tumours under active surveillance*
- *Distribution of surgeries for non-primary VHL-related tumours by specific tumour type*"[4, 27]

Note that all most comparator patients received immediate surgery (see Section 2.3), which is why these are associated with active surveillance. Clinical effectiveness estimates from the pre-treatment phase of MK-6482-004 were not provided as part of the company's response to clarification questions.

3.3.3 The Optum Clinformatics Data Mart Claims Study

The company stated that the Optum Clinformatics Data Mart Claims Study was used to align effectiveness data sourced from the VHL Natural History Study with real world UK SoC for the purposes of the CEA.[5] Again, no estimates relating to clinical effectiveness were provided in the CS[5] or in the company's response to clarification (clarification question A16).[4, 27]

EAG comment:

It is unclear why the VHL Natural History Study was used as the only source of comparator data in the clinical effectiveness section, given the use by the company of two other datasets: the pre-treatment phase of MK-6482-004 to inform rates of pre-surgery-> surgery; and the Optum Clinformatics Data Mart Claims Study, which was used in the CEA to adjust these rates on the basis that this dataset would align better with UK clinical practice than the VHL Natural History Study, which was US-based. Note that in the clarification letter response (to question A.32) the company indicated the superiority of the

former for comparison with Belzutifan: *"…the patient population of the retrospective analysis of the pre-treatment phase of the MK-6482-004 is necessarily composed of the same patients as the MK-6482-004 study' post-treatment-initiation phase and so no matching or adjustment is required."*[4] This would seem to be a reason to extend the use of the comparison of pre- to post- treatment period to estimate all outcomes including TTS. The reason given (in the company's response to clarification question A.32) for not using the Optum Clinformatics Data Mart Claims Study for the indirect treatment comparison (ITC) was *"significant limitations…the limitations include the more limited availability of matching variables (e.g., baseline renal tumour size would not be available, and the ability to measure the number of prior surgeries would be dependent on each patient's length of continuous enrolment in the database)".*[4] However, the EAG would argue that this should not rule out use of the Optum Clinformatics Data Mart Claims Study, but suggests a trade-off given the company's argument of its greater applicability to UK clinical practice.

The company did present a scenario of the cost-effectiveness analysis, which substituted the pre-treatment phase of MK-6482-004 for the VHL Natural History Study to estimate pre-surgery->surgery in the RCC model. Therefore, it is clear that whilst those alternative data sources have limitations, there are clearly significant limitations to the VHL Natural History Study and so a comparison of all three data sources should have been presented as part of the clinical effectiveness evidence with the potential for all three assessed for an ITC of all relevant time to event outcomes. The lack of this evidence is therefore a key issue.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

The company included an ITC in the form of a MAIC, which was motivated by the following: *"For the purposes of the cost-effectiveness analyses described in section B.3, it is necessary to compare the outcomes of treatment with Belzutifan with the outcomes observed in the standard of care for this patient population."* (p.98)[5]

3.4.1 MAIC methods

The company stated that the VHL Natural History Study was used as source of evidence to inform the comparator arm, adjusted using a propensity score weighting-based MAIC in a comparison with Belzutifan, informed by the MK-6482-004 study. The purpose of this was described as being to inform the CEA. For the VHL RCC '*indication'*, the VHL Natural History Study data were first subjected to a set of eligibility criteria that the company stated closely matched the MK-6482-004 study (Table 3.33).

Table 3.33: Sample selection process: Trial Population Subgroup

Step #	Criterion	N Patients
1	INCLUSION: Patients with VHL syndrome who are residents of the US or Canada	776
2	INCLUSION: Patients with ≥1 renal solid tumor identified and measured during the study period (31 July 2004 to 31 June 2020)	313
3	INCLUSION: Patients with a diagnosis of Von Hippel-Lindau (VHL) syndrome based on germline VHL alteration*	297
4	EXCLUSION: Patients with any renal procedure in the 30 days on or prior to Patient-level index date	296
5	EXCLUSION: Patients whose follow-up date was on or prior to Patient-level index date	296

Step #	Criterion	N Patients
6	EXCLUSION: If the largest tumor at patient-level index date is ≥30 millimeters (mm), patients with a renal surgical procedure with therapeutic intent performed within 60 days on or after patient-level index date	278
7	EXCLUSION: Patients who received treatment with MK-6482 or another hypoxia inducible factor 2 alpha (HIF-2α) inhibitor any time prior to Patient-level index date*	278
8	EXCLUSION: Patients who received systemic oncologic or investigational therapy any time prior to Patient-level index date**	272
9	EXCLUSION: Patients with evidence of VHL disease-associated metastatic disease prior to Patient-level index date*	260
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL RCC cohort**	**260**
10.a	INCLUSION: Patients at step #9 with ≥1 concomitant CNS hemangioblastoma on or before patient-level index date	228
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL CNS Hb cohort *(subset of the 260 patients in the VHL RCC Natural History Study sample)***	**228**
10.b	INCLUSION: Patients at step #9 with ≥1 concomitant pNET on or before patient-level index date	94
-	**VHL Natural History Study sample used to estimate key cost-effectiveness model inputs for the VHL pNET cohort *(subset of the 260 patients in the VHL RCC Natural History Study sample)***	**94**

Based on Table 29, CS [30 in first CS]. Bold font is as it appears in the CS.[5]
*These criteria are stated to be ones that were applied to better match the Belzutifan trial, the others being those for the whole of the VHL Natural History Study.[5]
**This criterion is not listed as having been applied for matching in Appendix O, but it is expressed differently i.e.: [5]
"Patients who received systemic oncologic or investigational therapy within 30 days on or prior to Patient-level index date"
CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; HIF-2α = hypoxia inducible factor 2 alpha; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

The company stated that, for the VHL CNS Hb and VHL pNET indications, patients who met the above inclusion/exclusion criteria were further restricted to those with a recorded history of CNS hemangioblastoma and pNET, respectively. However, the company further stated that: *"As a limitation, it was not feasible using the available Natural History Study data to identify whether patients in these subsets had CNS hemangioblastoma and pNET at the patient-level index date (i.e., it was only feasible to identify patients with a recorded history of CNS hemangioblastoma or pNET at some point prior to the patient-level index date)."* (p.107 of the CS).[5] The company stated that because of this, TTS for the VHL-CNS and VHL pNET CEAs was estimated for SoC for those patients who assumed not to receive surgery immediately (see Section 2) using the pre-treatment period of the MK-6482-004 study. In fact, the only outcomes that were reported in the clinical effectiveness Section (B.2)[5] from the VHL Natural History Study were from the RCC cohort, and only the cause specific hazard of pre-surgery->1st surgery (appears to be otherwise known as TTS) and the "*incidence of non-RCC VHL-related surgeries with therapeutic intent*" (see Section 3.4.2). This is despite adjustment having been made according to the CS to all three cohorts according to the baseline characteristics not only the RCC cohort, but also the CNS and the pNET cohorts (see Table 3.34).

Table 3.34: Baseline characteristics of the VHL Natural History Study and MK-6482-004 trial populations before and after reweighting – VHL RCC cohort

Baseline characteristics	VHL Natural History Study				MK-6482-004	
	Before reweighting		After reweighting (effective) [a]			
RCC cohort						
Sample size	260		92.2		61	
Mean Age at patient-level index date (years)	42.1		41.0		41.0	
Standard deviation	12.3		13.5		13.5	
Female	120	46.2%	43.8	47.5%	29	47.5%
Mean Number of renal surgeries with therapeutic intent prior to patient-level index date[b]	1.4		2.4		2.4	
Standard deviation	1.5		1.6		1.6	
Mean Tumour size of the largest renal solid tumour at patient-level index date (cm)	2.1		2.5		2.5	
Standard deviation	1.0		0.9		0.9	
CNS cohort						
Sample size	228		37.9		61	
Mean Age at patient-level index date (years)	42.3		40.4		40.4	
Standard deviation	11.7		12.8		12.8	
Female	102	44.7%	15.1	40.0%	20	40.0%
Mean Number of CNS surgeries with therapeutic intent prior to patient-level index date	1.0		2.8		2.8	
Standard deviation	1.3		2.6		2.6	
pNET cohort						
Sample size						
Mean Age at patient-level index date (years)	45.5		42.7		42.7	
Standard deviation	11.7		15.1		15.0	
Female	54	57.5%	32.9	54.5%	12	54.5%
Mean Number of pancreatic surgeries with therapeutic intent prior to patient-level index date	0.3		0.2		0.2	
Standard deviation	0.6		0.5		0.5	

Based on Tables 34-36, CS.[5]

[a] Effective sample size is computed as the square of the summed weights divided by the sum of the squared weights.

[b] This "Number of renal surgeries with therapeutic intent prior to patient-level index date" variable and its definition/criteria is not the same as (it is more restrictive than) that of the "Number of Prior Surgeries per Subject" variable shown in Table 3.6.

CNS = central nervous system; CS = company submission; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

The characteristics were reported to have been chosen *"after eliciting input from clinical experts that are prognostic of transition probabilities starting from the pre-surgery state, or that may modify the effect of Belzutifan on these transition probabilities"* (p.111)[5] Other characteristics identified as potential relevant, but excluded due to *"data limitations"*[5] were: VHL type and VHL gene alteration type (excluded due to a high proportion of missing values in the Natural History Study); number of concomitant measured tumours (unavailable at the time of analysis); and size of the largest CNS tumour at the patient-level index date in the VHL CNS Hb cohort or size of the largest pancreatic tumour in the VHL pNET cohort (the presence/absence of CNS Hb and pNET at the patient-level index date could not be identified using the available data).

In the CE Section (B.3),[5] the VHL Natural History Study is reported to be the source for the cause-specific hazards for two other outcomes, pre-surgery->metastatic disease and pre-surgery->death for all three cohorts, RCC, CNS Hb and pNET (see Table 46, CS).[5] However, it appears that the propensity score weighting was only applied to the estimation of TTS for SoC given that it was only explicitly stated to be the method for this outcome.

3.4.2 MAIC results

As stated above, the only MAIC was based on the VHL Natural History Study, relating to the whole RCC cohort, and only the cause specific hazard of pre-surgery->1st surgery (appears to be otherwise known as TTS) and the 'incidence of non-RCC VHL-related surgeries with therapeutic intent' were estimated, the results for which are shown in Table 3.35.

Table 3.35: Reweighted VHL Natural History Study RCC cohort and the MK-6482-004 trial population outcomes

Outcomes	VHL Natural History Study	MK-6482-004
	After matching (effective N=92.2)	(N=61)
Exponential rate parameter for the cause-specific hazards of pre-surgery → 1st surgery		
Rate (events/person-year)	0.25324	0.03692
Standard error	(0.01768)	(0.0156)
Incidence of non-RCC VHL-related surgeries with therapeutic intent (events/person-year)		
Number of VHL-related surgeries	2116.4	208
Total person-years at risk	227.35	194.41
Incidence rate (events/person-week*)	0.178984	0.02119

Based on Table 37, CS.[5]
*This was reported as person-year, but the EAG have corrected this based on their own calculations.
EAG = Evidence Assessment Group; CS = company submission; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

EAG comment:

The EAG would first point out that the purpose of an ITC of any kind is to estimate the treatment effect i.e., the effectiveness of the intervention, in this case Belzutifan, versus the relevant comparator, which should be SoC. This might be instrumental to performing the CEA, but it is also necessary in itself for decision making, as stated in the NICE health technology evaluations manual:[28] *"Evaluating effectiveness needs quantification of the effect of the technology under evaluation and of the relevant comparators on appropriate outcome measures."* (p.46).

The EAG requested clarification as to whether it was known whether a patient had pNET or CNS Hb, to which the company replied that it had to be inferred only from a history of ever had this kind of tumour prior to the patient-level index date.[4] However, it appears that the MAIC was only used for the RCC cohort where identification of pNET or CNS at the index date would not be required.

The EAG requested clarification as to precisely which data (only 1st or whether 2nd or 3rd surgeries) were used to inform the exponential rate parameter for the cause-specific hazards of pre-surgery → 1st surgery (TTS), to which the company replied that it was only 1st surgeries.[4]

The EAG requested clarification as to the precise methodology of the so-called MAIC, given that it was the comparator data that appear to have been adjusted, which would only be done if only summary statistics as opposed to individual patient data (IPD) were available for the intervention, which cannot be the case given that the company own the Belzutifan trial.[29] What would make more sense is that the analysis was with IPD from both the VHL Natural History Study and MK-6482-004, but with propensity score weights applied only to the former in order to estimate the average treatment effect of the treated (ATT).[30] The EAG requested that adjustment using IPD be carried out following the methodology described in Technical Support Document (TSD) 17.[30] In response to clarification, the company stated that a third party, IQVIA, conducted the MAIC, with access to the IPD of the VHL Natural History Study, but only summary statistics from MK-6482-004, which was their explanation as to why an IPD method of adjustment was not feasible. However, this does not explain why MSD did not supply IQVIA with IPD necessary to perform IPD based analyses. The company did provide references for the performance of a MAIC, but did not cite either TSD 17, as requested by the EAG, which clearly recommends an IPD based method if IPD are available.[30] They also did not cite TSD 18 in describing the MAIC, nor follow its reporting recommendations, including:[29]

- Evidence of effect modifier status.
- Distribution of weights to show the degree to which there is lack of overlap, as indicated by extreme weights.
- Estimates of systematic error before and after population adjustment.

The EAG also requested clarification as to the differences in list of baseline characteristics between tumour type cohorts and why the number of all surgery types was not included, to which the company replied that pNET and CNS tumour size was not available.[4] They also stated that all surgery types was not included because of the reduction in effective sample size (ESS) and the *"…expectation that number of primary tumour surgeries was the most pertinent covariate in each population…"*.[4] However, as recommended in TSD 18, all potential prognostic variables and treatment effect modifiers should be considered for inclusion: *"For an unanchored (using two sets of single arm data, as the CS MAIC does) indirect comparison, both propensity score weighting and outcome regression methods should adjust for all effect modifiers and prognostic variables, in order to reliably predict absolute outcomes."*[29]

In conclusion, the EAG considers that the ITC performed by the company is limited in that the company have not provided adequate justification for the following aspects:

- Method of adjustment for confounding, which could have included the use of IPD from both the MK-6482-004 trial and the VHL Natural History Study.
- Choice of confounding characteristics, prognostic or treatment effect modifying, for which no objective evidence was provided.
- Choice of outcomes for which adjustment has been performed, which could include cause-specific hazards for pre-surgery->metastatic disease and pre-surgery->death for all three cohorts.

These limitations in the ITC hinder the assessment of the effectiveness of Belzutifan compared to SoC and thus constitute a key issue.

3.5 Additional work on clinical effectiveness undertaken by the EAG

The EAG did not undertake additional clinical effectiveness work in relation to this submission.

3.6 Conclusions of the clinical effectiveness section

3.6.1 Clinical effectiveness review methods

The CS[5] and response to clarification[4] provided sufficient details for the EAG to appraise the literature searches conducted to identify relevant clinical evidence for the efficacy and safety of Belzutifan for treating tumours associated with VHL disease. Searches were conducted in June 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases and trials registers were searched. Overall, the EAG has no major concerns about the literature searches conducted.

The eligible population for the SLR is broadly in line with that described in the NICE Final Scope and the DP. However, mismatches in terms of interventions, comparators and outcomes may mean that the clinical effectiveness SLR does not fully address either the NICE Final Scope or the DP. There is potential for study selection bias in relation to studies providing comparator data. The impact of limiting to English language publications only is unclear and the possibility of language bias cannot be excluded.

The study selection process was satisfactory and in line with recommended good practice in systematic reviews, based on the company's response to the clarification letter. However, the process applied to data extraction is unclear and there is a lack of transparency relating to actual data extracted from intervention and comparator studies.

The company described inappropriate quality assessment tools both in the CS and in their response to clarification. Only the intervention study (MK-6482-004) was assessed, and not any of the sources of comparator data. The review process for assessing RoB was unclear. The EAG conducted its own RoB assessment of the one included study (MK-6482-004).

A pairwise meta-analysis was not feasible, and an ITC was performed using data from the MK-6482-004 and VHL Natural History studies.

3.6.2 Clinical effectiveness review results

One single-arm study was identified (MK-6482-004) that assessed the clinical effectiveness of Belzutifan in adults with VHL disease, who have at least one measurable RCC tumour. Reported outcomes included ORR, TTR, PFS, TTS and AEs. MK-6482-004 was conducted at 11 centres in the US, Denmark, France, and the UK. One patient received treatment in the UK. In light of the single-arm study design, MK-6482-004 is deemed to be at high RoB because of the potential for confounding of the treatment effect. The median duration of follow-up among the 61 participants with RCC in the safety analysis set at the 1 April 2022 data cut-off date was 37.7 months (range: 4.2 to 46.1 months). The confirmed ORR among the 61 participants with RCC in the Efficacy Analysis Set was 63.9% (95% CI: 50.6, 75.8). The median DOR was not reached. The median TTR was 11.1 months (range: 2.7 to 30.5 months) among 39 participants with a confirmed BOR of complete response (CR) or PR and the median (95% CI) PFS was [redacted] months. The median TTS could not be evaluated. No data on OS (an outcome specified in the NICE Final Scope) were provided. Data on subgroups defined according to tumour type or combination of different tumour types were limited.

Regarding AEs, there is remaining uncertainty as to whether the two recorded deaths could have been linked to the study drug and also whether contamination could have occurred from the use of drugs other than Belzutifan during the study period.

3.6.3 Comparator data and indirect treatment comparison

The VHL Natural History Study was used as the only source of comparator data in the clinical effectiveness section. It was unclear why this was the only study considered, given the use by the company of two other datasets: the pre-treatment phase of MK-6482-004 to inform rates of pre-surgery-> surgery; and the Optum Clinformatics Data Mart Claims Study, which was used in the CEA to adjust these rates on the basis that this dataset would align better with UK clinical practice than the VHL Natural History Study, which was US-based. In their response to the clarification letter, the company indicated the superiority of the pre-treatment phase of MK-6482-004 as a source of comparator data and it is unclear why this data source was not used to estimate all TTS outcomes with full reporting of the evidence in the clinical effectiveness section. It is also unclear why the Optum Clinformatics Data Mart Claims Study was not used in the ITC given its potentially greater applicability to UK clinical practice. The EAG noted limitations to all three studies in terms of using them for comparator data versus the data from MK-6482-004. .This is why the EAG have identified the source of comparator evidence as a key issue and would argue for the use of all three sources and not only the VHL Natural History Study as sources of comparator data and potentially for an ITC.

An ITC was performed using data from the MK-6482-004 and VHL Natural History studies. The EAG considers that the ITC performed by the company is limited in that the company have not provided adequate justification for the following aspects:

- Method of adjustment for confounding, which could have included the use of IPD from both the MK-6482-004 trial and the VHL Natural History Study.
- Choice of confounding characteristics, prognostic or treatment effect modifying, for which no objective evidence was provided.
- Choice of outcomes for which adjustment has been performed, which could include cause-specific hazards for pre-surgery->metastatic disease and pre-surgery->death for all three cohorts.

These limitations in the ITC hinder the assessment of the effectiveness of Belzutifan compared to SoC and thus constitute a key issue.

3.6.4 Summary

The EAG noted misalignments between clinical effectiveness data from the MK-6482-004 study and the DP in terms of the population and outcomes. The MK-6482-004 study population is narrower than that of the DP in terms of tumour type (must have ≥1 RCC as opposed to must have ≥1 among RCC or CNS Hb or pNET respectively). In addition, it is likely that at least some patients recruited to the MK-6482-004 study had less severe disease compared to those in the DP population. This presents challenges in generalising the findings of the MK-6482-004 study to the target UK population. It is unclear to what extent the distribution of such subgroups in the MK-6482-004 study correspond to the distribution in the UK target population. In terms of the comparator evidence, there are significant limitations with how it was obtained, the choice of only the VHL Natural History Study and the methods of the ITC to compare Belzutifan with SoC using this study and MK-6482-004.

4. COST EFFECTIVENESS

4.1 EAG comment on company's review of CE evidence

This Section pertains mainly to the review of CEA studies. However, the search Section 4.1.1 also contains summaries and critiques of other searches related to CE presented in the CS. Therefore, the following Section includes searches for the CEA review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

4.1.1 Searches performed for CE section

The following Sections include a summary and critique of literature searches undertaken "*to identify published cost-effectiveness studies for Belzutifan or other VHL therapies.*" The company further explained that: "*The cost-effectiveness SLR was designed and executed in line with NICE guidance and was run as part of a broader SLR designed to identify (i) RCTs and non-RCTs, (ii) utility data, and (iii) cost and resource use data.*"[5]

Searches for CEA review

The following paragraphs contain summaries and critiques of all searches related to CE and resource identification presented in the CS.[5] The CADTH evidence-based checklist for the PRESS, was used to inform this critique.[9, 10] The CS[5] was checked against the STA specification for company/sponsor submission of evidence.[11] The EAG has presented only the major limitations of each search strategy in the report.

Appendix G of the CS provides details of an SLR conducted to identify previous health economic evaluations and cost and resource use studies to inform a CEA of Belzutifan in the UK setting.[5] Searches were undertaken in July 2020 and update searches were conducted in July 2022.

A summary of the sources searched is provided in Table 4.1.

Table 4.1: Data sources searched for economic evaluations/cost resource identification (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
Embase	Embase.com	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
Medline In-Process	PubMed	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
NHS EED DARE HTA	Internet	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
Conferences			
ISPOR/ISPOR-EU and International	Internet	2016-2018	Not stated
ASCO			
AACR			
ECC/ESMO			
NCCN			
ASCO (GU)			

Resource	Host/Source	Date Ranges	Date searched
AMCP			
AMCP Nexus			
Based on details in Section G.1.2 of Appendix G of the CS.[5] AACR = American Association for Cancer Research; AMCP = Academy of Managed Care Pharmacy (Annual Meeting); ASCO = American Society of Clinical Oncology; ASCO GU = ASCO Genitourinary Cancers Symposium; CS = company submission; DARE = Database of Abstracts of Reviews of Effects; ECC/ESMO = European Cancer Congress /European Society for Medical Oncology; HTA = Health Technology Assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research (Europe and International); NCCN = National Comprehensive Cancer Network; NHS EED = National Health Service Economic Evaluation Database			

EAG comment:

- A good range of databases, websites, grey literature resources and trials registers were searched. Reference checking was conducted.
- Searches were well structured, transparent, and reproducible. A good range of subject indexing terms (MeSH/EMTREE) and free text were used.
- The EAG noted that the PubMed search strategy (Table 124; Appendix G) appeared to be a search of all PubMed records, rather than just MEDLINE In-Process as stated in the submission (clarification question A.4). The company responded to this question with the following explanation:

  "*This search was not restricted to in process/ahead of print results hence, no "AND (inprocess[sb] OR pubstatusaheadofprint)" string was used however, please note the search results of this search strategy still included in process and ahead of print records. To clarify, PubMed was searched in addition to Embase to additionally identify 'in process' records; however, the search in PubMed was not restricted.*"[4]

  The EAG was satisfied with this response, as this will have increased the yield of records from this database.
- The EAG noted that the company's economic searches reported a joint search of MEDLINE and Embase via Ovid.com. The company confirmed Embase was searched on the understanding that it contains all MEDLINE content. Whilst the company stated that Embase's mapping of MEDLINE records to Embase's own Emtree terms removed the necessity of searching MEDLINE as a separate search it is unclear if this is the case for all potentially useful MeSH terms. A separate search allows the searches to fully utilise the power of database specific study design filters developed to make the most of an individual database's subject headings, for these reasons the EAG considers it preferable to conduct a separate MEDLINE search. However, the separate PubMed search conducted for MEDLINE In-Process records should have identified any additional MEDLINE records missed by this approach.
- The EAG noted in Appendices G and H that the update searches were for a narrower population than the original searches. The original searches were for all patients with VHL, whereas the update searches only identified records where VHL terms occurred in conjunction with CNS Hb, pNETs or RCC terms. The EAG asked the company to explain the discrepancy between these two approaches, and what effect this may have had on the update search results (clarification question A.5). The company responded as follows:

  "*The original searches were conducted in 2020 before the MHRA GB marketing authorisation was granted on 31-MAY-2022 with wording specifically for patients with VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET). Therefore, the inclusion criteria for the population was updated to include these manifestations; however, as stated in Table 129 [now Table 128] of*

the CS "In absence of a clear reporting about the diagnosis method, studies that mention 'VHL disease' will also be included". Hence, the updated search in July 2022 has not restricted the population but rather more closely aligned with the GB marketing authorisation and NICE scope for this appraisal."[4]

The EAG appreciated this clarification.

- The EAG noted that the updated searches conducted for the National Health Service Economic Evaluation Database (NHS EED), Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment (HTA) databases Table 130; Appendix G) were redundant as NHS EED and DARE were last updated in 2015, and the HTA database in 2018.

Searches for model input

Appendix H of the CS provides details of an SLR conducted to identify published HRQoL utilities to inform a CEA of Belzutifan in the UK setting.[5] Searches were undertaken in July 2020 and update searches were conducted in July 2022.

A summary of the sources searched is provided in Table 4.2.

Table 4.2: Data sources searched for HRQoL studies (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
Embase	Embase.com	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
Medline In-Process	PubMed	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
CENTRAL	Cochrane Library	DB inception-1/7/2020 1/7/20-26/7/22	July 2020 July 2022
Conferences			
ISPOR/ISPOR-EU and International	Internet	2016-2018	Not stated
ASCO			
AACR			
ECC/ESMO			
NCCN			
ASCO (GU)			
AMCP			
AMCP Nexus			
AAO			
ARVO			

Based on details in Section H.1.2 of Appendix H of the CS.[5]
AACR = American Association for Cancer Research; AAO = American Academy of Ophthalmology; AMCP = Academy of Managed Care Pharmacy Annual Meeting; ARVO = Association for Research in Vision and Ophthalmology; ASCO = American Society of Clinical Oncology; ASCO (GU) = ASCO Genitourinary Cancers Symposium; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; ECC/ESMO = European Cancer Congress/European Society for Medical Oncology; EU = European Union; HRQoL = health-related quality of life; ISPOR = International Society for Pharmacoeconomics and Outcomes Research (Europe and International); NCCN = National Comprehensive Cancer Network

EAG comment:

- A good range of databases, websites, grey literature resources and trials registers were searched. Reference checking was conducted.
- Searches were well structured, transparent and reproducible, and a good range of subject indexing terms (MeSH/EMTREE) and free text was used.
- The EAG noted that the company's HRQoL searches reported a joint search of MEDLINE and Embase via Ovid.com. The company confirmed Embase was searched on the understanding that it contains all MEDLINE content. Whilst the company stated that Embase's mapping of MEDLINE records to Embase's own Emtree terms removed the necessity of searching MEDLINE as a separate search, it is unclear if this is the case for all potentially useful MeSH terms. A separate search allows the searches to fully utilise the power of database specific study design filters developed to make the most of an individual database's subject headings, for these reasons the EAG considers it preferable to conduct a separate MEDLINE search. However, the separate PubMed search conducted for MEDLINE In Process records should have identified any additional MEDLINE records missed by this approach.
- The EAG noted the study design filter used in Appendix H for the original PubMed search appears to be much narrower in scope than the filter used on other databases (clarification question A.8). For example, the filters used in line #2 of Table 141 of the PubMed search strategy are much narrower than the filters in line #2 of Table 139 of the Embase search. The company responded to this question as follows:

  "*We used only the search terms available in Pubmed, several of the terms such as the following were not found in Pubmed: "health year equivalent, "disutility", "disfigure", "123uropean organization for research and treatment of cancer questionnaire", "routine electronic monitoring of hrqol", "fact-ksi". Therefore, the search appears narrower, however given the fact that we used indexed terms and free-text searches, this search is not likely to have lesser sensitivity than the 'broader' one.*"[4]

  The EAG was satisfied with this response, as this will not have affected the results obtained from the search of this resource.

4.1.2 Inclusion/exclusion criteria

The in- and exclusion criteria used by the company are presented in Appendix G, Tables 127 (search date July 2020) and 128 (search date July 2022) for published CE studies, and in Appendix H, Tables 146 (search date July 2020) and 147 (search date July 2022) for HRQoL studies.[5] The EAG considers the in- and exclusion criteria suitable to capture all relevant evidence.

4.1.3 Findings of the CE review

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagrams for the CE studies can be found in Figures 35 (search date July 2020) and 36 (search date July 2022) of Appendix G, and for the quality of life (QoL) studies in Figures 38 (search date July 2020) and 39 (search date July 2022) of Appendix H.[5] A total of eight CE studies (six from the July 2020 search and two from the July 2022 search) and eight QoL studies (all from the July 2020 search) were included. The eight included CE studies were not suitable to assess the CE of a treatment for VHL in the UK. None of the HRQoL studies used the European Quality of Life-5 Dimensions (EQ-5D) for measuring HRQoL.

4.1.4 Conclusions of the CE review

The CS[5] and response to the clarification letter[4] provided sufficient details for the EAG to appraise the literature searches conducted to identify economic, HRQoL and cost data on Belzutifan for treating tumours associated with VHL disease. Searches were conducted in July 2020, with updates in July 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases were searched. Overall, the EAG has no major concerns about the literature searches conducted. Since no CE models to address the DP were identified by the company, a de novo model was built, which is discussed in the remainder of this Section.

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	EAG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	As per the reference case
Perspective on costs	NHS and PSS	As per the reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	As per the reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	As per the reference case
Synthesis of evidence on health effects	Based on systematic review	As per the reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Health effects expressed in QALYs. HRQoL measured using the EQ-5D-5L (mapped to EQ-5D-3L).
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	HRQoL collected from three different sources: • VHL RW QoL Disease Burden Study:[1] HRQoL reported by VHL patients: not limited to patients who require therapy and for whom localised procedures are unsuitable or undesirable and response status self-assessed by patients. • KEYNOTE-564:[3] HRQoL reported in adult patients with adjuvant pembrolizumab

Element of HTA	Reference case	EAG comment on CS
		treatment of RCC post-nephrectomy. The utility associated with CR in this trial is used for all three VHL cohorts since all patients in the MK-6482-004 trial had at least one measurable solid RCC tumour. • Kiebert et al. 2001:[2] reported utility values of patients with amyotrophic lateral sclerosis; a motor neurone disease that is considered an appropriate proxy for patients with VHL CNS Hb progressed disease.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Unclear if sample is representative for the UK population. The company was unable to provide a comparison of the trial data to the UK patient characteristics for all three subgroups due to the lack of published data on the population characteristics of patients with VHL disease in the UK. For both the VHL RW QoL Disease Burden Study and the KEYNOTE-564 trial, EQ-5D-5L scores mapped onto the UK EQ-5D-3L value set as per the NICE reference case.[31] Crosswalk and direct methods used to map the EQ-5D-3L value sets to EQ-5D-5L.[32]
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	As per the reference case. QALY weighting due to disease severity included as exploratory analyses. However, same QALY weight applied for three different patient cohorts where evidence suggests otherwise.

Element of HTA	Reference case	EAG comment on CS
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	As per the reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	As per the reference case

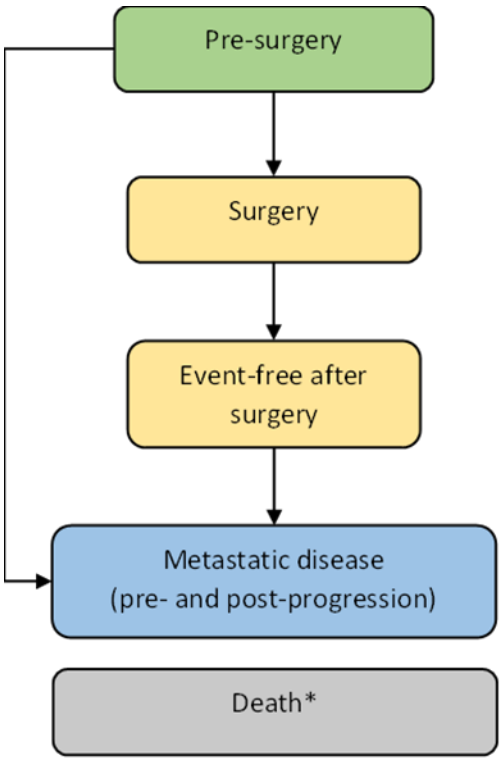
CNS Hb = central nervous system haemangioblastoma; CR = complete response; CS = company submission; EAG = Evidence Assessment Group; EQ-5D = EuroQoL-5 Dimensions; EQ-5D-3L = EuroQoL-5 Dimensions, 3 levels; EQ-5D-5L = EuroQoL-5 Dimensions, 5 levels; HTA = Health Technology Assessment; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; QoL = quality of life; RCC = renal cell carcinoma; RW = real-world; UK = United Kingdom; VHL = Von Hippel-Lindau

4.2.2 Model structure

The company developed a de novo Markov model in Microsoft Excel® to assess the CE of Belzutifan for UK adult patients with VHL-associated RCC, VHL-associated CNS Hb or VHL-associated pNET tumours who require treatment and for whom surgery is unsuitable or undesirable, compared to SoC.

The model consists of five mutually exclusive health states: pre-surgery, surgery (defined via a tunnel state), event-free after surgery, metastatic disease, and death. A schematic representation of the model structure is shown in Figure 4.1.

Figure 4.1: Schematic representation of the model structure



Based on Figure 18 of the CS.[5]
CS = company submission
*Arrows to the death health state, from each of the other health states, are omitted from the diagram for simplicity.

All patients start in the pre-surgery health state reflecting the treatment decision point of the patients and, from there, they can transition to the surgery, metastatic disease, or death health states. For the VHL RCC and VHL pNET cohorts, a tumour size threshold, combined with other clinical factors, were used to determine the need for surgery. For the VHL CNS Hb cohort, manifestation of symptoms, combined with other clinical factors, were used to determine the need for surgery. This implies that in the model some patients may be having a progressive disease (PD) while still remaining in the pre-surgery health state, and some patients may be presenting CR or PR to treatment or may be considered being in stable disease (SD), while being in the surgery and event-free after surgery states. Patients in the event-free after surgery can only transition to metastatic disease or death.

The surgery health state is a tunnel state in which patients remain for one week and it is assumed to lead to loss of organ function for VHL RCC and VHL pNET primary tumour patients, or to brain injury for VHL CNS Hb primary tumour patients. It is important to emphasise that in the model, the surgery health state refers to the surgery to the primary tumour and does not account for patients undergoing surgeries at non-primary tumour sites. The underlying assumption is that VHL patients may require surgeries at multiple sites, but clinicians would only focus on the highest risk tumour site. Therefore, as per response to clarification question A10, the company's referral to 'primary tumour' throughout the CS is used to define the tumour site that is driving treatment decisions and not the original site where it first arose.[4] Potential complications related to surgeries of non-primary tumours are captured in the model as one-time costs and decrements in QoL. Regarding the metastatic disease health state, transitions to metastatic disease are only allowed for VHL RCC and VHL pNET patients given CNS Hb tumours do not metastasise.

The model has a cycle length of one week and a half-cycle correction was applied to account for events happening at any time during the cycle. Costs and utilities are applied to each health state to calculate total costs and quality-adjusted life years (QALYs) per model cycle. The input values of the model, and their underlying assumptions, are further elaborated in the remainder of Section 4 of the EAG report.

EAG comment:

The EAG acknowledges that VHL is a complex disease to model. Despite this complexity, the current model structure seems to capture important outcomes in the disease progression of VHL, such as surgery and metastatic disease. The main concerns of the EAG relate to the model's ability to properly represent/capture "*those patients who require therapy and for whom localised procedures are unsuitable or undesirable*", as described in the DP. However, the EAG see this more as a data-related problem rather than a problem with the model structure itself.

The company explained (clarification question B2a) that the "*model was initially developed when the expectation for the marketing authorisation was for VHL-associated RCC only (in line with the population recruited into MK-6482-004 trial) and without restriction to those patients for whom localised procedures are unsuitable or undesirable.*"[4] The company then adapted the model by incorporating VHL-associated CNS Hb and pNET patients (with no restriction yet around localised procedures), following Belzutifan Food and Drug Administration (FDA) approval in August 2021. When the model was adapted at that time, it included 1st, 2nd or 3rd surgery states and the corresponding event-free after surgery states.

The EAG's main concern regarding the inclusion of other types of VHL relate to the assumptions the company made to accommodate them in the model. As the company mentioned in response to clarification question B2a for example, the model "*health states correspond to surgery and metastases of the primary tumour type only*".[4] However, the sources of evidence used to inform the effectiveness input parameters of the model, including the MK-6482-004 trial, do not distinguish whether the type of

tumour is primary or not. In response to clarification question A10d, the company explained that the assumption made in the model was that "p*atients with more than one tumour manifestation are therefore included in each cohort respective to their manifestation. For example, a patient with both RCC and pNET tumour manifestations provides data in two analyses: the VHL-RCC cohort in which RCC is the primary tumour (in the VHL-RCC cohort making pNET a non-primary tumour) and in the pNET cohort in which pNET is considered the primary tumour (in the VHL-pNET cohort making RCC a non-primary tumour)"*.[4] The EAG understands that in the absence of data that can be used to identify which is the primary tumour in patients with more than one tumour manifestation, (simplifying) assumptions are needed if the model has to include the population in the decision problem. However, the fact that there is overlap in the data informing input parameters for the three cohorts is problematic since, by definition, subgroups are expected to be mutually exclusive. Furthermore, in the model it is assumed that the rate of surgeries (i.e., the transitions from the pre-surgery to the surgery health state) should be informed by surgeries on the primary tumour, whereas in the data used to informed surgery rates, it is unknown whether the tumour is primary or not. This is a potential source of bias because it is unclear whether the trial results (e.g., surgery rates) are applicable to the model population. In response to clarification question B2c,[4] the company explained that because in the MK-6482-004 trial all patients had RCC, the observed incidence rates of *non-primary* tumour surgeries in the CNS Hb and pNET subgroups were adjusted (downwardly) using the proportions of patients with VHL-associated CNS Hb and pNET who do not have RCC, as reported in the VHL RW QoL Disease Burden Study.[1] Also, in both Belzutifan and SoC arms, the percentage breakdowns of *non-primary* tumour surgeries in these two cohorts were derived using pre-treatment period data from the MK-6482-004 trial, and were similarly adjusted to account for the proportions of patients who did not have RCC (i.e., the proportion of non-primary tumour surgeries attributable to RCC was decreased and the proportions attributable to other VHL-related tumours were proportionally increased). The company considered that, although the difference between the CNS Hb and pNET subgroups in the MK-6482-004 trial and in the model cohorts is a limitation in this submission, the impact on the model results for these cohorts is likely minor, given that:

- This difference would mostly impact raw incidences and percentage breakdowns of *non-primary* tumour surgeries in the CNS Hb and pNET subgroups of the MK-6482-004 trial, and these inputs have been adjusted.
- Data from the VHL RW QoL Disease Burden Study showed that the prevalence of RCC is 63.2% in the VHL CNS Hb cohort and 71.3% in the VHL pNET cohort.[1]
- The same limitation also applies to the SoC arm since both the MK-6482-004 trial (pre- and post-treatment periods) and the VHL Natural History Study were restricted to patients with VHL-related RCC at the baseline visit. Therefore, the company considers that this limitation should not differentially introduce bias for one arm compared to the other.

While the adjustments described above might address part of the bias, the EAG would like to emphasise that these are conducted on the surgery rates of *non-primary* tumours, but, as mentioned above, the incidence rates of *primary* tumour were directly sourced from the MK-6482-004 trial (as also explained in Section 4.2.6). For the CNS Hb cohort, there were only two CNS Hb surgeries observed with both performed on the same patient, while for the pNET cohort this is even more challenging since no pNET-related surgeries were observed in the MK-6482-004 trial. Therefore, surgery rates for the pNET cohort were indirectly estimated using the surgery rates in the SoC arm and the HR of Belzutifan versus SoC in the VHL RCC cohort, adjusted though for the better response rates of Belzutifan patients with respect to the pNET versus RCC tumours. Thus, the EAG considers that the impact on the model results of not being able to distinguish whether the type of tumour is primary or not in the available data is unknown,

and cannot be resolved, unless a different model structure is built to approach the DP , or another source of data is used to inform the current model.

A second issue is related to the narrow definition of the patient population in the DP by including "*those patients who require therapy and for whom localised procedures are unsuitable or undesirable*". As explained in Section 2.1, the DP population does not match the population in the MK-6482-004 trial. Since the trial did not include the restriction of localised procedures being unsuitable or undesirable, the trial population is less severe than the population in the DP. This issue follows MHRA approval of Belzutifan in May 2022, which included the restriction to patients "for whom localised procedures (e.g., surgery) are unsuitable or undesirable". In an attempt to accommodate this population, the company adapted their model by assuming that only one surgery was possible as a "last resort" intervention and that patients under SoC would receive this intervention immediately at the model start. In addition, transitions to subsequent surgeries (and event-free after subsequent surgery) were effectively removed from the model (transition probabilities equal zero). The EAG considers that the inclusion of patients for whom surgery is unsuitable or undesirable, but yet they would immediately get this "last resort" intervention under SoC only, is problematic because it represents a severely ill population for which no evidence has been presented. This is also evident throughout the CS, in which the company seem to make arbitrary assumptions and adjustments to define the model inputs that could be representative for this patient population. Thus, the EAG considers that the main limitation is that data from the MK-6482-004 trial are defining a different and less severe population than that in the decision problem, and despite that, were still used as the main source of evidence to inform inputs for patients in the decision problem. Given the severity of the disease of those patients included in the decision problem, which implies in the economic model that they are in need of immediate surgery, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial, and used in the model, underestimate the surgery rates for the population in the decision problem that would use Belzutifan in clinical practice. For instance, based on the evidence in Table 3.13, it is known that in the MK-6482-004 trial the overall response rate in the VHL RCC subgroup was 63.9%, with a median TTR of 11.1 months for patients with a response. The EAG is unclear whether these data are appropriately representing patients as severe as those included in the decision problem, who need immediate surgery. If patients need immediate surgery, the EAG wonders whether in daily practice these patients would be able to wait almost one year (in median) without surgery until a response to treatment is observed. This EAG concern is further strengthened when considering that 36.1% of the trial population did not respond at all to Belzutifan treatment. The EAG concern should in any case be clarified by clinical experts.

Also, as explained in detail in Section 4.2.8, another potential source of bias related to the model structure comes from having distinct baseline response distributions between treatment arms, which should be equal at the model start. This response distribution is used to implement HRQoL in the model.

In conclusion, the EAG acknowledges the difficulty of representing the population in the DP with the current evidence, which is mostly derived from the MK-6482-004 trial for Belzutifan. The EAG concludes that, in its current form, the company's model seems appropriate to reflect the initial MA VHL-associated RCC only and for the population recruited into MK-6482-004 trial but cannot provide reliable estimates of the CE of Belzutifan compared to SoC in the population defined in the DP.

4.2.3 Population

The population included in the economic analyses was defined by the company as adult patients (18 years or older) with VHL disease who require therapy for VHL-associated RCC, VHL-associated CNS Hb, or VHL associated pNET and for whom localised procedures are unsuitable or undesirable. The clinical evidence used to inform the Belzutifan arm in the economic model was obtained from the MK-

6482-004 trial, a phase two trial which investigated the efficacy and safety of Belzutifan in patients with VHL RCC disease.[25] The pre-treatment period of the MK-6482-004 trial combined with real-world data from the VHL Natural History Study and the Optum Clinformatics Data Mart Claims Study were used to inform input parameters in the SoC arm.[26, 33] Three cohorts of patients were therefore considered in the economic analyses. The key baseline patient characteristics included in the economic model are shown in Table 4.4. These characteristics were assumed to be the same across the three cohorts. The company indicated that the population assessed in the economic analysis is aligned to the population as specified in the marketing authorisation for Belzutifan. However, the company also noted that MHRA specified eligibility to adult patients with VHL "*who require therapy*" for VHL-associated RCC, CNS Hb, or pNET, "*and for whom localised procedures are unsuitable or undesirable*". The latter was not in the inclusion criteria of the MK-6482-004 trial, which did not require patients to be considered unsuitable or undesirable for localised procedures.

Table 4.4: Key baseline patient characteristics used in the economic model*

Patient characteristic	Mean (sd)	Source
Age (years)	41.0	MK-6482-004[25]
Weight (kg)	79.7 (23.4)	
Body surface area (m^2)	1.9 (0.3)	
Female (%)	47.5	

* Baseline characteristics were assumed to be the same for VHL RCC, VHL CNS Hb and VHL pNET cohorts.
CNS = central nervous system; Hb = hemangioblastomas; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; sd = standard deviation; VHL = Von Hippel-Lindau

EAG comment:

The main concerns of the EAG relate to the mismatch between the population in the DP and the population in the MK-6482-004 trial. The company seem to have made contradictory statements in this regard. In response to clarification questions A12b and A23, the company acknowledged that some patients in the MK-6482-004 trial may have had less severe disease relative to those in the DP/MA population.[4] Despite this, the company concluded in the CS that the population assessed in the economic analyses is in line with the population in the DP (MA) for Belzutifan.[5] The EAG does not agree with this conclusion and considers that this mismatch between the target population and the population in the MK-6482-004 study creates a baseline imbalance for treatment comparison within the evidence synthesis of the submission. For details we refer for example to the EAG comments in Sections 2.1 and 4.2.2.

To reflect the population in the DP in terms of each tumour type combination subgroup, the EAG asked the company (clarification question B1) to repeat the CEA using subgroup-specific model parameters (e.g., baseline characteristics and rates of surgery for each type of tumour).[4] The company indicated that distinguishing by different tumour type combinations was not feasible and, therefore, conducting CEA using subgroup-specific parameters was not possible. It remains unclear why this was not possible since, based on the evidence provided by the company in the CS, and in response to clarification question A10d, for example, the company mentioned that in the MK-6482-004 study "*all participants necessarily had VHL disease and RCC, 50 (82%) of these participants had RCC + CNS hemangioblastomas, 22 (36%) had pNETs, and 17 (28%) had RCC + CNS hemangioblastomas + pNETs at baseline as confirmed by independent review committee (IRC) assessment*".[4] The EAG wonders why it was not possible to derive input parameters for the CNS Hb cohort from the 50 patients in the MK-6482-004 study who had RCC + CNS, or for the pNET cohort from the 22 patients in the

MK-6482-004 study who had RCC + pNET. These input parameters would be in line with the subgroups/cohorts included in the economic model.

4.2.4 Interventions and comparators

The intervention considered in the CS was Belzutifan, administered at 120 mg orally once per day until disease progression or unacceptable toxicity, which is consistent with the anticipated licensed indication.[6] Duration of treatment with Belzutifan was informed based on the MK-6482-004 trial protocol, as explained in Section 3.2.2.

The comparator considered in the CS was SoC, defined as established clinical practice without Belzutifan.[5] For the VHL-associated RCC and VHL-associated pNET cohorts, this was assumed to be surgery resulting in loss of organ function in 90% of patients with the remaining 10% receiving symptom management. For the VHL-associated CNS Hb cohort, SoC was assumed to be surgery in 50% of patients with a risk of brain injury, while in the remaining 50% of patients for whom tumour location would not allow for operation, it was assumed that patients would undergo symptom management but with the same risk of brain injury due to tumour size or location as with surgery and, therefore, a similar impact on QoL as the serious complications from CNS surgery. For the VHL CNS Hb population, the last is essentially equivalent to assuming 100% of patients undergoing through immediate surgery in terms of health impact.

In the CS it was argued that, although in the UK there are localised procedures used, according to clinical experts, the patient population for whom 'localised procedures are unsuitable or undesirable', as per MA of Belzutifan, concerns actually for patients who have exhausted all alternative treatments and for whom localised procedures would be a 'last resort' option and likely result in loss of organ function with extremely poor outcomes.[5] Also, according to the CS, minimally invasive treatments, surgery, and radiotherapy that would preserve organ function would not be relevant comparators considering that if patients are to be eligible for these treatments they were assumed to accept them as such treatments would not lead to loss or organ function with extremely poor outcomes.[5] Treatment options for advanced or metastatic disease as defined in the NICE scope were also not considered as relevant comparators because these would only be used for metastatic patients following treatment with Belzutifan and were hence only incorporated in the economic model as subsequent treatments for metastatic disease.

EAG comment:

The main concerns of the EAG regarding the SoC definition are the following:

a) The EAG considered that the assumed percentages of patients undergoing immediate surgery or receiving symptom management treatment in the SoC arm to be arbitrary and for this reason asked the company to provide objective evidence for these assumptions (question B4d in the clarification letter).[4] The company's answer stated that the proportions of patients requiring immediate surgery were not informed by means of formal elicitation methods and suggested to estimate the impact of these parameters through scenario analyses. The scenarios provided in Table 4 of the clarification letter varied the proportion of patients undergoing immediate surgery from 90% to 80% and 100% in the VHL RCC and VHL pNET subgroups and from 100% to 90%-100% in the VHL CNS Hb subgroup.[4] Considering the lack of evidence for these assumptions the EAG considered the scenarios provided by the company too narrow to reflect the underlying uncertainty.
b) Furthermore, the EAG thinks that by allowing only patients in the SoC to undergo immediate surgery, the company might disproportionally be favouring the Belzutifan arm. In response to question B4b, the company stated that the need for immediate surgery is assumed as '*in the context of the MHRA label, patients are at a 'fork in the road' where they have run out of alternative*

treatment options yet still "require therapy". Therefore, they have a requirement for immediate surgery to treat their primary tumour of significant burden in the absence of Belzutifan as a treatment option'.[4] This statement does not preclude that patients eligible for Belzutifan are also in need of immediate surgery to treat their primary tumour, as treatment with Belzutifan does not imply an immediate treatment benefit. To further reinforce this point, the company in response to question B4c on the availability of the active surveillance for CNS Hb patients says that '*patients in the CNS Hb cohort can have active surveillance but not without experiencing significant sequelae associated with tumour burden which would otherwise be alleviated through localised procedures'*.[4] It is unclear to the EAG, why "similar" patients in the Belzutifan arm waiting for a response to Belzutifan treatment do not experience 'significant sequelae' which could otherwise be avoided through localised procedures. Moreover, this statement seems to imply that having surgery would result in an improvement with respect to not having surgery for CNS Hb patients. Therefore, patients in the Belzutifan arm, who do not get surgery, would not get this improvement until they start to respond to treatment. The company further mentioned that '*the population stipulated by the MHRA label "for whom localised procedures are unsuitable or undesirable" are patients experiencing either debilitating sequelae as a result of surgery or debilitating sequelae as a result of not undergoing needed surgery*'.[4] This suggests that the sequelae related to 'not undergoing needed surgery' would also be experienced by Belzutifan patients, until and if some response is achieved. Based on these statements from the company, the EAG is unclear whether patients waiting until they achieve response to Belzutifan treatment, might be in a worse state than patients in SoC undergoing immediate surgery. For this reason, in Question B4e the company were asked to clarify if patients in the Belzutifan arm would effectively suffer the harm entailed to not receiving immediate surgery. The company responded that '*it is logical that a patient would suffer some harm if needed surgery were not provided immediately. The harm in this case would be risk of metastatic disease due to tumour growth (for RCC and pNET) or symptomatic burden (in all cohorts but particularly in CNS Hb). Belzutifan works by shrinking tumours and therefore reducing the risk of these two types of "harm". This benefit is reflected in the economic model through the transitions within the health states as informed by the trial evidence for Belzutifan*'.[4] The EAG does not agree with this response and does not think this potential harm is completely captured in the model as patients in the trial did not need immediate surgery, meaning that their tumours were probably smaller than in the company's defined SoC arm and, therefore, the risk of metastatic disease would be smaller as well.

c) Also, allowing a higher proportion of patients in the SoC arm to undergo surgery when they are in need of immediate surgery would be expected to lead to some treatment benefit compared to not receiving surgery. However, the EAG noticed that in the economic model when reducing the proportion of patients in the SoC arm undergoing immediate surgery the total QALYs and life years gained (LYG) in the SoC arm are higher, with the results being strongly influenced by more QALYs and LYG in the pre-surgery health state. This translates to patients being in a better position by not having surgery despite the immediate need, which seems to be contradictory to current clinical practice.

Considering the above, the current model structure and the available data, the EAG is unable to change the model in a straightforward manner to account for patients requiring immediate surgery in the Belzutifan arm as assumed in SoC. However, to illustrate the impact on model outcomes of assuming only patients in SoC to be in need of immediate surgery, in Section 6.1 the EAG provided scenario analyses by changing these percentages to 0%.

4.2.5 Perspective, time horizon and discounting

The economic analysis is performed from the National Health Service (NHS) and Personal Social Services (PSS) perspective. The model has a time horizon of 59 years that is considered appropriate as a lifetime horizon, in line with the NICE reference case, given that the average age of patients at the start of treatment is 41 years. The model cycle length is one week, and a half-cycle correction is applied. Costs and QALYs were discounted at 3.5% as per the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main sources of evidence on treatment effectiveness used for the Belzutifan and SoC arms were the MK-6482-004 trial,[24] the VHL Natural History Study[26] and the Optum Clinformatics Data Mart Claims Study.[33] To estimate the relative treatment effectiveness of Belzutifan compared to SoC, the company indicated in the CS that the VHL Natural History Study, was adjusted using a propensity score weighting-based MAIC analysis, informed by the MK-6482-004 study (for further details please refer to Section 3.4).

4.2.6.1 Belzutifan arm: Transitions from non-metastatic disease health states

Transition probabilities from pre-surgery and event-free-after-surgery health states in each weekly cycle were calculated as a function of the cause-specific hazards for all three transitions from each of the states. To clarify this approach, the company explained the calculations of the cause-specific hazards using a 3-step approach in clarification question B6c.[4] In the pre-surgery health state for instance, the average cause-specific hazard within the cycle (the cause is defined as surgery, metastasis or death) was first estimated using trial data or other assumptions as explained in the following subsections. Next, the average hazard of any transition from pre-surgery was calculated as the sum of the average cause-specific hazard for all three causes within that cycle. In each cycle, the relative contribution to the overall hazard of each cause of transitioning from pre-surgery was then derived and used to estimate the probability of having a transition to one of the three health states given that a transition from the pre-surgery has occurred within the cycle. Cause-specific hazards for transitions from the surgery and pre-surgery health state are presented in Table 4.5. Further details on how transition probabilities were derived are provided in the remaining part of this Section.

Table 4.5: Cause-specific hazard rates for transitions from the pre-surgery and event-free-after-surgery states, by model subgroup and treatment arm

Subgroup/Treatment arm	Hazard rates per year				
	Pre-surgery → surgery	**Pre-surgery → metastatic disease**	**Pre-surgery → death**	**Event-free after surgery → metastatic disease**	**Event-free after surgery → death**
VHL RCC					
Belzutifan	0.03692	0.000312	0.00364	0.000468	0.00728
	MK-6482-004 trial[24]	Estimated using an HR equal to HR of pre-surgery → surgery for Belzutifan versus SoC	SoC rate, adjusted for reduced death cases attributable to VHL CNS Hb	Estimated using an HR equal to HR of pre-surgery → surgery for Belzutifan versus SoC	Assumed equal to SoC
SoC	0.25324*	0.00208	0.00624	0.00312	0.01196
	VHL Natural History Study (2021)[26]				
VHL pNET					
Belzutifan	0.000312	0.00026	0.00624	0.00026	0.00624
	Assumed equal to HR of pre-surgery → surgery of Belzutifan vs. SoC in the VHL RCC, multiplied by $(1-ORR_{pNET})/(1-ORR_{RCC})$	Estimated using an HR equal to HR of pre-surgery → surgery for Belzutifan versus SoC	SoC rate, adjusted for reduced death cases attributable to VHL CNS Hb	Assumed equal to pre-surgery → metastatic disease	Assumed equal to pre-surgery → death
SoC	0.00884*	0.00676	0.01092	0.00676	0.01092
	Pre-treatment period data from MK-6482-004 trial[24]	VHL Natural History Study (2021)[26]	VHL Natural History Study (2021)[26]	Assumed equal to pre-surgery → metastatic disease	Assumed equal to pre-surgery → death

Subgroup/Treatment arm	Hazard rates per year				
	Pre-surgery → surgery	**Pre-surgery → metastatic disease**	**Pre-surgery → death**	**Event-free after surgery → metastatic disease**	**Event-free after surgery → death**
VHL CNS Hb					
Belzutifan	0.0052	0.000156	0.00728	0.000156	0.00728
	MK-6482-004 trial[24]	Estimated using an HR equal to HR of pre-surgery → surgery for Belzutifan versus SoC	SoC rate, adjusted for reduced death cases attributable to VHL CNS Hb	Assumed equal to pre-surgery →metastatic disease	SoC rate, adjusted for reduced death cases attributable to VHL CNS Hb
SoC	0.10504*	0.00312	0.01456	0.00312	0.01456
	Pre-treatment period data from MK-6482-004 trial[24]	VHL Natural History Study (2021)[26]	VHL Natural History Study (2021)[26]	Assumed equal to pre-surgery → metastatic disease	Assumed equal to pre-surgery → death

Based on Table 45 and Table 46 of the CS.[5]

* For the pre-surgery → surgery transition in the VHL RCC and pNET cohorts this cause-specific hazard is used for the remaining 10% who do not receive immediate surgery in the SoC arm. In the VHL CNS Hb cohort all patients are assumed to have the outcomes from surgery, therefore this cause-specific hazard is only used following treatment effect waning in the Belzutifan arm.

** The EAG was unable to find this calculation in the model and to verify this assumption.

CS = company submission; CNS Hb = central nervous system haemangioblastoma; HR = hazard ratio; pNET = pancreatic neuroendocrine tumour; ORR = overall response rates; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Pre-surgery health state

Transitions from the pre-surgery health state of the Belzutifan arm, for the three different cohorts, were informed from MK-6482-004 trial data.[24]

For the VHL RCC cohort, the pre-surgery → surgery transition probability was estimated using an exponential model fitted to observed time-to-event data from baseline visit to first post-baseline renal surgery of the overall MK-6482-004 trial population with VHL-related RCC tumours at baseline (n=61). The CS states that the exponential distribution was selected based on 1) the small number of renal surgeries observed (n=7), 2) the fewer assumptions in the exponential model for underlying risk over time, which according to the company can avoid overfitting or convergence issues and have been used in previous NICE Technology Assessments (Tas) in case of low number of observations (clarification response B6b),[4] and 3) the selected model fit for TTS data in the data for SoC arm.[5]

For the VHL CNS Hb cohort, the pre-surgery → surgery transition probability was estimated using an exponential model fitted to the subpopulation of the MK-6482-004 trial that had also CNS Hb tumours (n=50). Note, that only two CNS Hb surgeries were observed in this cohort (both performed on the same patient). The exponential distribution was stated to be selected based on visual inspection of model fit and statistical goodness-of-fit, and clinically plausibility for long-term extrapolations and the fewer assumptions in the exponential model for underlying risk over time (constant hazards).

There were no surgeries observed in the VHL pNET subpopulation of the MK-6482-004 trial (n=22). Therefore, to estimate the pre-surgery → surgery transition probability in the VHL pNET cohort, the company used the respective probability in the SoC arm and the hazard ratio (HR) of Belzutifan versus SoC in the VHL RCC cohort, adjusted though for the ORR of Belzutifan patients with respect to the pNET versus RCC tumours (91% for pNET compared with 64% for RCC tumours) (see Table 4.5).

There were no metastatic events prior to surgery observed in the MK-6482-004 trial. Therefore, the transition probabilities from pre-surgery → metastatic disease in the three cohorts were estimated by applying the HR of pre-surgery → surgery for Belzutifan versus SoC to the hazard of developing metastatic disease for SoC (see also response to question B3 of the clarification letter).[4] This assumption indicates the treatment effect of Belzutifan on the risk of surgery is equal to the treatment effect on the risk of metastases. According to the company, this assumption is clinically plausible given Belzutifan is expected to reduce both the risks of surgeries and metastatic disease by reducing the tumour size and/or inhibiting their growth.[5]

Also, VHL-related tumour deaths were not observed in the MK-6482-004 trial. Therefore, the transition probabilities pre-surgery → death in the three cohorts were informed from the respective probabilities in the SoC arms, adjusted though for a reduced risk of death in the Belzutifan arm on the basis of a reduced size of VHL CNS Hb tumours observed in the MK-6482-004 trial population. The reduced mortality rate for the VHL RCC cohort was assumed to be attributable to a lower rate of secondary VHL CNS Hb tumours progression (due to a lower size of CNS Hb tumours in the MK-6482-004 trial) and was set equal to the percentage reduction in the risk of pre-surgery → surgery with Belzutifan versus SoC in the VHL CNS Hb cohort. According to the CS, the weekly exponential hazard rate for the transition from pre-surgery → surgery for patients with CNS-Hb in the MK-6482-004 trial (n=50) was 95% lower in the period following treatment with Belzutifan compared to the pre-treatment period (transition probability was 0.00010 following Belzutifan in the MK-6482-004 trial versus 0.00202 during the pre-treatment period) and ORR was 44.0% among patients who had CNS Hb tumours at baseline visit (n=50). Note that two deaths due to other causes occurred in the trial population one due to suicide and one due to toxicity from fentanyl and other agents. In the VHL CNS Hb cohort, the

company assumed that 50% of the pre-surgery → death transitions were attributable to CNS Hb progression, using evidence from Lonser et al. 2014, a prospective study of CNS Hb in VHL disease (N=225) which reported that four out of eight deaths were caused by CNS Hb progression.[34] In the VHL RCC and VHL pNET cohorts, 82% and 89% of patients were assumed to also have CNS Hb, respectively, based on the cross-sectional survey VHL RW QoL Disease Burden Study. Based on these percentages the pre-surgery → death transitions attributable to CNS Hb progression was calculated at 41% and 44.5% in the VHL RCC and VHL pNET cohorts, respectively.

EAG comment:

The main concerns of the EAG regarding the derivation of transition probabilities from the pre-surgery health state in the Belzutifan arm are the following:

a) Transition probabilities for the Belzutifan arm were estimated using data from the MK-6482-004 trial. As mentioned in previous Sections, the population in the trial does not match the population in the DP. In addition, these transitions were either based on a small number of observed events or derived from other assumptions. For instance, for the VHL RCC cohort, which was the population of the MK-6482-004 trial, there were only seven renal surgeries observed (pre-surgery → surgery transition), in the CNS-Hb cohort there were only two surgeries observed on the same patient (pre-surgery → surgery transition), whilst there were no surgeries observed in the VHL pNET cohort. Furthermore, there were no metastatic events or deaths observed in any of the cohorts. This indicates that data from the MK-6482-004 trial (also for the pre-treatment period), and consequently, the survival analyses conducted by the company are subject to great uncertainty (see Section 5.3.2.2 for details). As explained below, this uncertainty is also applicable to the SoC arm for the VHL pNET and VHL CNS Hb cohorts, for which the pre-treatment period data from the MK-6482-004 trial were used to estimate pre-surgery → surgery transitions.
b) As mentioned above, no metastatic events prior to or following surgery were observed in the MK-6482-004 trial. Therefore, the transition probabilities from pre-surgery → metastatic disease in the three cohorts in the Belzutifan arm were estimated by using the HR of pre-surgery → surgery estimated when comparing Belzutifan versus SoC. This assumption implies that the treatment effect of Belzutifan on the risk of surgery is equal to the treatment effect on the risk of metastases. According to the company, this assumption is clinically plausible given Belzutifan is expected to reduce both the risks of surgeries and metastatic disease by reducing the tumour size and/or inhibiting their growth.[5] However, it is unclear to the EAG why it would be expected that the treatment effect on the risk of metastases and surgeries would be exactly of the same magnitude.

Surgery health state

Patients entering the tunnel state of surgery remain in it for one week (one model cycle) and thereafter are assumed to lose organ function if primary tumour was VHL RCC or VHL pNET, or to experience brain injury for VHL CNS Hb primary tumours. The company in response to the CL explained that the term primary tumour throughout the submission describes the tumour that defines treatment decision.[4] Patients in the surgery health state can then transition to either event-free after surgery or death, with the latter representing the risk of perioperative mortality (death as consequence of immediate surgery).

Event-free after surgery health state

As mentioned above, no metastatic disease events were observed in the MK-6482-004 trial. Therefore, the transition probability event-free after surgery → metastatic disease in the VHL RCC cohort was estimated by multiplying the HR of pre-surgery → surgery for Belzutifan versus SoC with the hazard rate of event-free after surgery → metastatic disease in the SoC arm (see Table 4.5). As explained below

in Section 4.2.6.2, the hazard rate of event-free after surgery → metastatic disease in the SoC arm was estimated using patient-level time-to-event data from the VHL Natural History Study.

For the VHL CNS Hb and VHL pNET cohorts, the transition probability for event-free after surgery → metastatic disease was set equal to the transition probability for pre-surgery → metastatic disease (see Table 4.5). Transitions from the event-free after surgery state → death in the VHL RCC cohort were informed from the SoC arm probabilities, adjusted for a reduced risk of death attributable to a reduced size of VHL CNS Hb tumours as explained above in the pre-surgery → death transition of the VHL RCC cohort (see Table 4.5). For the VHL CNS Hb and VHL pNET cohorts, transition probabilities from the event-free after surgery state → death were assumed equal to the probabilities from pre-surgery → death presented above (see Table 4.5).

4.2.6.2 SoC arm: Transitions from non-metastatic disease states

Pre-surgery health state

In the SoC arm, 90% of the VHL RCC and VHL pNET cohorts and 50% of the VHL CNS Hb cohort were assumed to undergo immediate surgery upon model entry.[5]

For the remaining 10% of the VHL RCC cohort, the pre-surgery → surgery transition probability was informed from TTS data from the VHL Natural History Study, with the Natural History data being first reweighted using propensity score matching for the baseline characteristics of patients in the MK-6482-004 trial (see Table 4.5 above). The CS indicated that the exponential distribution was the best model for TTS data from the VHL Natural History Study, grounded on consistency with the model selected for the respective data in the Belzutifan arm, and because this distribution showed the best fit compared to other parametric models based on visual inspection, statistical goodness-of-fit and clinical plausibility.[5]

For the VHL pNET and VHL CNS Hb cohorts, the pre-treatment period of the MK-6482-004 trial was used to inform the pre-surgery → surgery transitions in the SoC arm because patients with CNS Hb or pNET tumours in the VHL Natural History Study could not be identified on the patient-level index date (see response in B7 of the clarification letter).[4] Specifically, parametric survival models were fitted to TTS data from baseline to the most recent pre-baseline CNS Hb and pNET tumour surgeries for the VHL CNS Hb and VHL pNET subgroups of the MK-6482-004 trial, respectively. According to the CS, an exponential distribution was selected for the cohorts of VHL pNET, and VHL CNS Hb patients based on statistical goodness-of-fit, visual inspection and clinical plausibility. Note that for the VHL CNS Hb cohort, the pre-surgery → surgery transition probability in the SoC arm was estimated from the pre-treatment data of the MK-6482-004 trial, and it was only used to inform the surgery rates when the treatment effect derived from the Belzutifan treatment in the Belzutifan arm waned. That is because in the SoC arm, 100% of the patients in the VHL CNS Hb cohort are assumed to experience the outcomes associated with immediate surgery due to the tumour location, which is anticipated to create the same neurological disability for the remaining 50% not operated on as the serious complications from CNS surgery in immediately operated patients (see Table 4.5 above).

The transition probability from pre-surgery → metastatic disease in the VHL RCC cohort was estimated based on parametric models fitted to data from the VHL Natural History Study. Also, for the VHL CNS Hb and VHL pNET cohorts the transition probabilities from pre-surgery → metastatic disease were estimated based on parametric models fitted to data from the VHL Natural History Study using patients with a pre-index history of CNS Hb and pNET tumours, respectively, because the pre-treatment period data from MK-6482-004 could not be used to estimate any transition probabilities from pre-surgery → metastatic disease or from event-free after surgery → metastatic disease or death as per the trial

eligibility criteria (see Table 4.5 above). According to the CS, the exponential distribution was the preferred parametric model for the three cohorts based on statistical goodness-of-fit, visual inspection and clinical plausibility.[5]

The transition probabilities from pre-surgery → death in the SoC arm of the VHL RCC, VHL CNS Hb and VHL pNET cohorts were set equal to the maximum of the mortality rates of the respective RCC, CNS Hb and pNET cohorts of the VHL Natural History Study and the mortality of rate the general population using age- and gender-specific adjustments (see Table 4.5 above). The CS stated that only few pre-surgery deaths were observed in the VHL CNS Hb cohort of the VHL Natural History Study, and that this population had disease which was less severe compared to the Belzutifan-eligible population.

EAG comment:

The main concerns of the EAG regarding the derivation of transition probabilities from the pre-surgery health state in the SoC arm are the following:

a) To estimate the transition probability from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm, the company used a retrospective analysis of the MK-6482-004 trial based on data from the pre-treatment period of the trial. That is because patients with CNS Hb or pNET tumours in the VHL Natural History Study could not be identified on the patient-level index date (see response in B7 of the clarification letter).[4] The CS states that for this analysis patient-level data were selected "*(looking backwards) to the most recent primary tumour surgery prior to Belzutifan initiation in patients with VHL-CNS Hb and VHL-pNET tumours in the MK-6482-004 trial"*.[5] As the primary tumour surgery in these subgroups could be a surgery either due to RCC tumour or pNET or CNS Hb, it is unclear to the EAG why the company used the subgroups to build the economic model. Focusing on VHL pNET or VHL CNS Hb patients would only make sense if the primary tumour would be VHL pNET and VHL CNS Hb, respectively. Considering the company's definition of the primary tumour throughout the submission (the tumour that defines treatment decision), the previous does not seem to be the case. Therefore, it remains a question to the EAG, why did the company decide to model the impact of Belzutifan treatment by using three different cohorts as previously highlighted in the EAG comments of model structure (see Section 4.2.2 for additional details).
b) In response to clarification question B7b, the company stated that using the pre-treatment data from the MK-6482-004 trial to inform transitions from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm, entails potential biases due to using two different data sources to define transitions within the same cohort: the pre-treatment period of the MK-6482-004 trial and the VHL Natural History Study.[4] Potential biases can be attributed to different treatment options, disease management and pathways. Furthermore, the company acknowledged that there may also be implications when comparing these two cohorts with the VHL RCC cohort of the SoC arm for which transition probabilities were completely informed from the VHL Natural History Study. In that perspective, to illustrate the impact of using the pre-treatment period trial data, the EAG requested the company to provide estimates of the respective transition probabilities for the VHL RCC cohort using the pre-treatment data from MK-6482-004 trial instead of the VHL Natural History Study (clarification question B7c).[4] Table 4.6 summarises the model inputs that were estimated based on the pre-treatment MK-6482-004 trial data. The company noted that under this scenario, the incidence rate and distribution of non-RCC surgeries for SoC in the VHL RCC cohort should be also informed from the MK-6482-004 pre-treatment data. According to the company's statement, "*the inputs were not dramatically different under these data sources; when using pre-treatment period data, the rate of pre-surgery → first RCC surgery decreased, while the incidence*

of non-RCC surgeries increased. Additionally, the ICER improves when using this data source".[4] The EAG does not agree with the company's perspective on the "*inputs not being dramatically different*", as the rate of RCC surgeries decreased to more than half of the rate estimated from the VHL Natural History Study, which can be seen as a substantial decrease. The impact of this change is also reflected on the incremental cost-effectiveness ratio (ICER) which dropped by more than 40% compared to the base-case ICER. The EAG confirmed that this change in ICER is driven by the change in the incidence rate of pre-surgery → first RCC surgery, and not by the change in the incidence of non-RCC surgeries. This analysis is further reflecting the EAG concerns on the appropriateness of using the MK-6482-004 pre-treatment period data to estimate the pre-surgery to surgery transitions for the VHL CNS Hb and VHL pNET cohorts and the potential impact these parameters may have in the currently presented company's base-case analysis. Based on the difference in the incidence rate for the VHL RCC cohort as estimated by the VHL Natural History Study and the pre-treatment period data from the MK-6482-004 trial (shown in Table 4.6), it could be expected that the pre-surgery → surgery incidence rate for the VHL-CNS Hb and VHL-pNET cohorts in the SoC arm might also be underestimated when using the pre-treatment period data. To reflect the impact of this potential underestimation the EAG increased the pre-surgery → surgery incidence rate in the SoC arm of the VHL-CNS Hb and VHL-pNET subgroups is doubled considering the rate of RCC surgeries was estimated to be more than double when using the VHL Natural History Study as compared to the pre-treatment period of the trial (shown in Table 4.6). The results of this scenario are presented in Section 6.2.

Table 4.6: Parameter values for SoC in the VHL RCC subgroup using the MK-6482-004 pre-treatment data

Parameter	VHL Natural History Study	MK-6482-004 pre-treatment period data
Weekly rate of pre-surgery → RCC surgery	0.00487	0.00207
Weekly rate of non-RCC tumour surgeries	0.00344	0.00438
Distribution of surgeries for non-RCC tumours		
CNS Hb surgery	52.4%	67.4%
pNET surgery	3.4%	7.0%
Adrenal lesion surgery	25.1%	2.3%
Endolymphatic sac tumour surgery	4.5%	0.0%
Epididymal cystadenoma surgery	0.1%	2.3%
Retinal Hb surgery	14.5%	20.9%
ICER	**£73,095**	**£42,622**

Based on Table 8 of the CL response.[4]
CL = clarification letter; CNS = central nervous system; Hb = hemangioblastomas; ICER = incremental cost-effectiveness ratio; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Surgery health state

We refer to the description of this health state in the Belzutifan arm.

Event free after surgery health state

The transition probabilities from event-free-after surgery → metastatic disease or death in the SoC arm of the VHL RCC cohort were estimated using patient-level time-to-event data from the VHL Natural History Study, after the natural history data were reweighted using propensity score matching for the baseline characteristics of patients in the MK-6482-004 trial. According to the CS, an exponential distribution was fitted to patient-level data on time from the first post-index renal surgery date until each metastatic disease or death occurred.[5] The per cycle probability of death for the VHL RCC cohort was then set equal to the maximum of probability of death estimated by the exponential model and the mortality rate of the general population. In the CS, the company argued that for the CNS Hb and pNET cohorts, the number of patients with ≥1 post-index primary tumour surgery was too small (n=3.8 for CNS Hb and n=14.1 for pNET) in the VHL Natural History Study to fit parametric models.[5] Therefore, the respective transition probabilities from event-free-after surgery → metastatic disease and death health states for the VHL CNS Hb and VHL pNET cohorts were assumed to be equal to the respective transitions from the pre-surgery state, as estimated from the VHL Natural History Study.

EAG comment: The main concerns of the EAG regarding the derivation of transition probabilities from the event-free after surgery health state in the Belzutifan and SoC arms are the following:

a) The transition probability from event-free after surgery → metastatic disease in the VHL-RCC cohort in the Belzutifan arm was estimated using the HR of pre-surgery → surgery for Belzutifan versus SoC, whereas for the VHL-CNS Hb and VHL-pNET cohorts in both arms, the transition probabilities for event-free after surgery → metastatic disease were set equal to the transition probability for pre-surgery → metastatic disease. It is unclear to the EAG what the evidence basis for these assumptions is and if those have been validated by clinical experts.

4.2.6.3 Transition probabilities from metastatic disease to death

The transition probabilities from metastatic disease to death were assumed to be equal between the Belzutifan and SoC arms for the three cohorts. Metastases were assumed to originate either from RCC or pNET tumours. This assumption was based on data from the VHL Natural History Study showing that RCC and pNET tumours were the origin tumour in most metastatic patients. In the VHL CNS Hb cohort, it was assumed that all metastatic cases originated from non-primary VHL RCC or VHL pNET tumours, as CNS Hb tumours do not metastasise. For each cohort the probability of metastases by origin tumour is presented in Table 4.7.

Table 4.7: Probability of metastases by origin tumour

Subgroup	RCC	pNET
VHL RCC	97%	3%
VHL CNS Hb	78%	22%
VHL pNET	66%	34%

Based on Table 61 of the CS.[5]
CNS = central nervous system; Hb = hemangioblastomas; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

As tumour origin defines treatment options for metastatic disease, to estimate transition probabilities from metastatic disease → death, OS for metastatic patients in each treatment arm was estimated based

on the weighted average of OS data associated with first-line metastatic disease treatments. The weights of the OS curves were based on market shares of first-line treatments (see Table 62 and Table 63 of the CS),[5] which were informed from the NICE appraisal of pembrolizumab as an adjuvant treatment of RCC post-nephrectomy (TA830) for the VHL RCC cohort,[35] the European Society for Medical Oncology (ESMO) clinical practice guidelines for gastroenteropancreatic neuroendocrine neoplasms and input from clinical experts for the VHL pNET cohort.[36] This approach aligns with the approach taken in the NICE appraisal of pembrolizumab for adjuvant treatment of RCC (TA830).[37] For metastatic disease with RCC as the origin tumour, the first-line treatment options considered in the model were sunitinib (30%), tivozanib (14%), pazopanib (29%), cabozantinib (13%), and nivolumab combined with ipilimumab (14%) (refer to Table 63 of the CS).[5] For metastatic disease with pNET as the origin tumour, the first-line treatment options considered in the model were lanreotide (50%) and octreotide (50%) (refer to Table 64 of the CS).[5]

For first-line treatment options of metastatic disease with RCC as the origin tumour, the company used exponential distributions to model OS and PFS. For sunitinib, exponential distributions were fitted based on the observed median OS and PFS in the sunitinib arm of the KEYNOTE-426 trial as shown in Table 4.8, aligning with the parametric model selected in a previously published CEA of sunitinib based on the KEYNOTE-426 trial.[38] To estimate OS and PFS curves for the other advanced treatment regimens, HRs for OS and PFS versus sunitinib were used. The HRs were estimated from a fixed effects parameters of a network meta-analysis (NMA) conducted using RCTs of first-line treatments in patients with locally advanced unresectable or metastatic RCC. As pembrolizumab/lenvatinib was not included in this NMA, the HRs for OS and PFS of this treatment versus sunitinib were obtained from Motzer et al. (2021).[39] Table 4.9 presents the HRs of OS and PFS for all treatment regimens versus sunitinib as obtained from the NMA, as well as estimates of mean weekly OS and PFS for each regimen.

For the first-line treatment options of metastatic disease with pNET as the origin tumour, the company used exponential distributions to model OS and PFS. For streptozocin/5-fluorouracil and no active treatment in the advanced pNET setting, exponential rates were estimated based on OS and PFS data extracted from Study E1281 and NCT00428597 trial, as shown in Table 4.8.[40, 41] The OS and PFS of streptozocin/doxorubicin and temozolide/capecitabine, were assumed to be equal to that of streptozocin/5-fluorouracil, as these three combination regimens are indicated for higher-grade pNET. To estimate OS and PFS curves for the other advanced treatment regimens, HRs versus no active treatment were used which were obtained from an SLR and NMA of trials conducted in advanced pNET, as shown in Table 4.9.[42]

Table 4.8: Exponential models of OS and PFS in advanced RCC and pNET

Subgroup/ Advanced Regimen	Exponential model of OS		Exponential model of PFS		Source
	Rate	SE	Rate	SE	
VHL RCC					
Sunitinib	0.0040	(0.0003)	0.0144	(0.0013)	KEYNOTE-426 (Rini et al. 2021)[*3]
VHL pNET					
Streptozocin / 5-fluorouracil	0.0066	(0.0016)	0.0301	(0.0055)	Sun et al. (2005) [Study E1281][^40]
No active treatment	0.0055	(0.0014)	0.0275	(0.0051)	Faivre et al. (2017) [NCT00428597][#41]
Based on Table 64 and Table 66 of the CS.[5]					

Subgroup/ Advanced Regimen	Exponential model of OS		Exponential model of PFS		Source
	Rate	SE	Rate	SE	

* For sunitinib in the advanced RCC setting, exponential models of OS and PFS were calculated based on median PFS and OS reported from the KEYNOTE-426 trial.[3]
^ For streptozocin/5-fluorouracil in the advanced pNET setting, exponential models of OS and PFS were calculated based on median OS and PFS reported from Study E1281 (Sun et al. 2005).[40] Of note, the rates of OS and PFS failure are higher for streptozocin/5-fluorouracil than for no active treatment, as streptozocin/5-fluorouracil is indicated for higher-grade pNET.
# For patients who receive no active treatment in the advanced pNET setting, exponential models of OS and PFS were calculated based on median OS and PFS reported from NCT00428597 (Faivre et al. 2017).[41] 	
CS = company submission; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; OS = overall survival; PFS = progression-free survival; SE = standard error; SoC = standard of care; VHL = Von Hippel-Lindau					

Table 4.9: HRs of OS and PFS failure with other treatments vs. sunitinib in advanced RCC and pNET

Subgroup/ Advanced Regimen	HR of death vs. sunitinib		HR of progression or death vs. sunitinib		Expected survival in metastatic state (weeks)	
	HR	SE of ln(HR)	HR	SE of ln(HR)	OS	PFS
VHL RCC						
Sunitinib	1.00		1.00		252	70
Tivozanib	1.33	0.27	1.19	0.26	189	59
Pazopanib	0.92	0.08	1.05	0.08	273	66
Cabozantinib	0.80	0.21	0.48	0.22	314	145
Nivolumab/ipilimumab	0.72	0.08	0.89	0.08	349	78
Avelumab/axitinib	0.80	0.13	0.69	0.09	314	101
Pembrolizumab/lenvatinib	0.66	0.15	0.39	0.11	381	179
VHL pNET						
Streptozocin/5-fluorouracil	1.20	-	1.09	-	152	33
Streptozocin/ doxorubicin	1.20	-	1.09	-	152	33
Temozolomide/ capecitabine	1.20	-	1.09	-	152	33
Everolimus	0.35	0.12	0.35	0.12	522	104
Sunitinib	0.42	0.24	0.42	0.24	435	87
Interferon a2B	0.37	0.42	0.37	0.42	493	98
Lanreotide	0.46	0.18	0.46	0.18	397	79
Octreotide	0.46	0.18	0.46	0.18	397	79
No active treatment	1.00	-	1.00	-	183	36

Subgroup/ Advanced Regimen	HR of death vs. sunitinib		HR of progression or death vs. sunitinib		Expected survival in metastatic state (weeks)	
	HR	SE of ln(HR)	HR	SE of ln(HR)	OS	PFS
Based on Table 65 and Table 67 of the CS.[5] CS = company submission; HR = hazard ratio; OS = overall survival; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SE: standard error; VHL = Von Hippel-Lindau						

For the VHL CNS Hb cohort, the transition rate of metastasis to death was calculated as in the VHL RCC and VHL pNET cohorts, using the distribution of origin tumours in the VHL CNS Hb population (presented in Table 4.7), the market shares of first-line treatments for VHL RCC and VHL pNET, and the efficacy of the first-line treatments for VHL RCC and VHL pNET tumours. Note again that in this cohort, it was assumed that all metastatic cases originated from non-primary VHL RCC or VHL pNET tumours, given that CNS Hb tumours do not metastasise.

For each subgroup, mean OS (in weeks) within the metastatic disease state calculated as a weighted average of estimated OS associated with different first-line treatments for advanced RCC and pNET, based on the origin tumour distribution and market shares of first-line advanced treatments is shown in Table 4.10. Mean OS was then translated into a weekly HR and included in the economic model.

Table 4.10: Hazard rates (weeks) for transitions from metastatic disease to death by subgroup and treatment arm as included in the electronic model

Subgroup/ Treatment arm	Expected survival in metastatic state (weeks): Weighted average based on origin tumour and first-line advanced treatment market shares			Hazard rate of metastatic disease → death
	OS	PFS	Ratio of OS/PFS	
VHL RCC				
Belzutifan	275	78	0.28	0.0036
SoC	275	78	0.28	0.0036
VHL pNET				
Belzutifan	299	78	0.26	0.0033
SoC	299	78	0.26	0.0033
VHL CNS Hb				
Belzutifan	314	78	0.25	0.0032
SoC	314	78	0.25	0.0032
Based on Table 68 of the CS.[5] CS = company submission; CNS Hb = central nervous system haemangioblastoma; pNET = pancreatic neuroendocrine tumour; ORR = overall response rates; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau				

EAG comment:

The main concerns of the EAG regarding the derivation of transition probabilities from the metastatic disease health state to death in both arms are the following:

a) The company relied on many assumptions to model the transition probabilities from metastatic disease to death. The EAG was unable to verify all of them.

b) It is unclear why many treatment options are included, but then some of the market shares are set at 0%. For the VHL RCC cohort, the market shares of avelumab/axitinib and pembrolizumab/lenvatinib are set out in Table 62 of the CS.[5] For the VHL pNET cohort, the market shares for streptozocin/5-fluorouracil, streptozocin/doxorubicin, temozolomide/capecitabine, everolimus, sunitinib, interferon a2B and no active treatment are set to 0%, while these have been included in the NMA analysis.[42, 43] The EAG wonders whether this was performed to allow connections for treatment effectiveness in the NMA network and if a simpler approach would have been more transparent in this case.
c) In principle, AEs associated to the treatments in the metastatic disease health state should have also been included in the model, although usually the impact of AEs in model results is not major.

4.2.6.4 Adjusting the risk of surgery and metastatic disease

In an attempt to better reflect the expected level of care in the UK clinical practice the company further adjusted the transition probabilities from i) pre-surgery → surgery (except for those receiving immediate surgery in the SoC arm), ii) pre-surgery → metastatic disease, and iii) event-free after surgery → metastatic disease in both arms using data from the Optum Clinformatics Data Mart Claims database.[33] The justification for this adjustment was that patients in the VHL Natural History Study were thought to have received an elevated SoC compared with UK clinical practice because patients in the VHL Natural History Study were treated at the US NCI, which is a Centre of Excellence.[5] Similarly, according to the CS patients in the MK-6482-004 trial underwent through imaging procedures and surgeries more frequently than would have been expected in the UK clinical practice.[5] Therefore, rates of surgeries in the VHL Natural History Study and the MK-6482-004 trial were considered higher, whilst rates of metastatic disease were considered lower than would have been expected. The Optum Clinformatics Data Mart Claims database, was considered more reflective of real-world clinical practice as it consisted of real-world data from 160 patients with VHL from a wide geographic area in the US and it does not represent a Centre of Excellence.[33]

Therefore, the company added the difference in the cause-specific hazards of event-free after surgery → next surgery between the Optum Clinformatics Data Mart Claims database and the VHL Natural History Study population to the cause-specific hazards of pre-surgery → surgery (except for those receiving immediate surgery in the SoC arm) in the SoC arm. Similarly, the difference in the observed cause-specific hazards of event-free after surgery → metastatic disease between the Optum Clinformatics Data Mart Claims database and the VHL Natural History Study population was added to the cause-specific hazard rates of pre-surgery → metastatic disease and event-free after surgery → metastatic disease in the SoC arm of the model. In the Belzutifan arm, respective cause-specific hazard rates of these transitions were adjusted accordingly by using the HRs of these transitions with Belzutifan versus SoC as estimated based on the VHL Natural History Study or pre-treatment period of the MK-6482-004 trial, so that the relative treatment effects of Belzutifan versus SoC on these transitions remained unchanged when the adjustment is applied. Note, that the cause-specific hazard rates of pre-surgery → surgery were only adjusted for the VHL RCC cohort, as respective transitions for patients not receiving immediate surgery in the VHL CNS Hb and VHL pNET cohorts were informed from the pre-treatment period of MK-6482-004 trial, and it was deemed uncertain if patients received an elevated SoC during this period. The differences in hazard rates as estimated from the Optum Clinformatics Data Mart Claims database and the VHL Natural History Study are presented in Table 4.11.

Table 4.11: Adjustment factors applied to the hazard rates of surgery and metastatic disease in the three subgroups

Difference in hazard rates	VHL RCC	VHL CNS Hb	VHL pNET
Event-free after surgery to next surgery*	-0.00109	NA	NA
Event-free after surgery to metastatic disease	0.00115	0.00107	0.00255

Based on Table 61 of the CS.[5]
* Applied to pre-surgery → surgery
CNS Hb = central nervous system hemangioblastoma; CS = company submission NA = not applicable; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

EAG comment:

The main concerns of the EAG relate to:

a) The company adjusted the transition probabilities; i) pre-surgery → surgery, ii) pre-surgery → metastatic disease, and iii) event-free after surgery → metastatic disease in both treatment arms using data from the Optum Clinformatics Data Mart Claims database.[33] Adjustments to surgery and metastases rates were only conducted for the VHL RCC cohort, whilst adjustments to metastases rates only were made for the VHL CNS Hb and VHL pNET cohorts. The underlying assumption was that patients in the VHL Natural History Study and in the MK-6482-004 trial were subject to "high-quality" SoC. While maintaining the treatment benefit of Belzutifan compared to SoC, the surgery rates were further lowered compared to those observed in the MK-6482-004 trial. As mentioned in the EAG comment in Section 4.2.2, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial, underestimate the surgery rates for the population in the decision problem that would use Belzutifan in clinical practice. This additional adjustment would only reinforce that underestimation. All in all, the EAG considers it likely that more severe patients would be at greater risk of surgery and metastasis. It could also be that Belzutifan is more effective in more severe patients, but with the current evidence it is unknown if the treatment effect would be the same as the one modelled (based on the MK-6482-004 trial) in a more severe population, so the underlying adjustments are not actually resolving uncertainties, rather than imposing additional assumptions. Furthermore, to reflect the impact of aligning adjustments between cohorts, the EAG presented a scenario in which the surgery rates were also adjusted for the VHL-CNS Hb and VHL-pNET cohorts at a similar relative reduction as the rates in the VHL-RCC cohort.
b) The EAG also considers that there are inherent uncertainties in the estimation of these transition probabilities, imposed by using data from a different clinical practice (the US versus the UK), which cannot be resolved by the implementation of additional adjustments, as these assumptions on adjustment parameters would likely be uncertain as well.

4.2.6.5 Perioperative mortality and non-primary surgeries

For both the Belzutifan and SoC arms and all three cohorts, patients undergoing surgery enter a tunnel state for one cycle (one week). Patients transitioning from surgery → death represent the risk of perioperative mortality following a primary tumour surgery. Table 4.12 shows the perioperative mortality risks in each of the three cohorts, which were assumed to be the same for Belzutifan and SoC arms. For the VHL RCC cohort, the perioperative mortality risk was 1.96% (1/51) in a retrospective centre-based chart review of 51 repeat partial nephrectomies performed in patients with hereditary renal

cancer.[44] For the VHL CNS Hb cohort, perioperative mortality risk was 1.82% (1/55) in a retrospective chart review of 55 resections of spinal cord Hb performed at the US NIH in patients with VHL disease.[45] For the VHL pNET cohort, the perioperative mortality risk was 1.7% (2/117) in a multicentre international registry study of 273 patients with VHL pNET.[46] The company further adjusted the aforementioned perioperative mortality risks by a factor of 2.0 grounded on clinical expert opinion in an attempt to better reflect the MHRA population which defines Belzutifan treatment appropriate for patients "for whom localised procedures are unsuitable or undesirable".[5]

Table 4.12: Perioperative mortality risk in the three subgroups

Subgroup	Surgery → death	Source
	Risk (SE)	
VHL RCC	0.0196 (0.01941)	Johnson et al. (2008)[44]
VHL-CNS-Hb	0.0182 (0.01802)	Lonser et al. (2003)[45]
VHL pNET	0.0171 (0.01198)	Krauss et al. (2018)[46]

Based on Table 51 of the CS.[5]
CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SE = standard error; VHL = Von Hippel-Lindau

For the VHL RCC cohort, the incidence rate of non-primary tumour surgeries in the Belzutifan arm presented in Table 4.13 was informed from the MK-6482-004 trial, whilst in the SoC arm data from the VHL Natural History Study were used.[47] For the VHL RCC cohort of both arms, the overall incidence rate of non-primary tumour surgeries was proportionally attributed to specific non-RCC VHL manifestations (i.e., CNS Hb, pNET, adrenal lesion, endolymphatic sac tumour, epididymal cystadenoma, or retinal Hb) based on the observed percentage breakdown of non-primary tumour surgery events in the reweighted VHL Natural History Study population subgroup as shown in Table 4.14. For the VHL CNS Hb and VHL pNET cohorts, the incidence rates of non-primary tumour surgeries were derived from the respective subgroups of patients in the MK-6482-004 trial for both the Belzutifan and SoC arms using respective data from the post-treatment and pre-treatment period of Belzutifan treatment. As all patients in the MK-6482-004 trial had RCC tumours, the company used the percentage of non-primary tumour surgeries that were due to RCC during the pre-treatment period of the MK-6482-004 trial and the proportion of patients with VHL CNS Hb and without RCC from the VHL RW QoL Disease Burden Study (36.8% (70/190)) to adjust downwards the incidence of non-primary RCC surgeries in the VHL CNS Hb cohort. Similarly, to downsize the incidence of non-primary RCC surgeries in the VHL pNET cohort the proportion of patients with VHL pNET and without RCC from the VHL RW QoL Disease Burden Study (28.7% (35/122)) was used in model computations.

Table 4.13: Incidence rates per year of non-primary VHL-related tumour surgeries, by subgroup and treatment arm

Treatment arm	VHL RCC	VHL CNS Hb	VHL pNET	Source
	Rate of non-RCC surgeries	Rate of non-CNS Hb surgeries	Rate of non-pNET surgeries	
Belzutifan	0.021187	0.029388	0.036498	RCC cohort: MK-6482-004 trial.[24] CNS Hb and pNET cohorts: MK-6482-004 trial,[24] adjusted for RCC surgery incidence

Treatment arm	VHL RCC	VHL CNS Hb	VHL pNET	Source
	Rate of non-RCC surgeries	Rate of non-CNS Hb surgeries	Rate of non-pNET surgeries	
				based on VHL RW QoL Disease Burden Study.[1]
SoC*	0.178984	0.195923	0.340651	RCC cohort: VHL Natural History Study (2021),[26] matched to MK-6482-004 trial.[24] CNS Hb and pNET cohorts: Pre-treatment period data MK-6482-004 trial,[24] adjusted for RCC surgery incidence based on VHL RW QoL Disease Burden Study.[1]

Based on Table 49 of the CS.[5]
*These rates were also used in the Belzutifan arm after treatment effect waning.
CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SE = standard error; SoC = standard of care; VHL = Von Hippel-Lindau

Table 4.14: Distribution of surgeries for non-primary VHL-related tumours by subgroup

Surgery type	VHL RCC	VHL CNS Hb	VHL pNET	Source
RCC surgery	NA	56.7%	38.2%	For RCC cohort: VHL Natural History Study (2021),[26] reweighted to match the MK-6482-004 population.[24] For CNS Hb and pNET cohorts: Analysis of pre-treatment period data from the MK-6482-004 trial,[24] adjusted for RCC surgery incidence based on VHL RW QoL Disease Burden Study (2022).[1]
CNS Hb surgery	52.4%	NA	41.2%	
pNET surgery	3.4%	6.7%	NA	
Adrenal lesion surgery	25.1%	3.3%	0.0%	
Endolymphatic sac tumor surgery	4.5%	0.0%	0.0%	
Epididymal cystadenoma surgery	0.1%	3.3%	0.0%	
Retinal Hb surgery	14.5%	29.9%	20.6%	
Total:	*100%*	*100%*	*100%*	

Based on Table 50 of the CS.[5]
CNS Hb = central nervous system hemangioblastoma; CS = company submission; NA = not applicable; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; QoL = quality of life; VHL = Von Hippel-Lindau

Following treatment discontinuation with Belzutifan, the incidence rate for non-primary tumour surgeries in the Belzutifan arm was set equal to the respective incidence rate in the SoC arm for all three cohorts. Costs due to non-primary tumour surgeries and decrements in QoL due to non-primary tumour

surgery complications were estimated per cycle and added to the costs and QALYs estimated across tumour-related health states.

EAG comment:

The main concerns of the EAG relate to:

a) The EAG considers that doubling the perioperative mortality risk in the three cohorts is an arbitrary adjustment. In response to clarification question B9, the company justifies this increase referring to the population specified by the GB label as it includes patients with more severe manifestations of VHL disease for whom surgery is unsuitable or undesirable.[4] This implies that these patients would also experience a significantly increased risk of perioperative mortality risk. The EAG is of the opinion that these assumptions should be better justified, since considering the mismatch between the population in the trial and the MHRA authorisation, one should also adjust all transition probabilities estimated from the MK-6482-004 trial. Furthermore, the company refers to clinical expert opinion for the decision to make these adjustments, but no formal reference is provided. In response to clarification question C1, the company stated that "*we are unable to provide documentation of this as it contains confidential information and we have not sought permission from participating experts*".[4] For this reason, the EAG considered a scenario analysis in which the adjustment in the perioperative mortality risks was removed from the computations.

4.2.6.6 Surgery-related complications

Complications related to primary and non-primary tumour surgeries were also included in the model. The company distinguished between short-term and long-term complications and both were informed from the Optum Clinformatics Data Mart Claims Study.[33] Short-term complications were assumed to happen within 28 days post-surgery, whereas long-term risks were defined using a follow-up period of 180 days post-surgery and were assumed to be lifelong. Short- and long-term risks of surgery complications were set equal between Belzutifan and SoC arms. Nonetheless, the company stated in the CS that the Optum Clinformatics Data Mart Claims Study data provided an underestimation of the risks related to short- and long-term surgery complications considering the MHRA authorisation of Belzutifan for patients "for whom localised procedures are unsuitable or undesirable".[5] To address the expected underestimation, the company performed upward adjustments using clinical expert input. For short-term complications all risks were doubled as shown in Table 4.15. The risks of most of the long-term surgery complications presented in Table 4.16 were also doubled. Note, that for end stage renal disease (ESRD) and/or dialysis in the VHL RCC cohort and secondary diabetes and immuno-compromisation in the VHL pNET cohort, the risks were respectively increased from 4% to 80%, from 20% to 100% and from 0% to 100% (Table 4.16). Similarly, the risk of cerebral vascular occlusion/stroke was increased from 7.7.% to 85% of the patients in the VHL CNS Hb cohort. According to the CS, the risks of long-term metabolic consequences resulting from surgery were substantially adjusted upwards to capture the limited organ function following surgery in the licensed population and adjustments were made using clinical expert consultation.[5]

Table 4.15: Risks of short-term surgical complications per surgery

Complication	Risk of complication	Risk of complication following adjustment
VHL RCC cohort		
Acute renal failure	8.0%	16.0%
Cardiac complications	4.0%	8.0%
Erythroderma	0.8%	1.6%

Complication	Risk of complication	Risk of complication following adjustment
Kidney infection	1.6%	3.2%
Other genitourinary complications	9.6%	19.2%
Postoperative infection (RCC-related)	6.4%	12.8%
Respiratory complications	20.8%	41.6%
Thrombosis and/or embolism	4.8%	9.6%
Vascular injury or anaemia	13.6%	27.2%
VHL CNS Hb cohort		
Acute renal failure	7.7%	15.4%
CNS haemorrhage	12.8%	25.6%
Nerve palsy related to anaesthesia	5.1%	10.3%
Respiratory complications	20.5%	41.0%
Thrombosis and/or embolism	15.4%	30.8%
Vascular injury or anaemia	15.4%	30.8%
VHL pNET cohort		
Abdominal abscess	10.0%	20.0%
Postoperative infection (pNET-related)	20.0%	40.0%
Respiratory complications	40.0%	80.0%
Thrombosis and/or embolism	10.0%	20.0%
Urinary tract infection	10.0%	20.0%
Vascular injury or anaemia	10.0%	20.0%
Based on Table 53, Table 54, and Table 55 of the CS.[5] CNS Hb = central nervous system hemangioblastoma; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau		

Table 4.16: Risks of long-term surgical complications per surgery

Complication	Risk of complication	Risk of complication following adjustment
VHL RCC cohort		
End stage renal disease and/or dialysis	4.0%	80.0%
Chronic kidney disease	24.0%	20.0%
Hernia surgery	1.6%	3.2%
Chronic pain	8.8%	17.6%
Cerebral vasculature occlusion or stroke	3.2%	6.4%
VHL CNS Hb cohort		
Chronic pain (in CNS Hb population)	15.4%	30.8%
Cerebral vasculature occlusion or stroke	7.7%	85.0%
Seizure	10.3%	20.5%
Neurological complications	43.6%	87.2%
VHL pNET cohort		
Chronic pain (in pNET population)	10.0%	20.0%

Complication	Risk of complication	Risk of complication following adjustment
Secondary diabetes or exocrine pancreatic insufficiency	20.0%	100.0%
Immunocompromisation	0.0%	100.0%
Based on Table 57, Table 58, and Table 59 of the CS.[5] CNS Hb = central nervous system hemangioblastoma; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau		

EAG comment:

The main concerns of the EAG relate to:

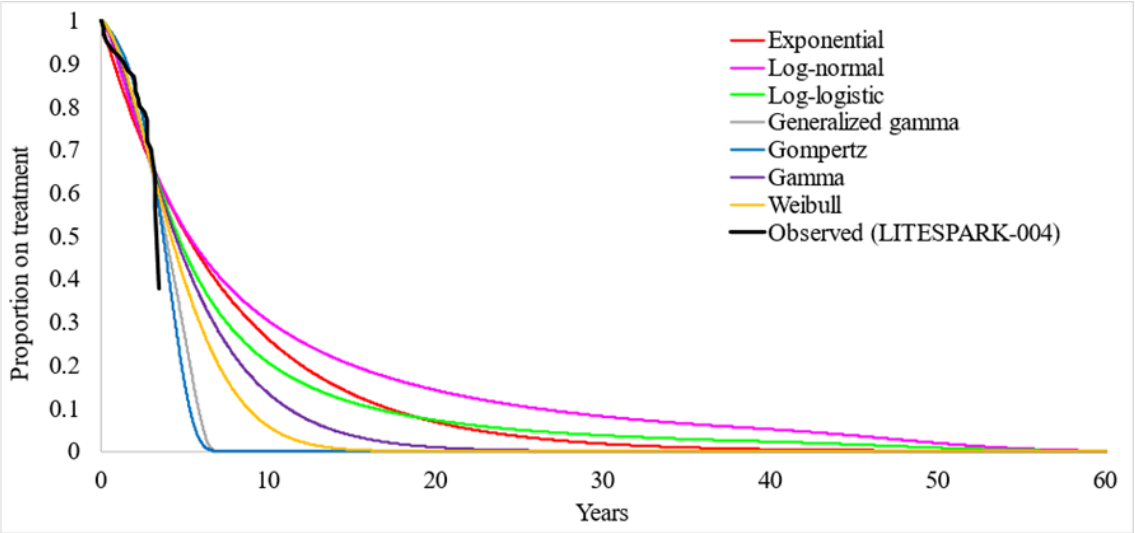
a) The risks of short- and long-term complications following surgery for the VHL RCC cohort are in the majority doubled when considering the MHRA label population compared to the risks estimated from the Optum Clinformatics Data Mart Claims Study. In response to clarification question B9, the company commented that this increase is justified by the nature of the population specified by the GB label.[4] According to the company, the target population in the decision problem including patients with more severe manifestations of VHL disease for whom surgery is unsuitable or undesirable, implies that these patients would also experience a significantly increased risk of surgery-related complications. The company refers to clinical expert opinion for the decision to make these adjustments, but no formal reference is provided. In response to clarification question C1, the company stated that "*we are unable to provide documentation of this as it contains confidential information and we have not sought permission from participating experts*".[4] Therefore, the EAG cannot assess the validity of these parameters.
b) Furthermore, as explained in Section 4.2.6.4 in their base-case analysis the company adjusted the transition probabilities from i) pre-surgery → surgery, ii) pre-surgery → metastatic disease, and iii) event-free after surgery → metastatic disease in both arms using data from the Optum Clinformatics Data Mart Claims database, because according to the company the adjusted values would better reflect the expected level of care in the UK clinical practice. The EAG considers these statements as potential source of contradiction given that if the Optum Clinformatics Data Mart Claims database reflects UK clinical practice better for the transitions of pre-surgery → surgery, it could also be argued that it would be an appropriate source for the estimation of short and long-term surgery complications. Therefore, this increase in the risk of short and long-term complications seems arbitrary, and for this reason the EAG considered a scenario analysis in which the risks of short- and long-term complications following surgery were set equal to the risks estimated from the Optum Clinformatics Data Mart Claims Study.
c) In response to clarification question B9, the company stated that the risk for chronic kidney disease is lower "*as it is expected after a 'last-resort' RCC surgery that all patients would have some form of renal impairment, of which 80% would have end-stage renal disease (ESRD) requiring dialysis and the remainder (20%) would have chronic kidney disease, an assumption which was informed by clinical expert opinion*".[4] Similarly, the risk of ESRD (in the VHL RCC cohort), secondary diabetes or exocrine pancreatic insufficiency and immunocompromisation (in the VHL pNET cohort) are sufficiently increased compared with the Optum Clinformatics Data Mart Claims Study grounded on surgery being a 'last resort' for the assessed patient population, and such a surgery is understood to lead to absent/limited organ function. The complications of cerebral vasculature occlusion or stroke in the VHL CNS Hb cohort was increased versus the Optum Clinformatics Data Mart Claims Study assuming a substantially higher risk of this in the target population given their unsuitability for surgery (see clarification response B9).[4] As also mentioned in previous EAG

comments, the company refer to clinical expert opinion to support these adjustments, but no formal reference is provided. Furthermore, similarly to adjustments reported for the risk of surgery and metastasis and the perioperative mortality risk, the EAG has concerns around the arbitrariness of these assumptions and cannot assess the validity of these parameters. For this reason, in Section 6.2 the EAG run a scenario analysis in which the surgery-related complications were set to their original values as estimated from the Optum Clinformatics Data Mart Claims database.

4.2.6.7 Time to treatment discontinuation and treatment waning

Patients in the Belzutifan arm are assumed to stay on treatment until unacceptable treatment-related toxicity or unequivocal disease progression. Different parametric models (exponential, Weibull, Gompertz, log-logistic, log-normal, gamma, and generalised gamma) were fitted to patient-level data on time-to-treatment discontinuation from the MK-6482-004 trial to inform Belzutifan time on treatment (ToT). The Gompertz curve was selected in the company base-case analysis based on model fit evaluated through Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics (see Table 72 of the CS),[5] visual inspection as shown in Figure 4.2 and clinical plausibility. The impact of using the second-best model option based on AIC and BIC values (the Weibull), was explored in the scenario analyses.

Figure 4.2: Modelled vs. observed ToT for Belzutifan in MK-6482-004 trial under different parametric distributions over long-term extrapolation



Based on Figure 23 of the CS.[5]
Note: LITESPARK-004 refers to the MK-6482-004 trial
CS = company submission; ToT = time on treatment

The CS reports that at a median follow-up of 37.7 months (3.14 years), 61% of patients in the MK-6482-004 trial remained on treatment.[5] A waning treatment effect for Belzutifan treatment was therefore included by the company in their base-case analysis to reflect the uncertainty around the long-term impact of treatment discontinuation. The waning effect of Belzutifan treatment is incorporated in the model by using "off-treatment" health states, which are defined separately for all health states, except the health states of metastatic disease and death. Patients transit in the "off-treatment" states in each cycle following the estimated ToT curve of Belzutifan. Transition probabilities from the "off-treatment" health states for patients treated with Belzutifan (including the efficacy of Belzutifan in preventing transitions to surgery, metastases, or death, reducing the incidences of non-primary tumour surgeries, and inducing primary tumour response) were assumed to linearly converge to those of the SoC arm are

assumed to linearly converge to the respective values in the SoC arm estimated using data from the VHL Natural History Study (see response B13 in the clarification letter).[4] The waning effect following treatment discontinuation is implemented for i) all transition probabilities starting from the pre-surgery and event-free after surgery state, ii) the overall response rates which were used to inform the composite utility values (see Section 4.2.8.4), and iii) the incidence rate of non-primary tumour surgeries. In the company base-case, treatment effect waning was assumed to start at 3.84 years (46.1 months), reflecting the maximum follow-up period from the MK-6482-004 trial. According to the CS, initiating treatment effect waning before the end of the trial period would not be appropriate as the estimation of transition probabilities based on the MK-6482-004 trial data used all available follow-up data accounting for patients discontinuing Belzutifan treatment. Belzutifan treatment waning was assumed to occur gradually over a 2.71-year period from the end of the maximum follow-up (i.e., 3.84 years). The 2.71-year period represents the amount of time until the largest RCC tumour reaches the baseline levels of growth (i.e., (24.9 - 15.36)/3.52 = 2.71 years).[34, 48] This was estimated using the average tumour size as estimated to the closest time point of discontinuing Belzutifan (average of 25.2 days from discontinuation) at 15.36 mm and the average size of the largest RCC tumour at baseline (24.9 mm). It was further assumed that the tumour growth rate would revert to pre-treatment levels at an average rate of 3.52 mm/year (as before treatment) after discontinuation, For the VHL pNET and VHL CNS Hb cohorts, the same residual benefit period was assumed due to the small sample size of patients discontinuing in these subgroups with an available CNS Hb and pNET measurement (respectively) near the time of discontinuation.

EAG comment:

The main concerns of the EAG relate to:

a) The choice of the parametric distribution to model ToT. As shown in Figure 4.2, there is uncertainty in the long-term extrapolations. This uncertainty has also been acknowledged by the company in response to clarification B13.[4] Despite this, the company presented only one additional scenario analysis using the second best more fit based on AIC and BIC values (the Weibull distribution). Therefore, the EAG explored the impact of using different parametric models in the scenario analyses.
b) The duration of the residual benefit is also uncertain. The 2.71-year period could be seen as a reasonable choice for the base-case, but it could also be argued that it could be different, especially for CNS Hb and pNET related tumours, for which no data are available. Therefore, also in this case, scenario analyses are appropriate to assess the impact of this assumption on the model results.
c) In response to clarification question B14, the company stated that applying treatment waning before the maximum observed trial period (3.84 years) would lead to a mismatch between "*observed versus predicted curves for time to surgery, metastatic disease, or death in the Belzutifan arm*" as these data account for patient discontinuation (response in question B14).[4] The company noted that "*it would therefore be inappropriate to consider a treatment effect waning assumption of no residual benefit (i.e. before the maximum follow-up period of the trial is complete)".*[4] However, as shown in Table 4.5 and discussed in the EAG comments of Section 4.2.6.1, survival data from the MK-6482-004 trial were only used to estimate transitions in the pre-surgery → surgery transitions for VHL RCC and VHL CNS Hb patients, which were also estimated using a small number of observed events. Other transitions were derived from other assumptions such as HR-based approaches. This indicates that data from the MK-6482-004 trial (also for the pre-treatment period), and consequently, the survival analyses conducted by the company are subject to great uncertainty. This in turn, points towards a great uncertainty also for the duration of the Belzutifan treatment

effect. For this reason, the EAG scenario analyses varied the duration of the residual benefit also to lower values than the maximum duration of the trial (3.84 years)

4.2.7 Adverse events (AEs)

The main source of evidence on treatment AEs is the MK-6482-004 trial.[24] The CS states that AEs of Grade ≥3 and at a frequency of ≥5% as well as TRAEs of Grade ≥3 and at a frequency of ≥0% are considered in the CE model as they were deemed by the company to impact the total costs of treatment and the patients' QoL.[5] The 5% and 0% threshold for AEs and TRAEs, respectively was set based on the thresholds accepted in previous NICE appraisals.

For the Belzutifan arm, AE rates for patients were sourced from the MK-6482-004 trial, based on the all-subjects-as-treated population. These consisted of anaemia (11.5%) and fatigue (4.9%) at a duration of 7.90 and 2.29 weeks (see Table 71 of the CS).[5] In the SoC arm, the risk of AEs was set to zero, which essentially means the model only used risks of drug-related Grade 3 to 5 AEs for the Belzutifan arm to approximate the incremental AE risks associated with Belzutifan versus SoC.

4.2.8 Health-related quality of life (HRQoL)

In line with NICE's preferences, EQ-5D data was used to inform utilities for the health states and disutilities related to surgical complications, disutilities due to AEs and disutilities due to VHL-associated tumours at non-primary tumour sides. Also, background disutilities due to ageing were considered in the economic analysis, based on Ara and Brazier.[49]

In terms of HRQoL, the health states were split in non-metastatic health states, which included pre-surgery, surgery, and event-free after surgery, and metastatic health states. Utility values were estimated contingent on treatment response for the following health states:

Non-metastatic disease

1. CR
2. PR and SD
3. PD for the VHL RCC and pNET cohorts
4. PD for the VHL CNS Hb cohorts

Metastatic disease

5. Pre-progression
6. Post-progression

In addition, the model includes utilities and (disutilities) related to clinical management of VHL, which is defined as a patient specific and time sensitive balancing trajectory of the following three components: (1) the prevention of metastatic disease originating from RCC or pNET tumours, (2) maintenance of organ function and (3) minimisation of burdensome symptoms, particularly for patients with VHL-associated CNS Hb. The dominant influence on a patient's QoL is assumed to be driven by the "worst" burden they are experiencing. Hence, the model includes utilities (and disutilities) representing each of these three components.

The company highlighted a limitation of EQ-5D data in their ability to capture HRQoL aspects for chronic conditions, like fatigue and impact on relationships and social life.[50] The company argued that this applies to VHL disease due to the frequency of disease monitoring and the impact of this chronic disease on loved ones. The company also stated that the benefit of Belzutifan in reducing the size of tumour manifestations and alleviating fear for patient's loved ones is not adequately captured by the EQ-5D; having no treatment option is significantly more negatively impactful on family mental health

than having a treatment option. Therefore, the impact Belzutifan by filling this unmet need, is not currently valued in the economic model.[5]

EAG comment:

It is not clear to the EAG if and how any disutility calculations in the economic model are influenced by the "worst" burden experienced by patients. From the model we noticed that for short-term (one-time) complications, a total one-time QALY decrement is calculated based on the prevalence of the complications, the duration of the complication and the disutility.

4.2.8.1 HRQoL data identified in the review

The primary sources of HRQoL data were the VHL RW QoL Disease Burden Study and KEYNOTE-564 (in the CR level of the pre-surgery, surgery and event-free after surgery states), since no HRQoL data were collected in the MK-6482-004 trial.[1, 3] The VHL RW QoL Disease Burden Study is a cross-sectional patient survey from 2022 conducted in 220 adult patients with self-reported healthcare professional diagnosed VHL in the US, Canada, the UK, France, and Germany.[1] Twenty-one out of the 220 patients were from the UK. Data collection was conducted by Adelphi Real World through a patient advocacy group, the VHL Alliance who operates in the US, and screened online. However, patients were self-reporting their tumour response status and did not a have a physician-confirmed diagnosis. In line with NICE guidance, the study collected EQ-5D scores. The survey included patients who ranged from relatively well, to severely unwell due to VHL. The mean utility score (EQ-5D Crosswalk from 5L to 3L) was 0.699 (SD: 0.27) using the UK value set. Importantly, the range was 0.240 to 0.988 reflecting the variation in impact VHL can have on patients. Because the survey was not limited to only those patients who require therapy and for whom localised procedures are unsuitable or undesirable, the VHL RW QoL Disease Burden Study is not fully generalisable to the MHRA license population. As a result, the utility estimates from the patients in the VHL RW QoL Disease Burden Study are likely an overestimate of the utilities that would be obtained from the licensed population.

The VHL RW QoL Disease Burden Study could not be used to inform CR utilities because only one patient reported having achieved CR and, therefore, the KEYNOTE-564 study was used for that.[3] Patients in KEYNOTE-564 (phase 3 placebo-controlled clinical trial) had adjuvant treatment pembrolizumab of RCC post-nephrectomy. Their utility values associated with CR were used for all three VHL cohorts since all patients in the MK-6482-004 trial had at least one measurable solid RCC tumour. The European Quality of Life-5 Dimensions-5 Levels (EQ-5D-5L) data from the VHL RW QoL Disease Burden Study and KEYNOTE-564 were mapped onto the UK European Quality of Life-5 Dimensions-3 Levels (EQ-5D-3L) values set as per the NICE reference case.[31] The 3L value set was used to derive utility values for the economic model.

The company decided to use these two sources after conducting a SLR to collect evidence for the humanistic burden of VHL disease, with focus on VHL-associated RCC, CNS Hb and pNET. The search covered the period July 2020 to July 2022 and found eight studies, four of which reported HRQoL data in patients with VHL disease (the other four studies reported QoL in patients with VHL and their partners, caregivers, or other family members). Of the four that reported QoL in patient with VHL disease, two presented HRQoL data in a tumour specific population (VHL CNS Hb),[51, 52] and two in a non-tumour specific VHL population.[53, 54] This means that no HRQoL studies were found for the VHL RCC and VHL pNET populations. In the absence of this information, the company considered the VHL RW QoL Disease Burden Study the most relevant source of HRQoL data for patients with VHL in each health state.

4.2.8.2 Utility in the non-metastatic health states (pre-surgery, surgery and event-free after surgery)

The economic analysis uses response-adjusted utility values for each primary tumour side population in the pre-surgery, surgery and event-free after surgery health states. This was done by applying the distributions of best OR level (i.e., CR, PR/SD, or PD) from the MK-6482-004 trial and VHL Natural History Study for the Belzutifan and SoC arms, respectively, to the utility scores by best response that were calculated based on data from the VHL RW QoL Disease Burden Study. This method is motivated by the assumption that in the pre-surgery, surgery and event-free after surgery health states a better response is assumed to be associated with a higher utility value, because a better response avoids the complications associated with tumour growth and the greater risk of metastases resulting from disease progression.

The utility values were directly estimated from patients who reported having non-metastatic disease in the VHL RW QoL Disease Burden Study and self-reported tumour response status. Table 4.17 describes the self-reported response and numbers after reclassification. This reclassification was done because one patient who reported CR and four patients who reported PR were not currently being treated with any medication for VHL-related cancer. Based on clinical expert feedback, the company considered that a spontaneous reduction in the size of VHL-related tumours was very unlikely and, therefore, these patients were assumed to have SD. Because of this reclassification, no utility could be estimated for CR. For PD and SD, the utility values are pooled. For PD the values were calculated separately. As mentioned above, the utility value for CR was sourced from patients in the KEYNOTE-564 trial who were treated with pembrolizumab in the adjuvant setting. Finally, the utility values estimated for PD in the VHL CNs Hb patients were not considered representative,[51, 52] because the VHL RW QoL Disease Burden Study did not select patients being unsuitable or undesirable for localised procedures. This seemed to be an issue, because the growth of CNS Hb can result in severe neurological disability. For VHL CNS Hb patients with PD who are unsuitable or undesirable for localised procedures, the utility associated with motor neurone disease was considered an appropriate proxy according to clinicians and, therefore, assumed by the company. This value was obtained from a structured interviews-based study by Kiebert et al. (2001) in 77 patients with amyotrophic lateral sclerosis.[2]

Table 4.17: Patient classification table and category after reclassification (last column).

Health state	Self-reported (N)	Included where	Motivation	N after reclassification
Complete response	1	Stable disease	Patient not on treatment; spontaneous reduction of tumour size very unlikely	0
Partial response	8	Stable disease (n=4) Progressed disease (n=4)	Patient not on treatment; spontaneous reduction of tumour size very unlikely.	4
Stable disease	107	Stable disease	Health state matches self-reported category	112 (107 + 1 CR + 4 PR)
Progressive disease	49	Progressive disease	Health state matches self-reported category	49
Based on text from page 193 of the CS.[5]				

Health state	Self-reported (N)	Included where	Motivation	N after reclassification
CR = complete response; CS = company submission; PR = partial response				

4.2.8.3 Utility in the metastatic disease health state

In the post-progression metastatic disease health states, utility values were based on the average EQ-5D-5L utility by self-reported progression status among patients with metastatic disease in the VHL RW QoL Disease Burden Study.[1] In each treatment arm, overall utility in the metastatic disease health state was calculated as a weighted average of the utilities associated with pre-progression and post-progression metastatic disease, based on the estimated proportion of time spent progression-free in the metastatic disease state, as determined by the ratio of PFS to OS in the metastatic disease state (estimated using an NMA as described in Section 4.2.6.3 of this report). The PFS to OS ratio is based on a weighted average of expected PFS and OS for each first-line metastatic disease treatment and the market shares of first-line metastatic disease treatments in each origin tumour. The overall utility in the metastatic disease health state was assumed to be the same in both treatment arms, because in both the Belzutifan and SoC arms patients were expected to receive the same mix of first-line treatments upon developing metastatic disease.

4.2.8.4 Health state utility values (summary)

A summary of all utility values used for the health states in the cost effectiveness analysis, with the justification for selection, is provided in Table 4.18. Given the small number of patients with PR and the high potential for misclassification amongst the PR and SD categories based on patients' responses in the VHL RW QoL Disease Burden Study, according to the company, a utility value pooled across the PD and SD categories was calculated (combined n=116, including four PR and 112 SD after the reclassification). However, this means that most of the subjects in this group had SD and the estimated utility values for the two groups combined is mainly representing the SD group.

Table 4.18: Health state utility values

Health state	Utility value	Reference	Justification
Non-metastatic states – pre-surgery, surgery and event-free after surgery states			
Complete response	*0.868*	KEYNOTE-564 (data cut-off date: 14 dec 2020)[3]	No data in the VHL RW QoL Disease Burden Study. The population in the KEYNOTE-564 was considered representative for the complete response population.
Partial response and stable disease	*0.754*	VHL RW QoL Disease Burden Study[1]	Pooled utility across progressed and stable disease, due to a high potential of misclassification
Progressed disease: VHL RCC and VHL pNET	*0.665*	VHL RW QoL Disease Burden Study[1]	Data available in VHL RW QoL Disease Burden Study appropriate
Progressed disease: VHL CNS Hb	*0.550*	Kiebert et al. 2001[2]	The VHL RW QoL Disease Burden Study did not select for being unsuitable for localised procedures and therefore likely to healthy compared to the population in the model. Patients with motor neurone disease are considered

Health state	Utility value	Reference	Justification
			an appropriate proxy according to clinicians.
Metastatic disease			
Pre-progression metastatic disease	0.525	VHL RW QoL Disease Burden Study[1]	Data available in VHL RW QoL Disease Burden Study appropriate
Post-progression metastatic disease	0.412	VHL RW QoL Disease Burden Study[1]	

Based on Table 74 in the CS.[5]
Note: Given the small number of patients with PR and the high potential for misclassification amongst the PR and Stable Disease categories based on patient responses, a utility value pooled across the PR and Stable Disease categories was calculated.
CS = company submission; CNS Hb = central nervous system hemangioblastoma; CR = complete response; PD = progressive disease; pNET = pancreatic neuroendocrine tumours; PR = partial response; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

In each treatment arm, the utilities in the non-metastatic health states (pre-surgery state and the surgery/event-free after surgery states) were computed as a weighted average of utility values by response level. The utility values for CR, PR, SD and PD as reported in Table 4.18 are weighted according to the ORR (Table 4.19). For the SoC arm, response rates were based on the VHL RW QoL Disease Burden Study (2022).[1] For the Belzutifan arm, the MK-6482-004 trial (cut-off 1 April 2022) was used.[24] Not-evaluable responses were excluded from the distribution, meaning that the not-evaluable category was proportionally redistributed to the other categories. This new distribution was used to weight the utilities, which is in line with the model implementation, as shown in Table 4.20.

Table 4.19: Distribution of objective response level by VHL cohort and treatment arm used to calculate utility values in the pre-surgery, surgery, and event-free after surgery states

Cohort/Treatment arm	Objective response level*				
	Complete response	*Partial response*	*Stable disease*	*Progressive disease*	*Not evaluable***
VHL RCC population					
Belzutifan	6.6%	57.4%	34.4%	0.0%	1.6%
SoC	0.0%	0.0%	57.9%	23.3%	18.9%
VHL CNS Hb					
Belzutifan	8.0%	36.0%	46.0%	6.0%	4.0%
SoC	0.0%	0.0%	57.9%	23.3%	18.9%
VHL pNET					
Belzutifan	31.8%	59.1%	9.1%	0.0%	0.0%
SoC	0.0%	0.0%	57.9%	23.3%	18.9%

Based on Table 75 CS.[5] For the Belzutifan arm, data comes from the MK-6482-004 trial.[24]
*For SoC, patients' distribution across response categories was approximated based on self-reported response status among patients in the VHL RW QoL Disease Burden Study (2022) who reported receiving no prescribed medication for VHL-related cancer (N=159). Untreated patients in the QoL study who reported complete response (1 patient) or partial response (4 patients) were assumed to have stable disease, based on clinical expert feedback that a spontaneous reduction in the size of VHL-related tumours is very unlikely in the absence of treatment.

Cohort/Treatment arm	Objective response level*				
	Complete response	*Partial response*	*Stable disease*	*Progressive disease*	*Not evaluable***
**When calculating the weighted average of utility in each non-metastatic health state, patients in the "not evaluable" category are proportionally redistributed to the other categories CS = company submission; CNS Hb = central nervous system hemangioblastoma; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau					

Table 4.20: Response-adjusted overall utility in non-metastatic health states (pre-surgery, surgery, and event-free after surgery)

Health state	Utility value Belzutifan	Utility value SoC
VHL RCC	0.762	0.728
VHL CNS Hb	0.751	0.695
VHL pNET	0.790	0.728
Based on Table 76 and 77 in the CS.[5] CS = company submission; CNS Hb = central nervous system hemangioblastoma; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau		

4.8.2.5 Disutility values

The model incorporates disutility due to surgeries and surgical complications for VHL-associated tumours. The disutility associated with the perioperative recovery from VHL pNET surgery is considered along with the risks of both short-term and long-term surgical complications for each VHL-associated surgery type and this utility is derived from a CE study comparing laparoscopic versus open distal pancreatectomy for pancreatic cancer.[55] Both the short and long-term complication were derived from real-world data from the Optum Clinformatics Data Mart Claims Study.[33] Short-term complications were measured over a 28-day period following each surgery, while long-term complications were measured over a 180-day period. The following long-term complications were considered: chronic pain, cerebral vasculature occlusion/stroke, seizure, neurological complications, and secondary diabetes or exocrine pancreatic insufficiency).

The perioperative recovery after a VHL-associated tumour procedures is assumed to have a significant disutility. These procedures operate on a primary tumour and are seen as a "last-resort" surgery, therefore, patients are expected to experience a period of disutility. For the VHL pNET associated surgeries, a disutility of -0.186 is assumed and applied for a 28-day period to reflect the perioperative recovery period. This disutility was sourced from a CE study assessing pancreatectomy for pancreatic cancer and was calculated by subtracting the subsequent stable period following distal pancreatectomy from the utility associated with complicated open distal pancreatectomy in the first 3 months.[55] For VHL RCC and VHL CNS Hb related-surgeries, no disutility associated with the perioperative recovery was assumed due to a lack of available data for this input. The company argue that this is a conservative assumption, as it will make the incremental benefit of Belzutifan compared to SoC smaller than what would be observed in reality.

The disutility of each short-term surgical complication and the surgery itself is applied to the 28-day period following the surgery, in accordance with the timeframe in which the risks of these complications were measured. As described above, long-term complications were measured over a 180-day period. The disutility of long-term complications is applied to the proportion of patients who experienced each complication in the Optum Clinformatics Data Mart Claims Study in all cycles starting from the first surgery until death or the end of the modelled time horizon.

A summary of these disutility values applied to long-term and short-term surgical complications in the base-case can be found in Table 4.21 and Table 4.22. respectively. For the long-term complication, disutilities were derived from the VHL RW QoL Disease Burden Study where available.[1] They were based on the difference in average utility between patient with versus without specific co-morbidities. For other, less-common long-term complications that were not assessed in the VHL RW QoL Disease Burden Study, disutilities were obtained from published literature sources.

Table 4.21: Long-term surgical complication disutility values in the base-case

Complication	Disutility	Source
VHL-associated RCC surgery		
End stage renal disease and/or dialysis*	-0.527	Lee et al. 2005 (weighted average of haemodialysis and peritoneal dialysis disutilities)[56]
Chronic kidney disease*	-0.136	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic kidney disease)[1]
Hernia surgery	-0.200	Simianu et al. 2020 (difference in utility with hernia complication vs. baseline state)[57]
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain)[1]
Cerebral vasculature occlusion or stroke	-0.370	Gandhi et al. 2012 (non-fatal stroke disutility)[58]
VHL-associated CNS Hb surgery		
Chronic pain (in CNS Hb population)	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain)[1]
Cerebral vasculature occlusion or stroke	-0.370	Gandhi et al. 2012 (non-fatal stroke disutility)[58]
Seizure	-0.270	Assumed equal to neurological complications
Neurological complications	-0.270	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without motor loss/ataxia)[1]
VHL-associated pNET surgery		
Chronic pain (in pNET population)	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain)[1]
Secondary diabetes or exocrine pancreatic insufficiency*	-0.042	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without diabetes)[1]
Immunocompromisation*	-0.081	NICE Committee Papers for GID-TA10024 (everolimus in neuroendocrine tumours), based on difference between the mean utility values for (stable disease without AE) minus (stable disease with leukopenia AE)[59]
Disutility due to surgery for other VHL-associated manifestations		
Complications of adrenal lesion surgery		
Adrenal insufficiency	-0.042	Assumed equal to diabetes
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain)[1]
Complications of retinal Hb surgery		
Chronic pain	-0.195	VHL RW QoL Disease Burden Study (2022) (difference in utility with vs. without chronic pain)[1]

Complication	Disutility	Source
Based on Table 77 CS.[5] *This is a metabolic complication resulting from limited/absent organ function following last-resort surgery CNS = central nervous system; Hb = hemangioblastoma; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau		

Table 4.22: Short-term surgical complication disutility values in the base-case

Complication	Disutility	Source
VHL-associated RCC surgery		
Acute renal failure	-0.150	Nisula et al. 2013,[60] as cited in Zargar et al. 2018 (acute kidney injury disutility)[60, 61]
Cardiac complications	-0.240	Gandhi et al. (2012) (non-fatal myocardial infarction disutility)[58]
Erythroderma	-0.335	Poole et al. 2010 (severe atopic dermatitis)[62]
Kidney infection	-0.340	Stevenson et al. 2014 (kidney infection disutility)[63]
Other genitourinary complications	-0.255	Stevenson et al. 2014 (simple average of urinary obstruction and incontinence disutilities)[63]
Postoperative infection (RCC-related)	-0.360	Stevenson et al. 2014 (abscess disutility)[63]
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility)[64]
Thrombosis and/or embolism	-0.330	Stevenson et al. 2014 (deep vein thrombosis disutility)[63]
Vascular injury or anaemia	-0.073	Nafees et al. 2008 (approximated by disutility of fatigue)[65]
VHL-associated CNS Hb surgery		
Acute renal failure	-0.150	Nisula et al. 2013,[60] as cited in Zargar et al. 2018 (acute kidney injury disutility)[60, 61]
CNS hemorrhage	-0.240	Wang et al. 2020 (minor intracranial hemorrhage)[66]
Nerve palsy related to anesthesia	-0.120	Memeh et al. 2020 (temporary unilateral laryngeal nerve injury)[67]
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility)[64]
Thrombosis and/or embolism	-0.330	Stevenson et al. 2014 (deep vein thrombosis disutility)[63]
Vascular injury or anaemia	-0.073	Nafees et al. 2008 (approximated by disutility of fatigue)[65]
VHL-associated pNET surgery		
Abdominal abscess	-0.360	Stevenson et al. 2014 (abscess disutility)[63]
Postoperative infection (pNET-related)	-0.360	Stevenson et al. 2014 (abscess disutility)[63]
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility)[64]
Thrombosis and/or embolism	-0.330	Stevenson et al. 2014 (deep vein thrombosis disutility)[63]
Urinary tract infection	-0.270	Stevenson et al. 2014 (urinary tract infection)[63]
Vascular injury or anaemia	-0.073	Nafees et al. 2008 (approximated by disutility of fatigue)[65]

Complication	Disutility	Source
Perioperative recovery after pNET surgery	-0.186	Gurusamy et al. 2017 (utility of complicated open distal pancreatectomy in first 3 months minus utility of subsequent stable period)[55]
Disutility due to surgery for other VHL-associated manifestations		
Complications of adrenal lesion surgery		
Acute renal failure	-0.150	Nisula et al. 2013,[60] as cited in Zargar et al. 2018 (acute kidney injury disutility)[60, 61]
Respiratory complications	-0.250	Sankar et al. 2020 (pulmonary embolus disutility)[64]
Thrombosis or embolism	-0.330	Stevenson et al. 2014 (deep vein thrombosis disutility)[63]
Vascular injury or anaemia	-0.073	Nafees et al. 2008 (approximated by disutility of fatigue)[65]
Complications of endolymphatic sac tumour surgery		
Acoustic impairment	-0.150	Verkleij et al. 2021 (moderate unilateral hearing loss)[68]
Complications of retinal Hb surgery		
Vitreous haemorrhage	-0.223	Assumed equal to vision loss disutility derived from Ament et al. 2018[a] (neurological complication: visual loss at 2 months)

Based on Table 78 CS.[5]
[a]The CS did not provide any further details of this reference.
CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

The cross-sectional VHL RW QoL Disease Burden Study could not be used of short-term complications, as short-term complication disutilities need to be measured while a patient is actively experiencing the complication. Therefore, all disutilities of short-term surgical complications were obtained from published sources. The incremental disutility resulting from long- and short-term complications of surgery for other VHL tumour was also determined by the incidence rate of surgeries and surgical complications for these tumours per cycle and sourced using the same methods described above.

As mentioned in Section 4.2.7, the model includes anaemia and fatigue as treatment related AEs for the Belzutifan arm. They are included in the model by applying a lump-sum QALY decrement at model entry. This QALY decrement was calculated as a function of the treatment specific AE risk, the mean duration of these AEs and the associated disutility. The disutilities associated with these AEs, as used in the base-case analysis, are presented in Table 4.23. These disutility values were obtained from an analysis of EQ-5D-5L data from KEYNOTE-564 that was previously conducted for NICE TA830.[35] and represents the difference in utility associated with disease-free without toxicity versus disease-free during any Grade 3+ AE.

Additionally, a background disutility related to aging of the cohort over time is applied within the model. This disutility values are based on the published linear regression model from Ara and Brazier 2010.[49] This model predicts mean utility values for individuals within the general population, conditional on age (in years), age-squared, and gender. The regression coefficients used for linear regression model are presented in Table 4.24.

Table 4.23: Summary of AE disutility values in the base-case

Adverse event	Disutility	Source
Anaemia	-0.06417	KEYNOTE-564 (data cut-off date: 14 Dec 2020)[35]
Fatigue	-0.06417	
Based on Table 79 CS.[5] CS = company submission		

Table 4.24: Regression coefficients used for the estimation of age-related disutility

Parameter	Coefficient	Source
Age (years)	-0.0002587	Ara and Brazier 2010[49]
Age 2	-0.0000332	
Male	0.0212126	
Intercept	0.9508566	
Based on Table 80 CS.[5] CS = company submission		

The model also includes disutilities associated to the burden experienced by caregivers. VHL disease is a severe condition with a profound impact on the health status and well-being of patients' caregivers (term used for caregivers and close family member). Caregivers are likely to experience anxiety for the patient, fear of tumour recurrence and bereavement in the event of the patient's premature death. In addition, over the patient's lifetime, caregivers are likely to carry responsibilities such as providing physical care and emotional support to the patient, scheduling and coordinating healthcare services, and managing disease-related finances. Therefore, the disutility of caregivers was considered in scenario analysis by the company. Caregiver disutility is modelled by patients' health state distribution in each cycle, based on published studies conducted among family members and caregivers of cancer patients, and applied to all cohorts. A caregiver bereavement disutility is applied as a one-time QALY decrement upon patient death. In the scenario analysis, the caregiver disutility value was assumed to conservatively apply to one caregiver only, despite patients potentially requiring more than one caregiver. In relatively severe health states before the patient's death the caregiver disutility may be conceivably worse than the disutility associated with bereavement. This is because while the patient is alive, the disutility includes both the disutility due to emotional distress/worry over the patient's condition, as well as the disutility due to the burden of providing care for the patient. The caregiver disutility values are summarised in Table 4.25. These disutility values were identified in a targeted review of published literature sources. The review identified no publications that examined caregiver HRQoL impairment for VHL RCC, VHL pNET or VHL CNS Hb. Therefore, estimates from studies in other oncology indications were used as proxies. Caring for a patient with severe neurological disability due to CNS Hb tumour growth is expected to be particularly burdensome. For caregivers of these patients, a disutility is using multiple sclerosis as a proxy is applied. This disutility is taken from the NICE TA for ocrelizumab for treating relapsing–remitting multiple sclerosis (TA533) taking the disutility associated with an Expanded Disability Status Scale (EDSS) score of 9.[69]

Table 4.25: Summary of caregiver disutility values in the base-case

Health state/response status	Disutility	Source
Pre-surgery	-0.030	Turner et al. 2013 [breast, colorectal, or prostate cancer survivor] [70]
Surgery and event-free after surgery	-0.065	Based on the midpoint between caregiver disutility in the pre-surgery state versus the metastatic disease state.
Metastatic disease	-0.100	Sjolander et al. 2012[a] [lung or GI cancer][71]

Health state/response status	Disutility	Source
After death of patient	-0.050	Song et al. 2012 [terminal cancer][72]
PD patients in the VHL CNS Hb cohort	-0.140	Ocrelizumab for treating relapsing–remitting multiple sclerosis [TA533] (caregiver disutility associated with EDSS score of 9)[69]

Based on Table 81 in the CS.[5]
[a]The bibliographic details for this reference were not provided in the CS[5] and the paper was not included within the reference pack. The cited reference is the suggestion of the EAG but has not been confirmed as correct.
CNS = central nervous system; CS = company submission; EAG = Evidence Assessment Group; EDSS = expanded disability status scale; GI = Gastrointestinal; Hb = haemangioblastoma; NSCLC = non-small-cell lung cancer; PD = progressive disease; VHL = Von Hippel-Lindau

EAG comment:

The main concerns of the EAG regarding the implementation of HRQoL in the model are the following:

a) The company have used response-adjusted health state utilities computed as a weighted average of utility values by response level observed in each treatment arm. The company explained that this is justified by the assumption that in the pre-surgery, surgery and event-free after surgery health states, a better response is assumed to be associated with a higher utility value. While the EAG considers this statement reasonable, its main concern relates to how this assumption has been operationalised in the economic model since as it is currently implemented in the model, patients in the Belzutifan arm obtain an immediate benefit in HRQoL compared to SoC from the first model cycle. This seems unrealistic. However, on p.123 of the CS (treatment decision point), it is mentioned that this "*patient population have exhausted alternative options to control VHL tumour manifestations and are at the "end of the road", they must have sufficient organ function as per the inclusion criteria for the trial*".[5] The company explained then that in clinical practice, a patient meeting the eligibility criteria for Belzutifan would have three options: 1) surgery, which based on the MHRA/DP eligibility criteria is unsuitable or undesirable for those patients because it will result in loss of organ function, 2) active surveillance, which is meant to monitor tumours that are larger than 3 cm (RCC) or 2 cm (pNETs) and, therefore, these patients are at an increased risk of experiencing metastatic disease and/or other symptoms of tumour burden (particularly in CNS Hb tumours), or 3) Belzutifan. Patients at this stage would have been monitored probably over many years and any treatment decision needs to be carefully made between patients and clinicians. For some patients this decision may not be immediate, however, in the economic model, it was assumed that at the treatment decision point the patient chooses between receiving immediately surgery, routine surveillance, or Belzutifan. The EAG considers this simplifying assumption appropriate given the lack of evidence, but it is not clear why patients in the Belzutifan arm start benefiting from its treatment right from the beginning. The EAG understands that at the treatment decision point patients are in a poor health condition, including those who receive Belzutifan. It seems therefore unlikely that before Belzutifan starts showing some effect, patients could experience any type of benefit. The EAG considers that this issue might have been resolved by including time to treatment response in the model and by linking the objective response level to time to response to calculate utility values in the pre-surgery, surgery, and event-free after surgery states. The EAG explored a scenario in which a fixed cut-off at the median time to treatment response was included to the QALY calculation (please refer to Scenario analyses set 6 in Section 6.1.2.6). The company also explored the impact of using fixed proportions at each response level in response to clarification question B21c.[4] However, these scenarios were only exploratory and their results should be

interpreted with caution. With the available data, the EAG was unable to remove Belzutifan's immediate effect on the response level which was used to estimate the weighted utility values as used in the economic model.

b) The company have acknowledged that the licensed population (the population in the DP) may have worse utility scores than those used in the model because the survey in the VHL RW QoL Disease Burden Study was not limited to only those patients who require therapy and for whom localised procedures are unsuitable or undesirable. Thus, the VHL RW QoL Disease Burden Study is not fully generalisable to the MHRA license population. As a result, the utility estimates from the patients in the VHL RW QoL Disease Burden Study are likely an overestimate of the utilities that would be obtained from the licensed population. The EAG wonders whether the company could have "adjusted" these utilities to better reflect the relevant patient population as they did with other parameters such as risks of surgery-related complications. However, even if the licensed population is expected to have worse utility scores, it does not mean that the effect of Belzutifan on HRQoL is underestimated in the model as the company claims. It could be the other way around, even if it is in an indirect way, e.g., some model assumptions could lead to an overestimation of Belzutifan HRQoL, for example when in the model it is assumed an immediate HRQoL effect associated to Belzutifan as explained above.
c) The distribution of objective response level by VHL cohort and treatment arm used to calculate utility values in the pre-surgery, surgery, and event-free after surgery states shown in Table 4.20 is applied in the model from the first cycle, resulting in an immediate HRQoL benefit for Belzutifan patients. This also implies that at model start, patients are not equal in the Belzutifan and the SoC arms, and based on Table 4.20, SoC patients in the model are more severe than patients in Belzutifan.

4.2.8.6 Benefits not captured in the QALY calculation

Besides reduced well-being of patients and direct costs to the NHS though increase resource use, the interventions used to treat VHL disease also cause a greater burden on social care services, family members, and other caregivers who are required to spend time and resources assisting patients as they lose independence due to VHL. Belzutifan is intended to slow down disease progression and prevent the requirement for severely detrimental surgeries and, therefore, has the potential to provide further benefits to patients, carers, and wider society that are not captured in the cost per QALY calculations.

According to the company, the value of Belzutifan that is not currently captured, or not sufficiently captured, in the QALY calculation includes the following:

1. The value of reducing symptoms across multiple systems. The model focuses on the 'worst' primary tumour. For the most adversely impacted patients, significant benefit will be gained by both preserving organ function and preventing tumours' advance at other tumour sites. This provides further relief for patients in knowing that their tumour manifestations as a whole are being controlled rather than requiring multiple different surgeries for different tumour sites.
2. The VHL RW QoL Disease Burden Study did not report utility scores for the types of interventions likely to be received by patients described in the MHRA indication.[1] The company believes this underestimate the negative impact on HRQoL for the target patient population.
3. Overall survival in the non-metastatic health state is based on either background mortality or VHL Natural History Study mortality, which reflects the minimum mortality risk of the general population. This is likely to substantially overestimate life expectancy in the patient population described in the MHRA indication, therefore underestimating the potential value of Belzutifan.
4. The company consider that the EQ-5D is not a sensitive enough tool to accurately reflect the mental health impact of this disease. The requirement of patients to undergo repeated surgical procedures,

which in themselves result in deterioration in health, can often cause patients to lose hope for the future and reduce their ability to take part in their normal daily activities. The two deaths, one suicide and one fentanyl overdose death, in a small patient population is a warning sign. These deaths were not attributed to Belzutifan, but the result of living with VHL. By delaying and slowing the need for surgeries, Belzutifan is therefore expected to improve the mental well-being (including a reduction in patient anxiety associated with scans) and offer optimism and much-needed hope to patients, which may not be captured in the stringent framework of the EQ-5D questionnaire and QALY measures.

5. Reduced anxiety associated with frequent scans, fear of disability or death from surgery, and relief in knowing tumour manifestations as a whole are controlled in comparison to multiple different surgeries for different tumour sites.
6. Not having a treatment option is significantly more negatively impactful on family mental health than having a treatment option. This negative impact of not having a treatment option is not currently valued in the economic model.
7. Increase work productivity for both patients and caregivers. It is expected that patients may increase their work productivity and might be able to work longer in life, due to the reduced disease burden of the surgical procedures when treated with Belzutifan and the delay of the most severe complication. The work productivity of caregivers of patients with VHL is also expected to improve due to the impact of Belzutifan on delaying and/or reducing surgeries over a patient's lifetime, and these benefits may spill over to other government bodies. Costs not reimbursed by the NHS, such as the costs of travelling to medical appointments, are likely to be reduced for patients treated with Belzutifan due to reduced need for surgical or other localised procedures, and cost savings would also accrue from reduced need for medical treatment for the complications of such procedures.

EAG comment:

While the EAG may agree with some of the points raised by the company above, it should be noted that none of these were supported by additional evidence. The EAG also wonders whether these points could have been somehow captured in the model by running additional scenario analyses. Furthermore, as mentioned above, even if some benefits are not currently captured in the QALY calculation, it does not necessarily mean that the effect of Belzutifan on HRQoL is underestimated in the model.

4.2.9 Resources and costs

The following cost categories were included in the analysis: drug acquisition and administration costs for Belzutifan and metastatic disease therapies, health states costs, costs associated to surgery and its complications, costs for the treatment of AEs, and other (miscellaneous) costs. Unit prices were based on the NHS Reference Costs, British National Formulary (BNF) and Personal Social Services Research Unit (PSSRU).[73-75]

4.2.9.1 Resource use and costs data identified in the review

According to the CS, a SLR was conducted in July 2020, and updated in July 2022 to identify cost and resource use data for VHL-associated RCC, pNET, and CNS Hb tumours for use in the economic model. The SLR identified two relevant publications, namely Jonasch et al. 2022 and Sundaram et al. 2022.[76, 77]

4.2.9.2 Belzutifan acquisition and administration costs

The drug acquisition costs for Belzutifan at list price are £11,936.70 for a pack of 90 oral 40 mg tablets. Also, an oral drug dispensing cost of £245.23 once every 4 weeks (i.e., assuming a 4-week fill at each dispensing), sourced from NHS 2020/21 Reference Costs,[73] is assumed for Belzutifan. The SoC is

defined as established clinical management without Belzutifan, which is assumed to be surgery for most of patients. Therefore, no drug administration or acquisition costs are included in the SoC arm. Belzutifan acquisition and administration costs are summarised in Table 4.26.

Belzutifan dosing schedule is consistent with the treatment protocol used in the MK-6482-004 trial and the MHRA license and consists of three (40 mg) tablets daily (i.e., daily dose of 120 mg).[6, 24] No vial sharing was assumed in the base-case, with vial sharing assumed for all treatments administered by intravenous (IV) infusion in scenario analyses.

In the company's base-case, the mean relative dose intensity (RDI) observed in the MK-6482-004 trial ([redacted]%) was applied to Belzutifan acquisition cost per 90-tablet pack to account for delays or interruptions in treatment.

Table 4.26: Belzutifan acquisition and administration costs per 28-day cycle

Item	Cost	Source
Belzutifan 40 mg – 90 tablet pack (list price)*	£11,936.70	BNF[74]
Administration cost – oral drug dispensing (per pharmacy dispensing)	£10.80	Band 6 Hospital Pharmacist based on 12 minutes of time for each dispensing, PSSRU 2021[75]

Based on Table 82 of CS.[5]
* Belzutifan administered at a dose of 120 mg daily.
BNF = British National Formulary; CS = company submission; mg = milligrams; PSSRU = Personal Social Services Research Unit

4.2.9.3 Metastatic disease therapies acquisition and administration costs

The company also considered drug acquisition and administration cost associated with metastatic disease therapies (for both first-line and second-line options). The company assumed that most patients entering the metastatic disease state received an active first-line treatment for advanced RCC or pNET. A subset of these patients is assumed to receive no active metastatic disease treatment, since in practice not all patients with metastatic disease receive active treatment. Since no cases of metastatic disease originated from CNS Hb patients were observed in the VHL Natural History Study, the company assumed that no patients received metastatic therapy for CNS Hb. Costs associated with second and further lines of metastatic therapies for patients progressing are also included in the model.

The first-line metastatic therapies included in the economic analysis were based on those recommended by NICE or listed as a preferred or recommended first-line regimen according to the National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines) for the treatment of advanced RCC and pNET. The company stated that this is in line with previous NICE appraisals such as pembrolizumab for adjuvant treatment of RCC (TA830) and pembrolizumab for adjuvant treatment of completely resected stage 3 melanoma (TA766).[37, 78]

For both first- and second-line therapies, market shares of metastatic therapies are assumed to be equal for Belzutifan and SoC. Drug acquisition costs for treating advanced RCC and pNET are then calculated in the model as a function of the unit drug cost, Patient Access Scheme (PAS) discount (if these are not non-confidential), defined dosing schedule, and RDI. For treatments with a confidential PAS discount, list prices are assumed. Unit drug costs per vial or capsule are sourced from BNF for branded agents, and the electronic market information tool (eMIT) for generic agents.[74, 79] These are summarised in Table 4.27.

Table 4.27: Unit drug costs for first- and second-line therapies for advanced RCC and pNET

Drug	Strength per unit (mg or MU)	Cost per unit (£)
Pembrolizumab	100	2,630.00
Sunitinib* 12.5 mg	12.5	28.03
Sunitinib* 37.5 mg	37.5	84.00
Sunitinib* 50 mg	50	112.10
Axitinib	5	62.80
Tivozanib	1.34	97.71
Pazopanib**	400	37.37
Cabozantinib	60	171.43
Nivolumab	40	439.00
Ipilimumab	50	3,750.00
Avelumab	200	768.00
Lenvatinib	10	47.90
Everolimus 5 mg	5	75.00
Everolimus 10 mg	10	89.10
Temsirolimus	30	620.00
Interferon a2B	25	103.94
Cisplatin	50	6.03
Etoposide	100	3.84
Irinotecan	500	15.51
Leucovorin	350	5.50
Oxaliplatin	100	7.28
Streptozocin	1000	570.00
5-fluorouracil	2500	4.21
Doxorubicin	200	20.02
Temozolomide	180	3.47
Capecitabine	300	0.13
Lanreotide	120	937.00
Octreotide	30	656.88

Based on Table 83 of CS.[5]
* Sunitinib price based on PAS discount for this therapy: first treatment cycle of sunitinib is free to the NHS.[80]
** Pazopanib price based on 12.5% PAS discount.[81]
CS = company submission; mg = milligrams; MU = million units; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma

For all first- and second-line treatments of advanced RCC, the mean RDI was applied to the drug acquisition costs. The company indicated that these were obtained from pivotal clinical trials and HTA appraisals in advanced RCC settings. For advanced pNET therapies dosing schedules were obtained from prescribing information, trial publications, and clinical expert opinion. The RDI of advanced pNET treatments was assumed to be 100%. For all IV drugs where dosage is calculated based on weight or body surface area (BSA), the company base-case assumed that vial-sharing was allowed. The number of vials required per infusion was estimated based on the mean body weight or mean BSA of patients

in the MK-6482-004 trial (e.g., average patient weight in kilogrammes (kg) multiplied by the required dose per kg – milligrams (mg)/kg - divided by the strength per vial - mg/vial, with the vial strength associated with the lowest cost per mg). Under the assumption that vial-sharing is not allowed scenario, the number of vials required per infusion was estimated based on a log-normal distribution of patient weight or BSA, using the mean and standard deviation from the MK-6482-004 trial. The proportion of patients requiring different numbers of vials was calculated as the estimated percentage of patients falling into the corresponding weight or BSA interval. The modelled dosing schedules and RDI for advanced RCC and advanced pNET therapies are summarised in Table 4.28 and 4.29, respectively.

Table 4.28: Dosing schedules and RDI for first- and second-line therapies for advanced RCC

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
First line					
Sunitinib	Sunitinib	50 mg QD orally for 4 weeks, then 2 weeks off treatment	No max	74.7%	KEYNOTE-426 (data cut-off date: 24 Aug 2018)[82]
Tivozanib	Tivozanib	1.34 mg QD orally for 3 weeks followed by 1 week without treatment	No max	94.0%	NICE TA512[83]
Pazopanib	Pazopanib	800 mg QD orally	No max	86.0%	NICE TA215[81]
Cabozantinib	Cabozantinib	20/40/60 mg QD orally	No max	94.3%	NICE TA542[84]
Nivolumab/ipilimumab	Nivolumab (in combination)	3 mg/kg IV Q3W for up to 4 doses	12	94.8%	Equal to pembrolizumab
	Ipilimumab	1 mg/kg IV Q3W for up to 4 doses	12	94.8%	
	Nivolumab (maintenance)	480 mg IV Q4W starting 6 weeks after the last combination dose	No max	94.8%	
Avelumab/axitinib	Avelumab	800 mg Q2W	No max	91.5%	Motzer et al. 2019 [JAVELIN Renal 101][85]
	Axitinib	5 mg BID orally	No max	89.4%	
Pembrolizumab/lenvatinib	Pembrolizumab	200 mg IV Q3W	104	94.8%	Assumed equal to that of pembrolizumab
	Lenvatinib	20 mg orally QD	No max	69.9%	Motzer et al. 2021 [KEYNOTE-581][39]
Second line					
Nivolumab	Nivolumab	480 mg IV Q4W or 240 mg IV Q2W	No max	92.0%	NICE TA417[86]
Pembrolizumab	Pembrolizumab	200 mg IV Q3W	No max	94.8%	Assume same as in first line

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
Axitinib	Axitinib	5 mg orally BID	No max	102.0%	NICE TA333/TA417[86, 87]
Cabozantinib	Cabozantinib	60 mg orally QD	No max	100.0%	NICE TA463[88]
Lenvatinib/everolimus	Lenvatinib	18 mg orally QD	No max	75.0%	Motzer et al. 2015 [NCT01136733][89]
	Everolimus	5 mg orally QD	No max	85.0%	
Pazopanib	Pazopanib	800 mg orally QD	No max	86.0%	Assume same as in first line
Sunitinib	Sunitinib	50 mg orally QD for 4 weeks, then 2 weeks off treatment	No max	74.7%	Assume same as in first line
Everolimus	Everolimus	10 mg orally QD	No max	91.8%	NICE TA219/TA432[90, 91]
Temsirolimus	Temsirolimus	25 mg IV QW	No max	92.4%	Hudes et al. 2007 [NCT00065468][92]
Cytokines (interferon)	Interferon a2B	10 MU SC three days per week	No max	100.0%	Assumption

Based on Table 84 of CS.[5]
BID = twice a day; CS = company submission; IV = intravenous; MU = million units; PAS = Patient Access Scheme; Q#W = once every # weeks; QD = once a day, QW = once weekly; RCC = renal cell carcinoma; SC = subcutaneous; ToT = time on treatment

Table 4.29: Dosing schedules for first- and second- line therapies for advanced pNET

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
First line					
Streptozocin/5-fluorouracil	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Streptozocin	Sun et al. 2005[40]
	5-fluorouracil	400 mg/m2 IV on days 1 to 5 and days 36 to 40 Q10W	No max	5-fluorouracil	
Streptozocin/doxorubicin	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Streptozocin	Sun et al. 2005[40]
	Doxorubicin	40 mg/m2 IV Q5W	No max	Doxorubicin	
Temozolomide/capecitabine	Temozolomide	200 mg/m2 orally daily for 5 days Q4W	No max	Temozolomide	Strosberg et al. 2011[93]
	Capecitabine	750 mg/m2 orally twice daily for 14 days Q4W	No max	Capecitabine	
Everolimus	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, Afinitor (Everolimus)	Everolimus
Sunitinib	Sunitinib 12.5 mg	37.5 mg orally QD	No max	Prescribing information, Sutent (sunitinib)	Sunitinib
Interferon a2B	Interferon a2B	5 MU SC three days per week	No max	Faiss et al. (2003) (105)	Interferon a2B
Lanreotide	Lanreotide	120 mg SC Q4W	No max	Prescribing information, Lanreotide	Lanreotide
Octreotide	Octreotide	20 mg SC Q4W	No max	Clinical expert input	Octreotide
Second line					
Cisplatin/etoposide	Cisplatin	80 mg/m2 IV Q3W for up to 6 cycles	18	Cisplatin	Iwasa et al. 2010[94]

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
	Etoposide	100 mg/m2 IV on days 1-3 Q3W for up to 6 cycles	18	Etoposide	
Everolimus	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, afinitor (Everolimus)	Everolimus
FOLFIRI	5-fluorouracil	400 mg/m2 IV (in a 10-min bolus) + 1,200 mg/m2 (in a 44-h infusion) Q2W	No max	5-fluorouracil	Hentic et al. 2012[95]
	Irinotecan	180mg/m2 (on day 1) Q2W	No max	Irinotecan	
	Leucovorin	400mg/m2 (in a 2-h infusion) Q2W	No max	Leucovorin	
FOLFOX	5-fluorouracil	400 mg/m2 (bolus) + 2,400 mg/m2 (as a 46-h continuous infusion) Q2W	No max	5-fluorouracil	Faure et al. 2017[96]
	Oxaliplatin	85 mg/m2 IV infusion (over 120 minutes) Q2W	No max	Oxaliplatin	
	Leucovorin	100 mg/m2 IV infusion (over 120 minutes on day 1) Q2W	No max	Leucovorin	
Streptozocin/5-fluorouracil	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Streptozocin	Sun et al. 2005[40]
	5-fluorouracil	400 mg/m2 IV on days 1 to 5 and days 36 to 40 Q10W	No max	5-fluorouracil	
Streptozocin/doxorubicin	Streptozocin	500 mg/m2 IV on days 1 to 5 Q10W	No max	Streptozocin	Sun et al. 2005[40]
	Doxorubicin	40 mg/m2 IV Q5W	No max	Doxorubicin	

Regimen	Drug component	Dosing schedule	Maximum ToT (weeks)	RDI (%)	Sources for RDI
Sunitinib	Sunitinib 12.5 mg	37.5 mg orally QD	No max	Prescribing information, sutent (sunitinib)	Sunitinib
Temozolomide/capecitabine	Temozolomide	200 mg/m2 orally daily for 5 days Q4W	No max	Temozolomide	Strosberg et al. 2011[93]
	Capecitabine	750 mg/m2 orally twice daily for 14 days Q4W	No max	Capecitabine	
Interferon a2B	Interferon a2B	5 MU SC three days per week	No max	Faiss et al. 2003 (105)	Interferon a2B
Lanreotide/everolimus	Lanreotide	120 mg SC Q4W	No max	Prescribing information, lanreotide	Lanreotide
	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, afinitor (Everolimus)	Everolimus 10 mg
Octreotide/everolimus	Octreotide	20 mg SC Q4W	No max	Clinical expert input	Octreotide
	Everolimus 10 mg	10 mg orally QD	No max	Prescribing information, afinitor (Everolimus)	Everolimus 10 mg
Lanreotide	Lanreotide	120 mg SC Q4W	No max	Prescribing information, lanreotide	Lanreotide
Octreotide	Octreotide	20 mg SC Q4W	No max	Clinical expert input	Octreotide

Based on Table 85 of CS.[5]

BID = twice a day; CS = company submission; IV = intravenous; MU = million units; PAS = Patient Access Scheme; pNET = pancreatic neuroendocrine tumour; Q#W = once every # weeks; QD = once a day, QW = once weekly; SC = subcutaneous; ToT = time on treatment

Intravenous and oral drug administration unit costs were sourced from the 2020/2021 NHS Reference Costs[73] and they are summarised in Table 4.30. For each treatment in the advanced RCC and pNET settings, the following assumptions were made:

- Intravenous and subcutaneous (SC) single-agent regimens: unit costs per infusion based on SB12Z (simple parenteral chemotherapy).[73]
- Intravenous combination regimens that either do not contain cisplatin or do not require multiple infusions per chemotherapy cycle: unit costs per chemotherapy cycle (covered all drug components) based on SB13Z (complex parenteral chemotherapy).[73]
- Intravenous combination regimens in which one or more drug components is administered more than once per chemotherapy cycle: unit costs per chemotherapy cycle based on the sum of SB13Z (complex parenteral chemotherapy) and SB15Z (subsequent elements of a chemotherapy cycle).[73] For the combination of cisplatin with etoposide: unit costs per chemotherapy cycle based on the sum of SB14Z (complex parenteral chemotherapy, including prolonged infusion) and SB15Z (subsequent elements of a chemotherapy cycle), given the prolonged 6- to 8-hour infusion time required for cisplatin and the multiple days of administration required for etoposide per 3-week chemotherapy cycle.[73]
- Orally administered single-agent or combination regimens: assumed to require one oral drug dispensing cost based on SB11Z (deliver exclusively oral chemotherapy) once every 4 weeks (or once every 6 weeks for sunitinib in the advanced RCC setting).[73]
- Combination regimens including both orally administered and IV-administered drug components: administration costs associated with the oral component assumed to be covered by the IV administration costs.

Table 4.30: Unit costs of drug administration in the advanced RCC setting

Route	Type of administration	Unit cost per administration or pharmacy dispensing (£)	Source
IV or SC	Simple parenteral chemotherapy	361.53	NHS Reference Costs 2020/2021 - SB12Z (deliver simple parenteral chemotherapy at first attendance)[73]
IV	Complex parenteral chemotherapy	426.80	NHS Reference Costs 2020/2021 - SB13Z (deliver more complex parenteral chemotherapy at first attendance)[73]
IV	Complex parenteral chemotherapy with prolonged infusion	526.52	NHS Reference Costs 2020/2021 - SB13Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance)[73]
IV	Complex parenteral chemotherapy with subsequent infusion(s)	897.42	NHS Reference Costs 2020/2021 - SB13Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance) + SB15Z (deliver subsequent elements of a chemotherapy cycle)[73]
IV	Complex parenteral chemotherapy with prolonged infusion and subsequent infusion(s)	997.14	NHS Reference Costs 2020/2021 - SB14Z (deliver complex parenteral chemotherapy, including prolonged infusion, at first attendance) + SB15Z (deliver subsequent elements of a chemotherapy cycle)[73]
Oral	Oral drug dispensing	245.23	NHS Reference Costs 2020/2021 - SB11Z (deliver exclusively oral chemotherapy)[73]

Based on Table 86 of CS.[5]

CS = company submission; IV = intravenous; NHS = National Health Service; RCC = renal cell carcinoma; SC = subcutaneous

In addition, the company used exponential rates of PFS failure to approximate discontinuation rates for first-line metastatic treatments for advanced RCC and advanced pNET, as explained in Section 4.2.6.3. Some regimens are also constrained by a maximum treatment duration following dosing schedules recommended by NICE, as shown in Table 4.31 and 4.32. The median ToT for each second-line metastatic treatment of RCC was sourced from relevant second-line clinical trials conducted in advanced RCC populations, as shown in Table 4.31, except for interferon a2B which was extracted from a first-line trial (no second-line setting clinical data were available). For the pNET cohort, the median ToT for all second-line metastatic treatments was assumed to be 4 months (17.4) weeks. This was based on the median PFS reported by Hentic et al. 2012,[95] a clinical trial investigating the use of folic acid, fluorouracil, and irinotecan (FOLFIRI) as second-line treatment of pNET. The mean ToT for each subsequent therapy was calculated as a function of the median ToT and assuming constant hazards. Finally, the estimated discontinuation rate and (where applicable) the maximum ToT of each component in a treatment regimen was used to estimate the mean total costs in the first- and second-line setting. These were calculated as a weighted average based on first- and second-line market shares, as can be seen in Table 4.32.

Table 4.31: ToT for second-line treatment regimens in the advanced RCC setting

Second-line treatment regimen	Component	ToT (months)		Source
		Median	Mean	
Nivolumab	Nivolumab	23.9	34.5	Motzer et al. 2015 [CheckMate 025][97]
Axitinib	Axitinib	35.7	51.4	Motzer et al. 2013 [AXIS][98]
Cabozantinib	Cabozantinib	36.5	52.7	Motzer et al. 2018 [METEOR][99]
Lenvatinib/ Everolimus	Lenvatinib	33.0	47.7	Motzer et al. 2015 [NCT01136733][89]
	Everolimus	33.0	47.7	
Pazopanib	Pazopanib	32.2	46.4	Sternberg et al. 2013 [VEG105192][100]
Sunitinib	Sunitinib	32.2	46.4	Assume same median ToT as pazopanib
Everolimus	Everolimus	19.1	27.6	Motzer et al. 2018 [METEOR][99]
Temsirolimus	Temsirolimus	19.1	27.6	Hutson et al. 2014 [INTORSECT][101]
Cytokines (interferon)	Interferon a2B	12.0	17.3	Rini et al. 2008 [CALGB 90206][102]

Based on Table 87 of CS.[5]
CS = company submission; RCC = renal cell carcinoma; ToT = time on treatment

Table 4.32: Metastatic treatment market shares and costs

Metastatic therapy regimens	Market share*		Total cost (£)	
	Belzutifan	SoC	Acquisition	Administration
First-line metastatic therapy (metastatic RCC)				
Sunitinib	30.0%	30.0%	24,866	2,846
Tivozanib	14.0%	14.0%	28,217	3,587
Pazopanib	29.0%	29.0%	26,106	4,066

Metastatic therapy regimens	Market share*		Total cost (£)	
	Belzutifan	SoC	Acquisition	Administration
Cabozantinib	13.0%	13.0%	164,162	8,894
Nivolumab/ipilimumab	14.0%	14.0%	110,894	7,420
Avelumab/axitinib	0.0%	0.0%	221,156	18,242
No active treatment	0.0%	0.0%	0	0
Second-line metastatic therapy (metastatic RCC)				
Nivolumab	0.0%	0.0%	41,804	3,118
Axitinib	7.0%	7.0%	46,133	3,154
Cabozantinib	32.0%	32.0%	63,235	3,231
Lenvatinib/everolimus	0.0%	0.0%	42,856	2,923
Pazopanib	4.0%	4.0%	18,274	2,846
Sunitinib	0.0%	0.0%	15,796	1,897
Tivozanib	0.0%	0.0%	22,083	2,807
Everolimus	7.0%	7.0%	15,804	1,692
Sorafenib	0.0%	0.0%	19,385	1,385
Cytokines (interferon)	0.0%	0.0%	2,159	6,259
No active treatment	50.0%	50.0%	0	0
First-line metastatic therapy (metastatic pNET)				
Streptozocin/5-fluorouracil	0.0%	0.0%	9,044	2,984
Streptozocin/doxorubicin	0.0%	0.0%	9,052	2,984
Temozolomide/capecitabine	0.0%	0.0%	448	2,038
Everolimus	0.0%	0.0%	64,837	6,373
Sunitinib	0.0%	0.0%	48,629	5,311
Interferon a2B	0.0%	0.0%	6,133	35,551
Lanreotide	50.0%	50.0%	18,528	7,149
Octreotide	50.0%	50.0%	12,989	7,149
No active treatment	0.0%	0.0%	0	0
Second-line metastatic therapy (metastatic pNET)				
Cisplatin/etoposide	0.0%	0.0%	172	4,270
Everolimus	25.0%	25.0%	15,650	1,538
FOLFIRI	0.0%	0.0%	347	11,259
FOLFOX	0.0%	0.0%	297	11,259
Streptozocin/5-fluorouracil	25.0%	25.0%	6,826	2,252
Streptozocin/doxorubicin	25.0%	25.0%	6,832	2,252
Sunitinib	25.0%	25.0%	12,414	1,538
Temozolomide/capecitabine	0.0%	0.0%	338	1,538
Interferon a2B	0.0%	0.0%	1,565	9,072
Lanreotide/everolimus	0.0%	0.0%	21,528	2,268
Octreotide/everolimus	0.0%	0.0%	19,771	2,268

Metastatic therapy regimens	Market share*		Total cost (£)	
	Belzutifan	SoC	Acquisition	Administration
Lanreotide	0.0%	0.0%	5,878	2,268
Octreotide	0.0%	0.0%	4,121	2,268
No active treatment	0.0%	0.0%	0	0

Based on Table 88 of CS.[5]

* Source for market shares not provided in the CS.

CS = company submission; FOLFIRI = folic acid, fluorouracil, and irinotecan; FOLFOX = folic acid, fluorouracil, and oxaliplatin; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; SoC = standard of care

4.2.9.4 Health state costs

Health state costs are incurred in the pre-surgery, surgery, event-free after surgery, and metastatic disease modelled health states. Health sates costs are sourced from the NHS Reference Costs 2020/21 and the PSSRU Unit Costs of Health and Social Care 2021.[73, 75] Given the severity of the disease, the company used the highest complication and comorbidity (CC) score for costs when applicable. Costs included in the non-metastatic health states consist of costs for outpatient visits, laboratory tests and radiologic exams. In the metastatic health state, the analyses considered costs for salvage surgery, outpatient visits, laboratory tests and radiologic exams. The metastatic health state includes both pre- and post-progression metastases, therefore, costs are calculated as a weighted average of resource use pre- and post-progression metastases, with the weight based on the proportion of time spent progression-free within the metastatic disease state. Resource use frequencies in the pre-surgery and event-free after surgery health states were sourced from Maher et al. 2011.[103] Kanno et al. 2014[104] was used for the event-free after surgery health state for the CNS Hb cohort. In the metastatic disease health state, resource use was sourced from the KEYNOTE-564 trial (data cut-off date: 14 June 2021), and NICE TA542 (cabozantinib in the untreated locally advanced or metastatic RCC setting) for the RCC cohort, and TA476 (nab-paclitaxel with gemcitabine for untreated metastatic pancreatic cancer) for the pNET cohort.[84, 105] In addition, for a proportion of patients with metastatic RCC, one-time costs associated with salvage surgery were assumed to be incurred upon entering the metastatic disease state. The corresponding proportion of patients was estimated as the observed percentage of patients with surgery, among those who experienced distant metastases as their first DFS failure type, in the KEYNOTE-564 trial.[35] Finally, a one-time cost associated with palliative/terminal care was included for patients who died. Terminal care costs were estimated as £7,220.05 (inflation-adjusted using the health component of the Consumer Price Index (CPI) from the Office of National Statistics [ONS]), based on costs during the last 90 days before death as reported by Georghiou and Bardsley 2014.[106]. This was in line with NICE TA542 for cabozantinib in the untreated locally advanced or metastatic RCC setting.[84] Health state costs and resource use included in the economic model are summarised in Table 4.33.

Table 4.33: Health state costs

Cost item	Unit cost (£)	Patients (%)	Monthly resource use	Source
Pre-surgery and event-free after surgery states *				
GP visit	39.00	100%	0.08	GP costs – PSSRU – Unit Costs of Health and Social Care 2021[75]
Ophthalmologist visit	166.35	100%	0.08	WF01A - Service Code: 130 - Ophthalmology - Non-Admitted Face-to-Face Attendance, Follow-up - NHS Reference Costs 2020/21[73]
Complete blood count test	3.63	100%	0.08	DAPS05 - Haematology - Directly accessed pathology services - NHS Reference Costs 2020/21[73]
Urinalysis	1.85	100%	0.08	DAPS04 - Clinical Biochemistry - Directly accessed pathology services - NHS Reference Costs 2020/21[73]
CT scan of abdomen/pelvis	133.80	71%	0.08	Weighted average of total HRG activity for RD20A, RD21A, and RD22Z - NHS Reference Costs 2020/21[73]
Brain MRI (RCC and pNET cohort)	230.62	100%	0.04	Weighted average of total HRG activity for RD01A, RD02A, and RD03Z - NHS Reference Costs 2020/21[73]
MRI of brain (in CNS Hb cohort)	230.62	100%	0.08	Weighted average of total HRG activity for RD01A, RD02A, and RD03Z - NHS Reference Costs 2020/21[73]
Ultrasound	230.62	58%	0.08	Weighted average of total HRG activity for RD01A, RD02A, and RD03Z - NHS Reference Costs 2020/21[73]
Metastatic disease state (RCC origin tumour) **				
Salvage surgery	7,850.92	21%	1.00 (one-time upon entry only)	NHS. Robot-assisted nephrectomy: Evidence summary report (2014), inflation-adjusted to 2021 GBP***
Medical oncologist visit	224.55	100%	1.00	WF01A - Service Code: 370 - Medical Oncologist - Total Outpatient Attendances - NHS Reference Costs 2020/21[73]
Complete blood count test	3.63	100%	1.00	DAPS05 - Haematology - Directly accessed pathology services - NHS Reference Costs 2020/21[73]
CT scan of abdomen/pelvis	133.80	100%	1.00 (one-time upon entry)	Weighted average of total HRG activity for RD20A, RD21A, and RD22Z - NHS Reference Costs 2020/21[73]

Cost item	Unit cost (£)	Patients (%)	Monthly resource use	Source
Metastatic disease state (pNET origin tumour)[**]				
Medical oncologist visit	224.55	100%	1.00	WF01A - Service Code: 370 - Medical Oncologist - Total Outpatient Attendances - NHS Reference Costs 2020/21[73]
Cancer specialist nurse	90.49	50%	1.00	N10AF - Specialist Nursing, Cancer Related, Adult, Face to face - NHS Reference Costs 2020/21[73]
Complete blood count test	3.63	100%	6.00	DAPS05 - Haematology - Directly accessed pathology services - NHS Reference Costs 2020/21[73]
CT scan of abdomen/pelvis	133.80	100%	1.00 (one-time upon entry) 0.33 thereafter	Weighted average of total HRG activity for RD20A, RD21A, and RD22Z - NHS Reference Costs 2020/21[73]
MRI of abdomen/pelvis	230.62	10%	1.00 (one-time upon entry)	Weighted average of total HRG activity for RD01A, RD02A, and RD03Z - NHS Reference Costs 2020/21[73]
Ultrasound	230.62	5%	1.00 (one-time upon entry)	Weighted average of total HRG activity for RD01A, RD02A, and RD03Z - NHS Reference Costs 2020/21[73]
Death				
Terminal care	£7,220.05	100%	1.00 (one-time upon death)	Georghiou and Bardsley. Exploring the cost of care at the end of life. September 2014. Nuffield Trust. With inflation-adjustment to 2021 GBP.[106]

Based on Table 89 of CS.[5]

[*] Pre-surgery and event-free after surgery resource use is assumed to be the same for all cohorts in the model with the exception of MRI scan of brain

[**] Unless stated, for all metastatic states (i.e., pre- and post-progression), resource use is assumed to be the same for all cohorts in the model.

[***] No bibliographic details of this reference were available from the CS[5] and the paper was not included within the reference pack.

Notes: Frequencies of salvage surgery are based on observed percentages of patients with surgery among those who experienced distant metastases as their first DFS failure type in KEYNOTE-564. For the metastatic disease state in the pNET cohort, the 6 complete blood count tests include a total of 5 liver function tests and one blood test.

CS = company submission; CT = computed tomography; DFS = disease-free survival; GBP = Great British Pound; GP = general practitioner; MRI = magnetic resonance imaging; NHS = National Health Service; pNET = pancreatic neuroendocrine tumour; PSSRU = Personal Social Services Research Unit; RCC = renal cell carcinoma

4.2.9.5 Surgery and complications costs

The economic model also includes costs of surgery related to the primary tumour (i.e., RCC, pNET or CNS Hb) and surgery costs not related to the primary tumour. Costs incurred due to surgical complications are also considered. The company assumed that the risks of surgical complications are equal between the Belzutifan and SoC arms of the model. The model also distinguishes between short- and long-term surgical complications. Short-term complications costs are applied as a one-off cost per surgery. Long-term complications costs are considered annually and then converted into weekly costs, that were applied per model cycle to the cumulative proportion of patients who are assumed to develop long-term complications and are still alive. The costs of all surgical procedures, short-term and long-term complications can be seen in Tables 4.34, 4.35 and 4.36, respectively.

Table 4.34: Unit costs per surgical procedure

Surgery	Unit cost per procedure (£)	Cost source
Surgical procedure for VHL-associated RCC	7,850.92	Solutions for Public Health (on behalf of NHS England) (2014), inflation-adjusted to 2021 GBP*
Surgical procedure for VHL-associated CNS Hb	20,573.29	AA82Z – Total HRGs, Intracranial Telemetry, with Cortical Mapping or Resection of Brain NHS Reference Costs 2020/21[73]
Surgical procedure for VHL-associated–pNET	23,922.25	GA03C - Total HRGs, Very Complex, Hepatobiliary or Pancreatic Procedures, with CC Score 4+ - NHS Reference Costs 2020/21[73]
Adrenal lesion surgery	10,369.44	KA04 – Total HRGs, Adrenal Procedures with CC Score 2+ - NHS reference costs 2020/21[73]
Endolymphatic sac tumour surgery	5,029.35	Weighted average CB02A, CB02B–and CB02C Total HRGs - Non-Malignant, Ear, Nose, Mouth, Throat or Neck Disorder–, with Interventions - NHS reference costs 2020/21[73]
Epididymal cystadenoma surgery	6,609.57	Weighted average LB35C and LB35D - Scrotum, Testis or Vas Deferens Disorder–, with Interventions - NHS reference costs 2020/21[73]
Retinal Hb surgery	3,970.02	Weighted average BZ80A, B–80B, BZ81A and BZ81B - Complex/Very Complex Vitreous Retinal Procedures, 19 years and over - NHS reference costs 2020/21[73]

Based on Table 90 of CS.[5]

* No bibliographic details of this reference were available from the CS[5] and the paper was not included within the reference pack.

CNS = central nervous system; CS = company submission; GBP = Great British pounds; Hb = haemangioblastoma; NHS = National Health Service; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau

Table 4.35: Costs of short-term surgical complications

Complication	Cost per complication* (£)	Cost source
Short-term complications of surgical procedures for VHL RCC		
Acute renal failure	7,534.29	NHS Reference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21[73]
Cardiac complications	3,685.32	NHS Reference Costs (Weighted costs of EB03A, EB05A, E–10A, EB13A and EB14A - Total HRGs) 2020/21[73]

Complication	Cost per complication* (£)	Cost source
Erythroderma	8,559.58	NHS Reference Costs (JD07A Skin Disorders with Interventions, with CC Score 12+ - Total HRGs) 2020/21[73]
Kidney infection	7,612.90	NHS Reference Costs (LA04H Kidney or Urinary Tract Infections, with Interventions, with CC Score 12+ - Total HRGs) 2020/21[73]
Other genitourinary complications	1,375.85	NHS Reference Costs (LA09J General Renal Disorders with Interventions, with CC Score 6+ - Total HRGs) 2020/21[73]
Postoperative infection (RCC-related)	13,139.64	NHS Reference Costs (WH07A Infections or Other Complications of Procedures, with Multiple Interventions, with CC Score 2+ - Total HRGs) 2020/21[73]
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21[73]
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Short-term complications of surgical procedures for VHL CNS Hb		
Acute renal failure	7,534.29	NHS Reference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21[73]
CNS haemorrhage	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21[73]
Nerve palsy related to anaesthesia	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21[73]
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21[73]
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]

Complication	Cost per complication* (£)	Cost source
Short-term complications of surgical procedures for VHL pNET		
Abdominal abscess	10,881.28	NHS Reference Costs (FD10A Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions, with CC Score 8+ - Total HRGs) 2020/21[73]
Postoperative infection (pNET-related)	13,139.64	NHS Reference Costs (WH07A Infections or Other Complications of Procedures, with Multiple Interventions, with CC Score 2+ - Total HRGs) 2020/21[73]
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21[73]
Thrombosis and/or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Urinary tract infection	1,715.45	NHS Reference Costs (LA04 - Total HRGs) 2020-2021[73]
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Short-term complications of adrenal lesion surgery		
Acute renal failure	7,534.29	NHS Reference Costs (LA07H - Acute Kidney Injury with Interventions, with CC Score 11+, Total HRGs) 2020/21[73]
Respiratory complications	7,640.92	NHS Reference Costs (Weighted costs of D–16H, DZ26G and DZ27M - Total HRGs) 2020/21[73]
Thrombosis or embolism	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Vascular injury or anaemia	5,543.32	NHS Reference Costs (YQ50A Peripheral Vascular Disorders with CC Score 15+ - Total HRGs) 2020/21[73]
Short-term complications of retinal Hb surgery		
Vitreous haemorrhage	1,009.15	BZ24 - Non-surgical ophthalmology - NHS reference costs 2020/21[73]
Complications of endolymphatic sac tumour surgery		
Acoustic impairment	956.29	CB02 - Non-Malignant, Ear, Nose, Mouth, Throat or Neck Disorders - NHS reference costs 2020/21[73]

Based on Table 91 of CS.[5]
*Costs are applied as one-time costs for patients undergoing surgery in each cycle for short-term/acute complications

Complication	Cost per complication* (£)	Cost source
CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; HRG = Healthcare Resource Group; NA = not applicable; NHS = National Health Service; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma; VHL = Von Hippel-Lindau		

Table 4.36: Annual costs of long-term surgical complications

Complication	Cost per complication* (£)	Cost source
Annual cost of long-term complications for VHL RCC surgeries		
End stage renal disease and/or dialysis**	30,477.27	Kerr et al. 2012 (estimated expenditure on dialysis per patient requiring dialysis), inflation-adjusted to 2021 GBP[107]
Chronic kidney disease**	1,034.32	Kerr et al. 2012 (overall annual cost of CKD per patient diagnosed with CKD), inflation-adjusted to 2021 GBP[107]
Hernia surgery	2,021.79	Coronini-Cronberg et al. 2013, inflation-adjusted to 2021 GBP [108]
Chronic pain	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21[73]
Cerebral vasculature occlusion or stroke†	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21[73]
Annual cost of long-term complications for VHL CNS Hb surgeries		
Chronic pain (CNS Hb)	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21[73]
Cerebral vasculature occlusion or stroke†	7,883.91	NHS Reference Costs (AA35A Stroke with CC Score 16+ - Total HRGs) 2020/21[73]
Seizure	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21[73]
Neurological complications†	4,705.35	NHS Reference Costs (AA26C Muscular, Balance, Cranial or Peripheral Nerve Disorders, Epilepsy or Head Injury, with CC Score 15+ - Total HRGs) 2020/21[73]
Annual cost of long-term complications for VHL pNET surgeries		
Chronic pain (pNET)	1,872.11	NHS Reference Costs (WH08A Unspecified Pain with CC Score 1+ - Total HRGs) 2020/21[73]

Complication	Cost per complication* (£)	Cost source
Secondary diabetes or exocrine pancreatic insufficiency**	9,681.57	NHS Reference Costs (GC17A Non-Malignant, Hepatobiliary or Pancreatic Disorders, with Multiple Interventions, with CC Score 9+ - Total HRGs) 2020/21[73]
Immunocompromisation**	794.74	NHS Reference Costs (WJ11Z Other Disorders Of Immunity- Total HRGs) 2020/21[73]
Annual cost of long-term complications for adrenal lesion surgeries		
Adrenal insufficiency	876.28	KA08 - Other Endocrine Disorders - NHS reference costs 2020/21[73]
Chronic pain	649.85	–D05 - Abdominal Pain - NHS reference costs 2020/21[73]
Annual cost of long-term complications for Renal Hb surgeries		
Chronic pain	1,009.15	BZ24 - Non-surgical ophthalmology - NHS reference costs 2020/21[73]

Based on Table 92 of CS.[5]

*Costs are applied annually (recurring) from the time of surgery until death or the end of the modelled time horizon for long-term complications.

**This is a metabolic complication resulting from limited/absent organ function following last-resort surgery

†Social care costs are added to this unit cost in the cost-effectiveness model.

CNS = central nervous system; CS = company submission; Hb = haemangioblastoma; HRG = Healthcare Resource Group; NHS = National Health Service; pNET = pancreatic neuroendocrine tumour; RCC = renal cell carcinoma

4.2.9.6 AE costs

As mentioned in Section 4.2.7, the model includes costs and resource use associated with Grade ≥3 AEs occurring in ≥5% of patients treated with Belzutifan, and all Grade ≥3 TRAEs associated with Belzutifan observed in the MK-6482-004 trial. These AEs are also expected to incur a utility decrement, and the AEs durations were also sourced from the MK-6482-004 trial (see Section 3.2.8). In the SoC arm, the company assumed no AEs, the company assumed treatment-related approach to AEs instead of an all-cause approach. Unit costs were obtained from NHS 2020/21 Reference Costs and weighted by the risk of each AE among patients in the Belzutifan arm.[73] Costs associated to AE management are applied as one-time costs in the first model cycle. A summary of the AE-related costs included in the models is shown in Table 4.37.

Table 4.37: AEs (%) and costs included in the economic model

AEs	AE risk (%)	AE cost (£)	Source
Total cost of AEs	NA	46.62	Weighted average of rate and cost of individual AEs
Anaemia	[redacted]	356.39	NHS Reference Cost 2020/21, WH13: Abnormal Findings without Diagnosis - Regular Day or Night Admissions (weighted average) [73]
Fatigue	[redacted]	116.45	NHS Reference Cost 2020/21, WH17: Admission Related to Social Factors - Regular Day or Night Admissions (weighted average) [73]

Based on Table 93 of CS.[5]
AE = adverse event; CS = company submission; NA = not applicable; NHS = National Health Service

4.2.9.7 Miscellaneous unit costs and resource use

Finally, the company included social care costs for patients with sequalae/complications that require social care adaptations. The company indicated that social care is required mostly for patients with neurological disability resulting from CNS Hb and from debilitating surgical complications, because these patients are unable to perform standard activities of daily living and, therefore, require social care support for everyday needs.

The company mentioned that the NICE reference case stipulates that costs should be considered from an NHS and PSS perspective.[31] Despite this, the company included social care costs associated with stroke and neurological dysfunction as a complication of surgery associated with VHL have in their base-case analysis. Furthermore, for patients with progressive disease in the VHL CNS Hb cohort, social care costs associated with disease management are also included in the model to reflect the social care required for these patients who are expected to experience significant morbidity. In line with utility and caregiver disutility estimation approaches described in Section 4.8.2.5, motor neurone disease was used as proxy health condition. A summary of the social care costs included in the model are summarised in Table 4.38.

Table 4.38: Social care costs included in the economic model

Complication/Patient population	Annual cost (£)	Source and estimation method
Stroke	£3,232.45	Estimated societal costs of stroke in the UK based on a discrete event simulation (Patel et al., 2020)[109] Course of community rehabilitation reported in the supplementary material.

Complication/Patient population	Annual cost (£)	Source and estimation method
Neurological complications	£849.11	The size, burden and cost of disorders of the brain in the UK (Fineberg et al., 2013)[110] Proportion of estimated total UK per-subject cost of brain disorders attributed to direct non-medical costs converted to GBP and inflated to 2021.
Disease management of PD patients in VHL CNS Hb cohort	£1,085.31	Health Utilities and Costs for Motor Neurone Disease (Moore et al., 2019)[111] Community services cost over a 3-month period reported in the supplementary material inflated to 2021.
Based on Table 95 of CS.[5] CNS = central nervous system; CS = company submission; GBP = Great British Pound; Hb = hemangioblastomas; PD = progressed disease; UK = United Kingdom; VHL = Von Hippel-Lindau		

EAG comment:

The company have approached the implementation of costs and resource use in a comprehensive way. Appropriate sources seem to have been used in general, rationale for assumptions and references to previous appraisals have been provided. The only EAG concern relates to its inability to check (and therefore to validated) all cost items included in this submission. Given the time constraints associated to this project, this was not feasible.

4.2.10 Disease severity

The NICE reference case stipulates that the committee will regard all QALYs as being of equal weight. However, the committee may consider the severity of the condition, as determined by the absolute and proportional QALY shortfall (including discounting at the reference case rate), as decision modifier. Severity can be then taken into account quantitatively in the cost effectiveness analyses through QALY weighting, based on the absolute and proportional shortfall, as shown in Table 4.39. Whichever implies the greater severity level will be considered, and if either the proportional or absolute QALY shortfall falls exactly on the cut-off between two severity levels, the higher level will apply.[31]

Table 4.39: QALY weightings for disease severity

QALY weight	Proportional QALY shortfall	Absolute QALY shortfall
1.0	Less than 0.85	Less than 12
1.2	From 0.85 to 0.95	From 12 to 18
1.7	At least 0.95	At least 18
QALY = quality adjusted life year		

The results of the QALY shortfall analysis are shown in Table 4.40, where the total lifetime QALYs associated with SoC were obtained from the model results of the base-case analysis, and the estimated total QALYs for the general population reflected the baseline characteristics of the MK-6482-004 trial and the economic analyses (47.5% female and 41.0 years). These results suggest that a QALY weight of 1.7 can be applied to both VHL-associated RCC and CNS Hb cohorts, whereas a QALY weight of 1.2 can be applied to the VHL-associated pNET cohort. The company refers to Figure 9 in Section B.2.7 of the CS, to indicate that "*in the real world these three VHL cohorts are not actually distinct cohorts and therefore the appropriate severity weighting is for the full GB-indicated population, not based on primary tumour*".[5] Furthermore, the company emphasised that all patients in the MK-6482-004 trial had more than one tumour manifestation, and that based on "*what is known about the significant burden of*

VHL-pNET on mortality and morbidity and the severe complications arising from surgeries of the pancreas, there is little doubt that VHL-pNET meets the definition of highly severe that the severity modifier has been designed to identify".[5] For these reasons, the company decided to apply a QALY weight of 1.7 QALY to the CEA for all three VHL cohorts.

Table 4.40: Summary of company QALY shortfall analysis

Cohort	Expected total QALYs for the general population	Total expected QALYs for people with VHL on current SoC	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
VHL GB marketing authorisation population (weighted cohort)	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated RCC	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated CNS Hb	18.15	[redacted]	[redacted]	[redacted]	1.7
VHL-associated pNET	18.15	[redacted]	[redacted]	[redacted]	1.2

Based on Table 96 in CS.[5]
CNS Hb = central nervous system hemangioblastoma; CS = company submission; GB = Great Britain; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

EAG comment:

The QALY shortfall results presented in Table 4.40 were validated by the EAG with the Institute for Medical Technology Assessment (iMTA) Disease Burden Calculator (iDBC), an online free tool to estimate the total (and proportional) QALYs lost. In addition, the iDBC tool also estimates the likelihood of the applicable QALY weight based on the probabilistic sensitivity analysis (PSA) results provided in the company's model, which can be used to estimate the severity adjusted probability of being cost-effective.[112] The iDBC tool can be found here: https://imtamodels.shinyapps.io/iDBCv2_1/. The QALY shortfall calculations conducted by the EAG are shown in Table 4.41. These results are broadly in line with those presented by the company in Table 4.40 for the RCC, CNS Hb and pNET cohorts. The minor differences observed are likely due to using different utility sources and/or life tables to estimate expected QALYs for the total population (however, the EAG cannot confirm this because this information does not seem to be provided by the company) and using the PSA results of the company's model to estimate the QALYs for VHL patients under SoC. The uncertainty around the QALY weights is also presented in Table 4.41. This shows for example that, for the RCC cohort, even though a weighted point estimate is 1.7, there is a 44.1% that the applicable QALY weight is 1.2, which may have a substantial impact on the severity adjusted results. For the RCC cohort thus, reporting a severity weight of 1.7 only could be misleading because there is a large chance that severity weight is 1.2. For the other two cohorts, the uncertainty around the QALY weight is expected to be minor since

in both cohorts the likelihood that the deterministic weight is applicable in the probabilistic setting is higher than 95%. Finally, note that the so-called VHL-GB MA population has not been included in Table 4.41 since the EAG considers this population irrelevant for this submission.

Table 4.41: Summary of EAG QALY shortfall analysis

Cohort	Expected total QALYs for the general population	Total expected QALYs for people with VHL on current SoC	Absolute QALY shortfall	Proportional QALY shortfall	Likelihood QALY Weight (probability weight applicable)
VHL-associated RCC	18.02	[redacted]	[redacted]	[redacted]	1.7 (55.9%) 1.2 (44.1%)
VHL-associated CNS Hb	18.02	[redacted]	[redacted]	[redacted]	1.7 (95.4%) 1.2 (4.6%)
VHL-associated pNET	18.02	[redacted]	[redacted]	[redacted]	1.7 (1.7%) 1.2 (97.5%) 1.0 (0.8%)

Based on Institute for Medical Technology Assessment Disease Burden Calculator and company electronic model.[113, 114]

CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Furthermore, the EAG does not agree with the company's rationale for selecting a QALY weight of 1.7 for all VHL cohorts for the reasons summarised below:

- If, as the company suggests, "*in the real world these three VHL cohorts are not actually distinct cohorts and therefore the appropriate severity weighting is for the full GB-indicated population, not based on primary tumour*",[5] the EAG wonders what the value of analysing and presenting results for the different cohorts is. If only the results for the so-called VHL-GB MA population are valid, the CS should have been focused on that population only. This would have avoided the issues associated with the definitions of the different subgroups/cohorts described in Sections 4.2.2 and 4.2.3.
- The statement that "*there is little doubt that VHL-pNET meets the definition of highly severe*",[5] should be supported by evidence, otherwise it becomes a subjective interpretation. In this case, the evidence provided by the company shows that the most likely associated severity QALY weight for the VHL pNET cohort is 1.2.
- Also, in the company PSA, the severity weights differ per PSA iteration. The EAG agrees with this approach, however, this seems to contradict the company's rationale of using a QALY weight of 1.7 in all cases.
- In relation to the application of QALY weighting based on disease severity, the new NICE methods indicate that: "*The data used to estimate both absolute and proportional QALY shortfall should focus on the specific population for which the new technology will be used and be based on established clinical practice in the NHS*".[31] The EAG understands that in this case, the "specific population" are the three cohorts shown in Tables 4.40 and 4.41, since the absolute and proportional QALY shortfall are estimated using three different subsets of data.

For the reasons mentioned above, the EAG does not agree with the company in using the same QALY weight (1.7) for all subgroups. This should be based on the results presented in Table 4.38 and Table 4.39 and, therefore, for the VHL pNET cohort the deterministic QALY weight should be 1.2. The uncertainty around the distribution of the QALY weights should be considered too.

Finally, while validating the company's QALY shortfall results, the EAG found an error in the PSA calculations. The severity weights change per PSA iteration, which is methodologically correct. However, they are fixed for all VHL cohorts, which is incorrect because as shown in Tables 4.40 and 4.41 the QALY outcomes for SoC vary per cohort. The error is that the QALY weights are calculated as a weighted average for the MA population in each PSA iteration, and then applied to all three cohorts. The severity weights should be calculated for each VHL cohort separately.

5. COST EFFECTIVENESS RESULTS

5.1 Company's CE results

Table 5.1 shows the company's base-case deterministic CE results for the RCC, CNS Hb and pNET cohorts. All results are discounted. Results indicated that Belzutifan was more costly and more effective than SoC in all cohorts. Compared to SoC, in the RCC cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER in the RCC cohort was £73,095 per QALY gained. In the CNS Hb cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs compared to SoC. Therefore, the ICER in the CNS Hb cohort was £56,933 per QALY gained. Finally, in the pNET cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER in the pNET cohort was £77,649 per QALY gained.

When accounting for disease severity, as explained in Section 4.2.10, the company assumed a QALY weight of 1.7 for all VHL cohorts, which results on ICERs equal to £42,997, £33,490 and £45,676 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. As also explained in Section 4.2.10, the EAG does not agree with the company's rationale for selecting a QALY weight of 1.7 for all cohorts and considers that the severity adjusted QALYs should be based on the weight likelihood estimated in Table 4.39. Based on this distribution the severity adjusted ICERs are equal to £49,359, £33,976 and £64,311 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively.

The disaggregated discounted QALYs, life years (LYs) and costs are shown in Tables 5.2, 5.3 and 5.4, respectively.

Table 5.1: Company base-case deterministic CE results (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – company*	Severity Adjusted ICER – EAG**
VHL RCC cohort									
SoC	[redacted]	[redacted]	[redacted]						
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	73,095	42,997	49,359
VHL CNS Hb cohort									
SoC	[redacted]	[redacted]	[redacted]						
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	56,933	33,490	33,976
VHL pNET cohort									
SoC	[redacted]	[redacted]	[redacted]						
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	77,649	45,676	64,311

Based on Tables 98, 99 and 100 in the CS v3.0,[5] Tables 98, 99 and 100 in the CS v2.0,[115] the economic model[114] and the iDBC tool.[113]

* Company's severity adjusted ICERs based on a QALY weight equal to 1.7 for all cohorts (see Table 4.38).

** EAG's severity adjusted ICERs based on distribution of QALY weights per cohort (see Table 4.39).

CE = cost effectiveness; CNS Hb = central nervous system hemangioblastoma; CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; iDBC = Institute for Medical Technology Assessment Disease Burden Calculator; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Table 5.2: Disaggregated QALYs results

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	Belzutifan	SoC	Belzutifan	SoC	Belzutifan	SoC
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	Belzutifan	SoC	Belzutifan	SoC	Belzutifan	SoC
Surgical complication disutility for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for other tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS.[5]
AE = adverse event; CS = company submission; CNS Hb = central nervous system hemangioblastoma; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Table 5.3: Disaggregated life years results

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	Belzutifan	SoC	Belzutifan	SoC	Belzutifan	SoC
Total life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS.[5]
CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Table 5.4: Disaggregated cost results (£)

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	Belzutifan	SoC	Belzutifan	SoC	Belzutifan	SoC
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Belzutifan treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	Belzutifan	SoC	Belzutifan	SoC	Belzutifan	SoC
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for other tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS.[5]

AE = adverse event; CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Overall, the new technology is modelled to affect QALYs by:

- Increasing the number of QALYs pre-surgery and reducing the number of disutilities associated to surgical complications for primary tumour.
- In all other health states, the difference in QALYs is not substantial.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments.
- Decreasing costs associated to surgery and surgical complication costs for primary, and to a lower extent, other types of tumours.
- A moderate reduction in costs due to advanced treatments.

EAG comment:

Following the EAG comments in Section 4.2.10 of this report, the EAG considers that the severity adjusted QALYs should be based on the weight likelihood estimated in Table 4.39. In all base-case scenarios, with and without severity weight, all ICERs are above the commonly used threshold ICER of £30,000 per QALY gained. Only the ICER for the CNS Hb cohort is close to the threshold ICER.

5.2 Company's sensitivity analyses

5.2.1 PSA

The company conducted a PSA in which all relevant input parameters were sampled simultaneously from their corresponding probability distributions over 1,000 iterations. The input parameters and the probability distributions used in the PSA can be found in Appendix J of the CS.[5]

The average PSA results are summarised in Table 5.5 for all three cohorts. The results without severity weights applied are in line with the deterministic ones shown in Table 5.1, but with slightly larger ICERs due to higher incremental costs and lower incremental QALYs obtained in the PSA. When accounting for disease severity, the company indicates in the CS that a QALY weight of 1.7 for all VHL cohorts was assumed.[5] However, this seems to contradict the model implementation where the severity weights change per PSA iteration (see model sheet "PSA Calculation" – Column EC). The EAG considers that this approach (i.e., changing the severity weights per PSA iteration) is methodologically correct. Therefore, the PSA severity weighted ICERs in the model do not match with those presented in the CS and in Table 5.5. The PSA severity weighted ICERs in the model equal to £51,622, £39,771 and £54,072 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. Thus, the PSA severity-weighted ICERs in the model are between £5,000-£7,000 larger than those presented in the CS and in Table 5.5. However, as already mentioned in Section 4.2.10, the severity weights in the model seem to be fixed per VHL cohort, which is incorrect because, as shown in Table 4.38 and 4.39, the QALY outcomes for SoC vary per cohort. The error is that the QALY weights are calculated for the RCC cohort, but then these weights are applied to all three cohorts. The PSA severity weighted ICERs calculated by the EAG, based on the model PSA outcomes, equal to £51,116, £34,788 and £66,136 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. These ICERs are similar to the deterministic ones shown in Table 5.1.

Table 5.5: Company base-case probabilistic CE results (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – company*	Severity Adjusted ICER – EAG**
VHL RCC cohort									
SoC	[redacted]	NR	[redacted]						
Belzutifan	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]	76,253	44,854	51,116
VHL CNS Hb cohort									
SoC	[redacted]	NR	[redacted]						
Belzutifan	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]	58,398	34,352	34,788
VHL pNET cohort									
SoC	[redacted]	NR	[redacted]						
Belzutifan	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]	79,842	46,966	66,136

Based on Table 101 in the CS v3.0,[5] Table 101 in the CS v2.0,[115] the economic model[114] and the iDBC tool.[113]

* Company's severity adjusted ICERs based on a QALY weight equal to 1.7 for all cohorts (see Table 4.38).

** EAG's severity adjusted ICERs based on distribution of QALY weights per PSA iteration and VHL-related cohort.

CE = cost effectiveness; CNS Hb = central nervous system hemangioblastoma; CS = company submission; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; iDBC = Institute for Medical Technology Assessment Disease Burden Calculator; Inc. = incremental; LYG = life years gained; NR = not reported; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

The company also plotted the PSA outcomes on a CE-plane for the three cohorts separately. These are shown in Figures 5.1, 5.2 and 5.3, for the RCC, CNS Hb and pNET cohorts, respectively. Note that in these plots, severity weighting for QALYs is not considered. It can be seen that for the three cohorts [redacted]. From the PSA results, a cost effectiveness acceptability curve (CEAC) was also calculated and plot in Figures 5.4, 5.5 and 5.6. The CEAC plot indicates that [redacted]. At the common thresholds of £20,000 and £30,000 per QALY gained, the estimated probability that Belzutifan is a cost-effective alternative to SoC was [redacted] for the three cohorts. When accounting for disease severity, as explained above in this Section, the company assumed a QALY weight of 1.7 for all VHL cohorts, but the EAG does not agree with this approach: severity weights should vary per PSA iteration and cohort. However, the EAG considers that in this case a plot of the PSA outcomes on a CE-plane might be misleading. This is because dots may be clustered depending on the distribution of the severity weights and its interpretation is unclear. Therefore, CE-plane plots based on severity weighting are not presented in the EAG report. As suggested by Versteegh et al. 2019,[112] the severity adjusted probability of being cost effective is more informative in this situation. This was calculated by the EAG based on the PSA model outcomes, and the threshold of £30,000 per QALY gained, it was equal to [redacted]% for the RCC and pNET cohorts, and [redacted]% for the CNS Hb cohort.

Figure 5.1: PSA CE-plane (Belzutifan list price): VHL RCC cohort

[redacted]

Based on the latest model version provided by the company.[114]
CE = cost effectiveness; PSA = probabilistic sensitivity analysis; RCC = renal cell carcinoma; QALY = quality-adjusted life year; VLH = Von Hippel-Lindau; WTP = willingness to pay

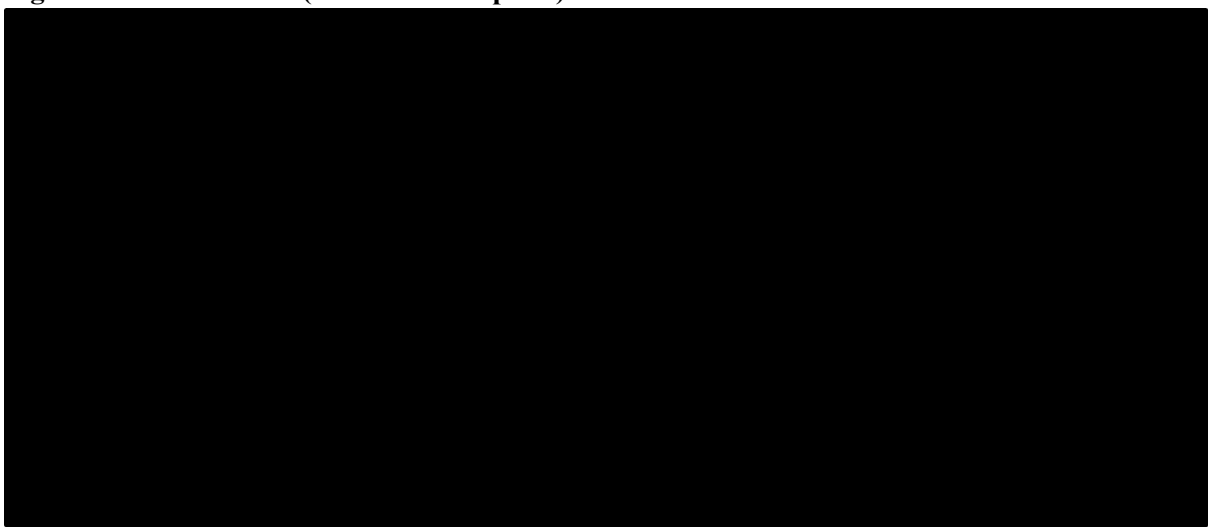
Figure 5.2: PSA CE-plane (Belzutifan list price): VHL CNS Hb cohort

Based on the latest model version provided by the company.[114]
CE = cost effectiveness; CNS = central nervous system; Hb = haemangioblastoma; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; VLH = Von Hippel-Lindau

Figure 5.3: PSA CE-plane (Belzutifan list price): VHL pNET cohort

Based on the latest model version provided by the company.[114]
CE = cost effectiveness; CNS = central nervous system; pNET = pancreatic neuroendocrine tumour; QALY = quality-adjusted life year; VLH = Von Hippel-Lindau

Figure 5.4: PSA CEAC (Belzutifan list price): VHL RCC cohort



Based on the initial model version provided by the company.[114]
CEAC = cost-effectiveness acceptability curve; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VLH = Von Hippel-Lindau

Figure 5.5: PSA CEAC (Belzutifan list price): VHL CNS Hb cohort

Based on the initial model version provided by the company.[114]
CEAC = cost-effectiveness acceptability curve; CNS = central nervous system; Hb = haemangioblastoma; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VLH = Von Hippel-Lindau

Figure 5.6: PSA CEAC (Belzutifan list price): VHL pNET cohort

Based on the initial model version provided by the company.[114]
CEAC = cost-effectiveness acceptability curve; pNET = pancreatic neuroendocrine tumour; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; SoC = standard of care; VLH = Von Hippel-Lindau

EAG comment:

The EAG noted, as explained above, that the company's implementation of the severity weights was incorrect. First, in the CS, it was assumed the same severity weight of 1.7 for all cohort and all PSA iterations. This does not match with the model implementation where the severity weight varies per PSA iteration. Therefore, the severity adjusted ICERs in the CS are incorrect. Furthermore, in the model, the severity weights were calculated for the RCC cohort but applied to the CNS Hb and pNET cohorts too. This is also incorrect. The largest impact is observed in the pNET cohort, because the company assumed a weight of 1.7, but for this cohort the majority of the PSA outcomes resulted in a weight of 1.2. As a consequence, the severity adjusted ICER calculated by the EAG is approximately £20,000 larger than the severity adjusted ICER calculated by the company. When severity weighting was not considered, at the common thresholds of £20,000 and £30,000 per QALY gained, the estimated probability that Belzutifan is a cost-effective alternative to SoC was [redacted] for the three cohorts. When accounting for disease severity, the severity adjusted probability of being cost effective at the threshold of £30,000 per QALY gained, was equal to [redacted]% for the RCC and pNET cohorts, and [redacted]% for the CNS Hb cohort.

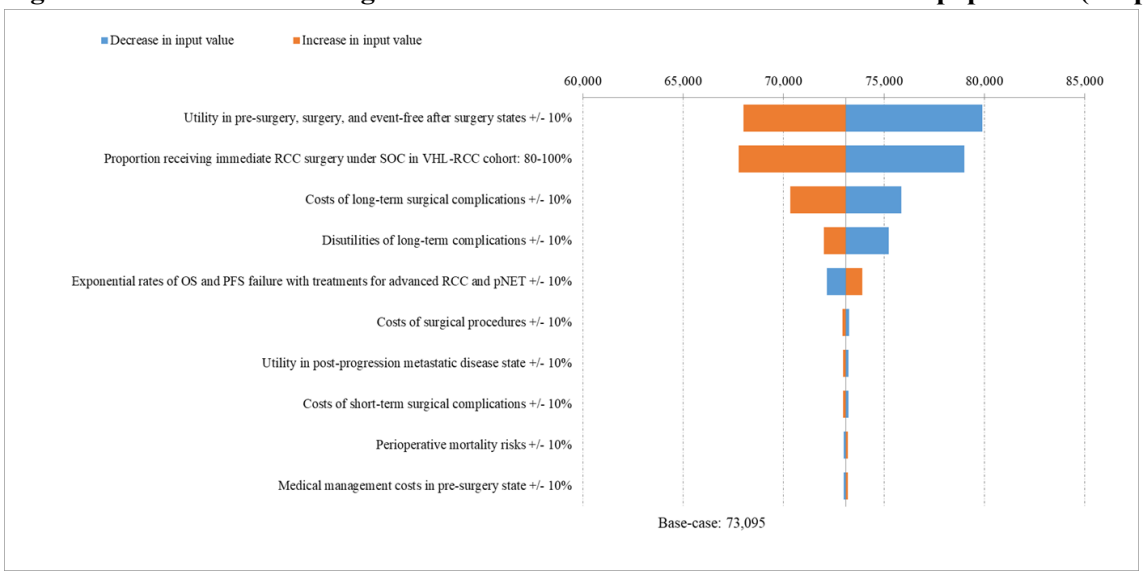
5.2.2 Deterministic sensitivity analysis (DSA)

The company also conducted deterministic sensitivity analyses (DSAs) comparing Belzutifan against SoC for the three relevant cohorts separately. Key parameters were individually varied at lower and upper bounds of values presented in Appendix J of the CS.[5]

The results of the DSAs were presented by the company in the form of tornado diagrams showing the 15 parameters that have the greatest influence on the ICER. The tornado diagram for each cohort, without accounting for QALY weighting, can be seen in Figure 5.7, 5.8 and 5.9. In general, the utility

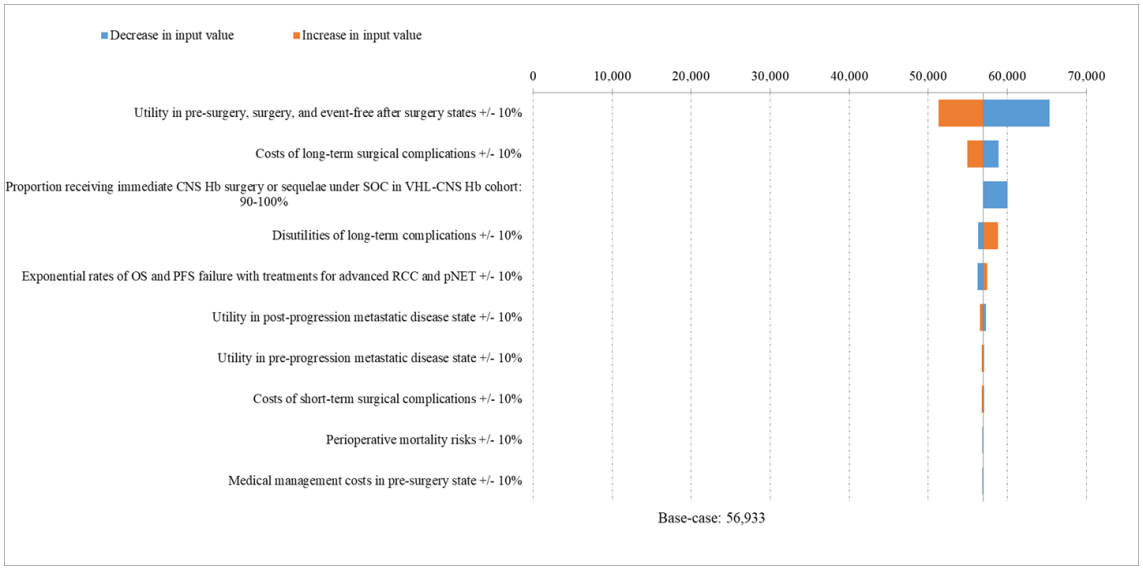
associated with the non-metastatic health states was the most sensitive parameter for the VHL RCC and VHL CNS Hb cohorts and the second most sensitive parameter for the VHL pNET cohort. The proportion of patients receiving immediate surgery was the most sensitive parameter for the VHL pNET cohort and the second and third most sensitive parameter for VHL RCC and VHL CNS Hb cohorts, respectively.

Figure 5.7: DSA tornado diagram for Belzutifan vs. SoC in the VHL RCC population (list price)



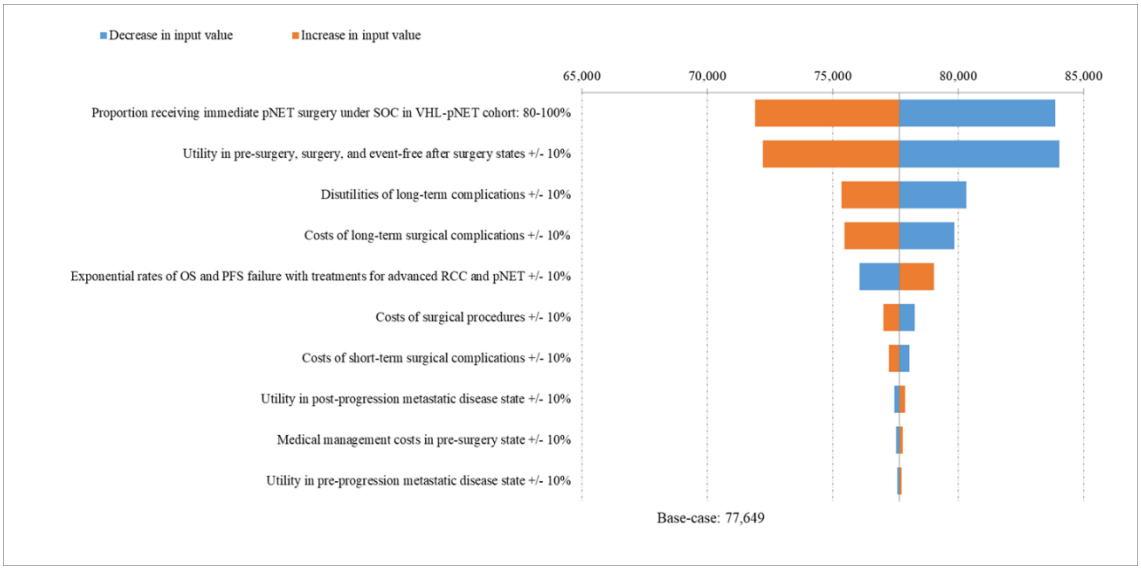
Based on Figure 31 in CS Document B – 19 May 2023[115].
CS = company submission; DSA = deterministic sensitivity analysis; OS = overall survival; PFS = progression-free survival; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

Figure 5.8: DSA tornado diagram for Belzutifan vs. SoC in the VHL CNS Hb population (list price)



Based on Figure 32 in CS Document B – 19 May 2023[115].
CNS = central nervous system; CS = company submission; DSA = deterministic sensitivity analysis; Hb = haemangioblastoma; OS = overall survival; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumour; SoC = standard of care; VHL = Von Hippel-Lindau

Figure 5.9: DSA tornado diagram for Belzutifan vs. SoC in the VHL pNET population (list price)



Based on Figure 33 in CS Document B – 19 May 2023[115].
CNS = central nervous system; CS = company submission; DSA = deterministic sensitivity analysis; Hb = haemangioblastoma; OS = overall survival; PFS = progression-free survival; pNET = pancreatic neuroendocrine tumour; SoC = standard of care; VHL = Von Hippel-Lindau

5.2.3 Scenario analysis

The company presented in total the results of 14 scenario analyses to assess the robustness of the model results to changes in some modelling assumptions. A summary of the results of these scenarios is provided in Table 5.6. These included exploring alternative long-term assumptions regarding Belzutifan treatment effect, changing utilities, considering alternative pre-surgery rates for the VHL RCC cohort in the Belzutifan arm, not adjusting surgery and metastases rates to account for real-world SoC, not applying relative dose intensity in the calculations, including indirect costs, considering shorter model time horizons or lower discount rates.

Note that the company only presented ICERs with and without severity weighting, but the QALY weights used in each scenario were not reported. In this respect, the company indicated that these "*results have the severity modifier applied in the QALY calculation, some of these results should be interpreted with caution as a lower weighted severity modifier is applied which is not appropriate for this appraisal*".[5] The EAG is unclear about how interpret this sentence since, as explained above, the company assumed for all three VHL associated cohorts a QALY weight based on the VHL-GB MA population (e.g., weighted cohort). The same approach is taking in the scenario analyses but in some scenarios the severity weighting is 1.2 instead of 1.7 assumed in the base-case. The EAG considers this approach correct since scenarios are based on different assumptions representing underlying uncertainties, which are also applicable to the severity weighting, which does not necessarily have to be equal 1.7 under all circumstances.

When severity weighting was not considered, all scenarios for all cohorts resulted in ICERs clearly above the threshold of £30,000 per QALY gained, except in the scenario where no treatment effect waning for Belzutifan is assumed. In that scenario, all ICERs without severity weighting are below the threshold of £30,000 per QALY gained and for the VHL CNS Hb cohort it is even below £20,000 per QALY gained. Even when severity weighting was applied, all scenarios for the RCC and pNET cohorts resulted in ICERs above the threshold of £30,000 per QALY gained, except in the scenario where no treatment effect waning for Belzutifan is assumed. For the CNS Hb cohort, most of the ICERs are still

above the threshold of £30,000 per QALY gained, but in the range of £30,000-£35,000 per QALY gained.

The EAG also presented in Table 5.6 ICERs based on EAG estimates for the severity weights. Unless otherwise stated, the EAG used the same weights as in the EAG base-case shown in Table 4.39. Note that the base-case weights were applied to all scenarios resulting in no change in the QALYs accrued in the SoC arm. In general, the EAG's severity adjusted ICERs were different to those estimated by the company. This is because the company used a fixed weight, either 1.7 or 1.2, whereas the EAG used a mix of weights based on the expected distribution of the QALY weights obtained with the iDBC tool.

In conclusion, the modelling assumptions explored by the company that had the greatest effect on the ICER were related to:

- Utility in the non-metastatic health states.
- The proportion to receive immediate surgery in the SoC arm.
- The removal of treatment effect waning.

EAG comment:

The main concerns of the EAG relate to:

- The (sometimes) unclear rationale for conducting scenarios.
- Many assumptions in the model were not tested (or not shown) by the company. Therefore, it is unclear the impact of other modelling assumptions. Some of these were explored by the EAG in Section 6.2 of this report.

Table 5.6: Summary of company scenario analyses

Scenario	Description	ICER VHL RCC cohort (£/QALY)			ICER VHL CNS Hb cohort (£/QALY)			ICER VHL pNET cohort (£/QALY)		
		Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*
Base-Case		**73,095**	**42,997**	49,359	**56,933**	**33,490**	33,976	**77,649**	**45,676**	64,311
1	Assume no treatment effect waning: Model efficacy and ToT separately	26,913	15,831	18,191	17,655	10,385	10,528	23,052	13,560	19,101
2	Distribution for Belzutifan ToT: Weibull	91,265	53,685	61,686	71,497	42,057	42,634	96,471	56,748	79,933
3	Annual discount rate: 3.5% for costs and 1.5% for effectiveness	67,209	39,535	39,535 (w = 1.7*100%)	52,095	30,644	(w = 1.7*100%)	67,880	39,930	39,930 (w = 1.7* 99.8% + 1.2*0.2%)
4	Annual discount rate 0%	73,359	43,153	43,153 (w = 1.7*100%)	52,864	31,097	31,097	65,377	38,457	38,457 (w = 1.7*100)

Scenario	Description	ICER VHL RCC cohort (£/QALY)			ICER VHL CNS Hb cohort (£/QALY)			ICER VHL pNET cohort (£/QALY)		
		Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*
5	Annual discount rate 1.5%	72,491	42,642	42,642 (w = 1.7*100%)	54,352	31,972	31,972 (w = 1.7*100%)	70,589	41,523	41,523 (w = 1.7*99.8% + 1.2*0.2%)
6	Time horizon: 20 years**	68,515	57,096 (w = 1.2)	53,843 (w = 1.7*25.7% + 1.2*46.3% + 1.0*28%)	55,230	46,025	41,198 (w = 1.7*34.4% + 1.2*49.9% + 1.0*15.7%)	79,193	65,994	74,012 (w = 1.7* 3.4% + 1.2*23.1% + 1.0*73.5%)
7	Time horizon: 30 years**	71,547	59,623	52,931 (w = 1.7*31.9% + 1.2*64.2% + 1.0*3.9%)	56,234	46,861	39,820 (w = 1.7*43.2% + 1.2*54.9% + 1.0*1.9%)	77,593	64,661	66,563 (w = 1.7* 6.3% + 1.2*60.8% + 1.0*32.9%)
8	Assume same utility for CR as PR/stable disease	73,710	43,359	49,821	57,392	33,760	34,223	81,086	47,698	67,185
9	Apply caregiver disutility	69,940	41,141	47,273	53,496	31,468	31,900	72,876	42,868	60,382
10	Do not apply age-adjusted disutility	70,192	41,290	51,366 (w = 1.7*33.3% + 1.2*66.7%)	54,042	31,789	32,225	75,002	44,119	61,176 (w = 1.7* 5.8% + 1.2*92.7% + 1*1.5%)

Scenario	Description	ICER VHL RCC cohort (£/QALY)			ICER VHL CNS Hb cohort (£/QALY)			ICER VHL pNET cohort (£/QALY)		
		Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*
11	Distribution for pre-surgery → surgery in the Belzutifan arm (VHL RCC cohort): Gamma	76,127	44,781	52,783	NR	NR	NR	NR	NR	NR
12	Do not adjust surgery and metastases rates to account for real-world standard of care	75,814	63,178	56,430 (w = 1.7*28.7% + 1.2*71.3%)	49,901	41,584	34,320 (w = 1.7*50.8% + 1.2*49.2%)	55,768	46,474	50,423 (w =1.2*53% + 1*47%)
13	Do not apply relative dose intensity	81,504	47,944	55,089	63,363	37,272	37,783	86,606	50,945	71,759
14	Include indirect costs (societal perspective)	60,516	35,597	40,899	47,670	28,041	28,400	60,124	35,367	49,716
Based on Tables 102, 103 and 104 in the CS,[5] Tables 102, 103 and 104 in CS Document B – 19 May 2023,[116] the economic model,[114] and the iDBC tool.[113]										

Scenario	Description	ICER VHL RCC cohort (£/QALY)			ICER VHL CNS Hb cohort (£/QALY)			ICER VHL pNET cohort (£/QALY)		
		Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*	Unadjusted	SA company	SA EAG*
* Same weight distribution as in EAG base-case unless otherwise indicated (w_{RCC} = 1.7*55.9% + 1.2*44.1%, W_{CNS} = 1.7*95.4% + 1.2*4.6%, W_{pNET} = 1.7*1.7% + 1.2*97.5% + 1.0*0.8%) ** To estimate the distribution of QALY weights in these scenarios the EAG adjusted the cohort age on the iDBC tool in such a way that that number of QALYs without the disease (i.e., for the general population) was similar to that estimated in the model. For the 20-year time horizon scenario age in the iDBC was 59 years and in the 30-year time horizon scenario age was 50 years. CNS = central nervous system; CR = complete response; CS = company submission; EAG = Evidence Assessment Group; Hb = hemangioblastoma; ICER = incremental cost-effectiveness ratio; NR = not reported; PD = progressed disease; pNET = pancreatic neuroendocrine tumour; PR = partial response; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SA = severity adjusted; SD = stable disease; ToT = time on treatment; VHL = Von Hippel-Lindau									

5.3 Model validation and face validity check

Some of the validation efforts conducted on the economic model were discussed in the validation Section of the CS (B.3.14).[5] The validation efforts discussed in Section B.3.14 of the CS technical verification, external and cross-validation. Other validation aspects, such as the validation of the transition probabilities used in the model or how clinical experts feedback was used to validate other modelling features are scattered over Document B of the CS.[5] In addition, more details about model validation were provided by the company in response to some clarification questions.[4] In the remaining of this section, the validation efforts performed on the model, as presented by the company, are categorised according to the types of validation used in the Assessment of the Validation Status of Health-Economic decision models (AdViSHE) tool.[117]

5.3.1 Validation of the conceptual model

5.3.1.1 Face validity testing (conceptual model)

United Kingdom clinical experts were consulted via an advisory board (and in individual consultation meetings) to validate model assumptions from a clinical perspective. Experts were selected based on experience treating VHL, ensuring that a broad geographic range in the UK was covered. Experts were engaged through individual consultation meetings. One expert was contracted for a series of consultations where an honorarium was paid and some of the experts were later engaged through an advisory board, where an honorarium was also paid. The company considered the experts suitably qualified to provide input on the evidence used for this submission, having the following experience:

- A consultant endocrinologist who runs a VHL multidisciplinary team (MDT) that manages 50-60 VHL patients, inclusive of RCC, CNS Hb and pNET patients.
- Three consultant endocrinologists that attend VHL MDTs, who have expertise in recommending and referring VHL RCC, CNS Hb and pNET patients for treatment.
- Four consultant urological surgeons, the respective regional leads for VHL surgery, specialising in treatment for VHL RCC.
- A consultant neurologist, neurosurgeon and a neuro-oncologist who manage patients with VHL CNS Hb.
- Three consultant clinical geneticists, respective leads of regional genetics services that manage patients with VHL.
- A professor of medical genetics and lead author on an evaluation of tumour surveillance protocols and outcomes in VHL disease in the UK.
- An interventional radiologist who provides non-surgical treatment for patients with VHL RCC.
- Two medical oncologists that have experience with treating VHL tumours.

The company explained in Section B.3.14 of the CS that individual consultation meetings and the discussion at the advisory board, experts were asked questions about their experience managing VHL, typical patient profiles and patient management, how they would interpret the MHRA label population, expected consequences of surgery in the MHRA label population, interpretation of Belzutifan clinical data, and validation of model assumptions. In particular, the following points were of interest:[5]

- General insights and unmet need on VHL.
- Treatment pathways and insights for each of the three VHL tumour manifestations included in the MHRA label.
- Other common VHL manifestations.
- Primary goals of VHL treatment.

- Interpretation of the MHRA label.

In addition, in response to clarification question B31,[4] the company explained that the initial model was conceptualised to include VHL-related RCC tumours only. At that stage, a clinical panel with one external expert physician, Dr Eric Jonasch, MD (Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center), supplemented by an internal therapy area clinician, validated the conceptual model and other aspects of the model development plan. Experts confirmed that a Markov model structure was reasonable and reflected the VHL disease trajectory and its impact on costs and health outcomes over time. Experts also confirmed that tumour reduction surgeries are the most relevant clinical events in VHL disease process. Finally, they also confirmed the relevance of the metastatic disease health state, noting that metastases should be relatively infrequent for patients who are appropriately managed by SoC.

5.3.1.2 Cross-validity testing (conceptual model)

According to the company, this submission represents the first CE assessment of any treatment for VHL-associated RCC, pNET, and CNS Hb. Therefore, there is no evidence that can be used for cross-validation against other, independently developed economic models in the same indication. Previous NICE appraisals provide justification for some of the assumptions used in the economic analysis, but these mostly relate to the choice of input parameters and, therefore, are discussed below.

5.3.2 Input data validation

5.3.2.1 Face validity testing (input data)

Input parameters are estimated from different sources of data. The main source of evidence for Belzutifan is the MK-6482-004 trial. As discussed in previous Sections of this report, there are concerns regarding the mismatch between the population in the DP and the population in the MK-6482-004 trial. In this respect, it was acknowledged in response to clarification question A10d that the company was not able to validate whether the proportions of patients with VHL-associated RCC + CNS Hb, or with VHL-associated RCC + pNET in the MK-6482-004 trial reflect those expected to be observed in UK clinical practice with any external data sources which report this data. In addition, in response to clarification question B31, the company explained that the operationalisation in the model of patients for whom surgery is unsuitable or undesirable as receiving a 'last resort' surgery (leading loss of organ function and/or extremely poor outcomes) was validated with an UK consultant endocrinologist who runs a VHL MDT. However, the company also acknowledged the challenge to adapt the model to include the specifics of the decision problem population, namely patients who "require therapy" and for whom "localised procedures are unsuitable or undesirable".

Treatment effectiveness was estimated from a MAIC, as explained in Section 3.4. The set of baseline variables used for matching in the MAIC, was selected based on expert feedback to identify baseline characteristics that are likely to be prognostic variables for the transition probabilities from the pre-surgery state, or that may modify the effect of Belzutifan on these transition probabilities (see clarification question B31).[4]

The main source of evidence for SoC is the VHL Natural History Study. The company explained though that the surgery rates observed in the VHL Natural History Study are high which results in lower rates of metastases compared to what would be expected under SoC in UK clinical practice. This was validated by clinical experts, who also justified using real-world data from the Optum Clinformatics Data Mart Claims Study database to better reflect SoC in the UK, as described in Section B.3.3 of the CS.[5] In response to clarification question B5,[4] the company further explained that from the Optum

Clinformatics Data Mart Claims database, a clear difference in surgery rates for VHL-related tumours was observed compared to the surgery rates estimated from the VHL Natural History Study. Clinicians suggested that the cause of this difference was due to less proactive surveillance, which would result in less disease control and, therefore, higher rates of metastatic disease. The company thus used the Optum Clinformatics Data Mart Claims database to adjust the surgery and metastases rates estimated from both the MK-6482-004 trial and the VHL Natural History Study, as described in Sections 3.3 and 3.4. These adjustments imply that the surgery rates in the model are lower than those observed in the MK-6482-004 trial and the VHL Natural History Study, whereas the risks of metastasis are higher.

Once patients transition to the surgery health state, these are assumed to be at risk of experiencing complications directly related to the surgical procedure. As mentioned in Table 97 in the CS,[5] the company assumed that the risk of surgical complications is equal between both arms (i.e., Belzutifan is not expected to affect the rate of surgical complications, its expected benefit consists of avoiding surgeries). The risks of experiencing different surgical complications were also informed by the Optum Clinformatics Data Mart Claims database and validated by clinical experts at the company and an academic medical centre (Dr Eric Jonasch from MD Anderson). These complication rates were also adjusted to align with the UK MHRA label population. Furthermore, for non-primary tumours, it was assumed that patients are only at risk of complications from the first surgery since the model does not track the number of non-primary surgeries that patients undergo. This assumption was also validated by clinical experts.

The company discussed the validity of the utilities used in the economic model. In clarification question B19,[4] the EAG asked the company to discuss the (face) validity of the EQ-5D values presented in the CS, which were sourced from the VHL RW QoL Disease Burden Study.[1] In their response, the company indicated that the utility values presented in the CS show internal consistency, since worse disease status is consistently associated lower utility scores. The company stated that the utility values reported in the CS are all lower than the age- and sex-match utility values estimated for the general population in the UK. However, this was not explicitly shown by the company. Unfortunately, the SLR conducted by the company did not identify any studies reporting utility scores in a VHL population. The utility values obtained from the VHL RW QoL Disease Burden Study were presented and discussed in engagements with clinical experts. The utility values were deemed plausible, despite the differences in terms of severity of tumour manifestations in the VHL RW QoL Disease Burden Study and the population eligible for Belzutifan in the decision problem. However, the company acknowledged the difficulty to capture the disease severity in the decision problem population with data from the VHL RW QoL Disease Burden Study and as such it was recognised as a limitation of the current submission. The company also emphasised that comparing the utility values included in the model to those used similar diseases is challenging given the rarity of VHL. Utility values accepted in previous NICE appraisals in disease areas that might be used as a proxy to represent specific VHL health states (e.g., motor neurone disease as a proxy for the event-free post-surgery health state in patients VHL CNS Hb for whom localised procedures are unsuitable or undesirable) were used in the absence of alternative sources of input data. The utility values obtained for patients with metastatic disease in the VHL RW QoL Disease Burden Study (0.412) are substantially lower to utility values previously accepted in metastatic RCC (0.772; TA830),[35] and metastatic pancreatic cancer (0.67; TA476).[105] However, the company considered that the target population in the current appraisal suffers from severe manifestations of VHL, which includes multi-systemic disease associated with a greater disease burden compared to metastatic disease related to one of the affected organs. Experts mentioned that VHL-related surgeries (particularly for RCC, CNS Hb, and retinal Hb), including surgery-related complications, and disability due to CNS Hb lesions, are expected to be among the most important drivers of QoL impairment. The economic model considers the direct impact of tumours on QoL by linking utility in the non-metastatic health

states to levels of tumour response. One potential issue is that the VHL RW QoL Disease Burden Study provides data self-reported by patients. Thus, the company indicated that there is potential for patients to confuse SD for PR in the recollection of their disease status, and that it was not possible for the company to validate this with clinicians or their notes (clarification question B21).[4] The company also argued that a spontaneous reduction in size of VHL-related tumours is highly unlikely without an active VHL treatment. For this reason and given the small number of patients who classified themselves as achieving PR, the company reclassified them as having achieved SD. Due to time constraints, the company were unable to rerun analyses in the VHL RW QoL Disease Burden Study to retrieve utility values before pooling the PR and SD categories. Finally, the impact of surgery-related complications on HRQoL was model through utility decrements. The utility decrement values used in the model were not validated by the company.

In response to clarification question B31,[4] the company explained that clinical expert feedback was considered to validate the residual efficacy of Belzutifan after treatment discontinuation and the approach for estimating survival after the metastatic health state has been reached. In relation to a potential residual treatment effect following Belzutifan discontinuation, two experts agreed that time to surgery will likely continue to be delayed compared to SoC for some time. This is based on the expectation that at the time of discontinuing Belzutifan the size of the patients' tumours is likely to be smaller than what it would have been if they were under SoC up to that point. Therefore, the expectation is that it will take a longer time for the largest tumour to reach the threshold size that warrants surgery, even if the growth rate (mm/year) immediately reverts to pre-treatment levels after Belzutifan discontinuation. Regarding survival in the metastatic health state, the experts agreed that it is reasonable to use evidence from trials assessing the effectiveness of first-line treatment in metastatic RCC to estimate transition probabilities from the metastatic disease health state to the death health state, even though these trials were conducted in a different metastatic RCC population (i.e., not specifically in VHL patients who developed metastatic RCC), for example TA830.[35] The TA830 was also used to validate the inclusion of Grade ≥3 AEs occurring in ≥5% of patients treated with Belzutifan or Grade ≥3 TRAEs occurring in >0% of patients treated with Belzutifan in the model.

Finally, on clarification question B23,[4] the EAG asked the company to provide further details on the numbers of metastatic patients not receiving active treatment, the sources used to inform these parameters, and on whether any validation steps were taken for these parameters. The company replied that the presented market shares in the VHL RCC cohort were obtained from TA830 (the NICE appraisal of pembrolizumab as an adjuvant treatment of RCC post-nephrectomy), NICE,[35] and for the VHL pNET cohort these were obtained from the ESMO clinical practice guidelines for gastroenteropancreatic neuroendocrine neoplasms,[36] and input from clinical experts. Patients are eligible to receive no active treatment only after receiving a first-line therapy for metastatic RCC, which is in line with the assumption made in TA830.[35]

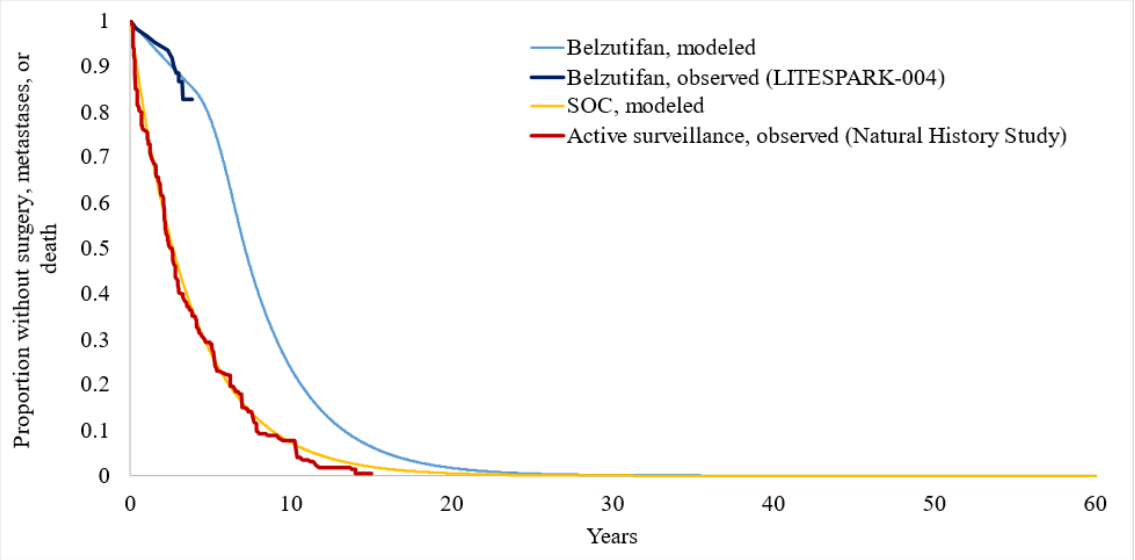
5.3.2.2 Model fit testing

In relation to model fit testing, the company acknowledged that there is no available evidence for the population in the DP to allow proper validation of the efficacy inputs included in the model and, therefore, this type of validation could not be conducted. As explained in previous Sections of this report, the EAG is concerned that this is a major limitation of this submission. Thus, even if input parameters for the trial population could be properly validated, there is no guarantee that the same would apply to the DP population.

Instead, the company discussed in the CS validation efforts conducted on the transition probabilities from the pre-surgery health state to surgery, metastatic disease, or death against the original sources

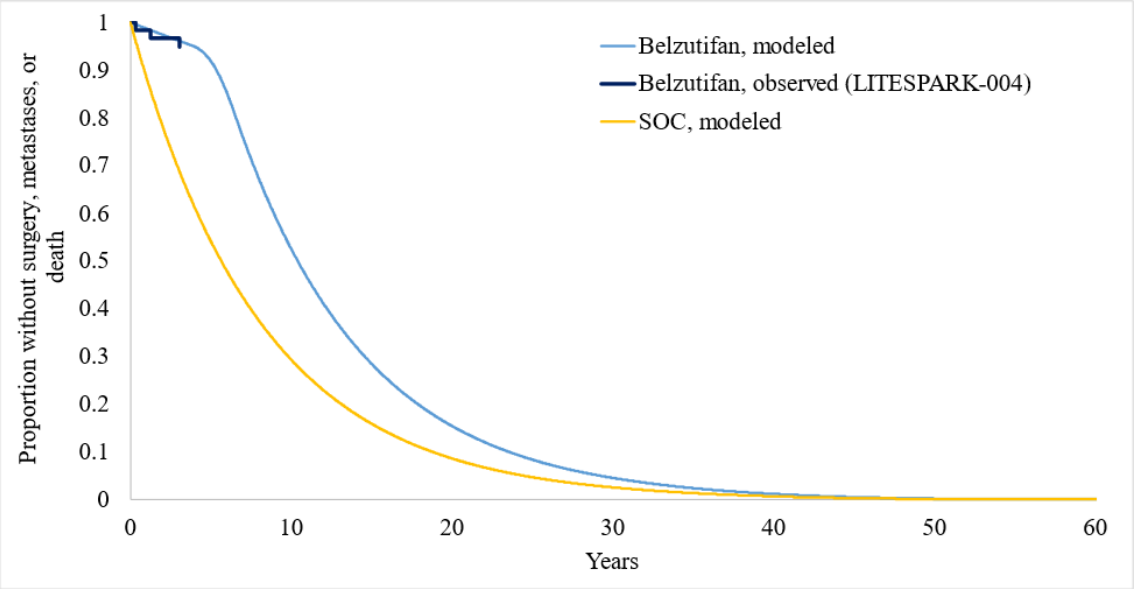
(thus on the MK-6482-004 trial population). Note that, in order to do this, the adjustments to account for real-world SoC and the assumption of immediate surgery for 90% (VHL RCC and VHL pNET cohorts) or 100% (VHL CNS Hb) cohorts were not considered. Survival curves showing long-term extrapolations of time to first surgery, metastases, or death, for Belzutifan and SoC in each VHL cohort are presented in Figures 5.10 to 5.14. Other than presenting these survival curves, and landmark estimates for time to first surgery, metastases, or death in Table 69 of the CS (table not shown here),[5] the company have not discussed how to interpret these data nor the implications for the validity of the fitted parameters. The EAG considers that, based on the presented evidence, the modelled time to surgery, metastases, or death in the SoC arm of the VHL RCC cohort fits reasonably well the active surveillance curve observed in the VHL Natural History Study (see Figure 5.10). However, the fit is worse for the CNS Hb cohort (see Figure 5.12) and especially poor for the pNET cohort (see Figure 5.14). This raises concerns about the validity of the SoC extrapolations of the pNET cohort. For the Belzutifan arm, given the immaturity of the data, it is challenging to draw reliable conclusions about long-term extrapolations. For the RCC cohort, the modelled time to surgery, metastases, or death seems to underestimate the short-term and overestimate the long-term trial observations. While the overall impact is unclear, in general long-term assumptions carry more weight on the model results than short-term ones. As it happened with the SoC arm, the CNS Hb and pNET cohorts (see Figures 5.11 and 5.13, respectively) are more challenging to interpret, especially the pNET cohort (see Figure 5.13) for which no events were observed in the trial. Again, this raises concerns about the validity of the Belzutifan extrapolations in general and of the pNET cohort in particular.

Figure 5.10: Validation of time to surgery, metastases, or death against MK-6482-004 trial data and VHL Natural History Study in the VHL RCC cohort



Based on Figure 20 in the CS.[5]
CS = company submission; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.11: Validation of time to surgery, metastases, or death against MK-6482-004 trial data (Belzutifan arm) in the VHL CNS Hb cohort

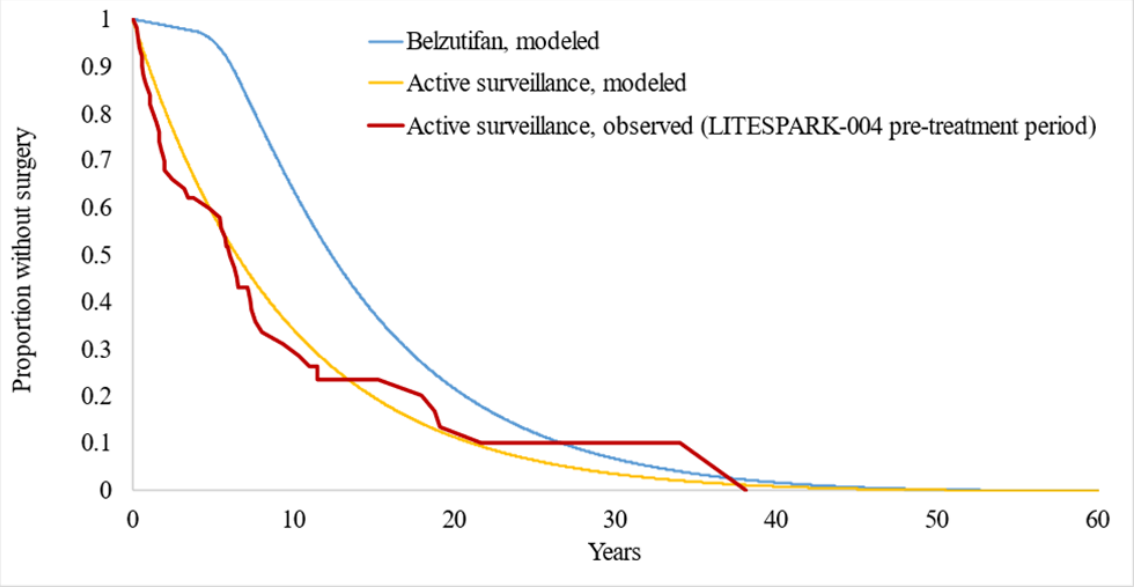


Based on Figure 20 in the CS.[5]

CNS Hb = central nervous system hemangioblastoma; CS = company submission; SoC = standard of care; VHL = Von Hippel-Lindau

LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.12: Validation of time to surgery, metastases, or death against VHL Natural History Study (SoC arm) in the VHL CNS Hb cohort†

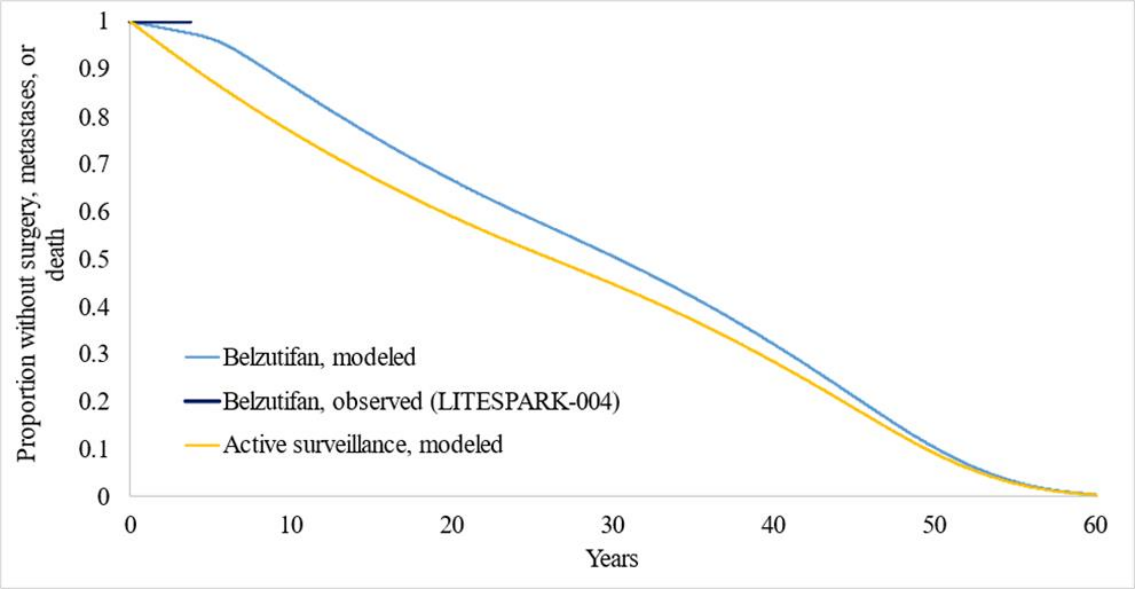


Based on Figure 20 in the CS.[5]

CNS Hb = central nervous system hemangioblastoma; CS = company submission SoC = standard of care; VHL = Von Hippel-Lindau

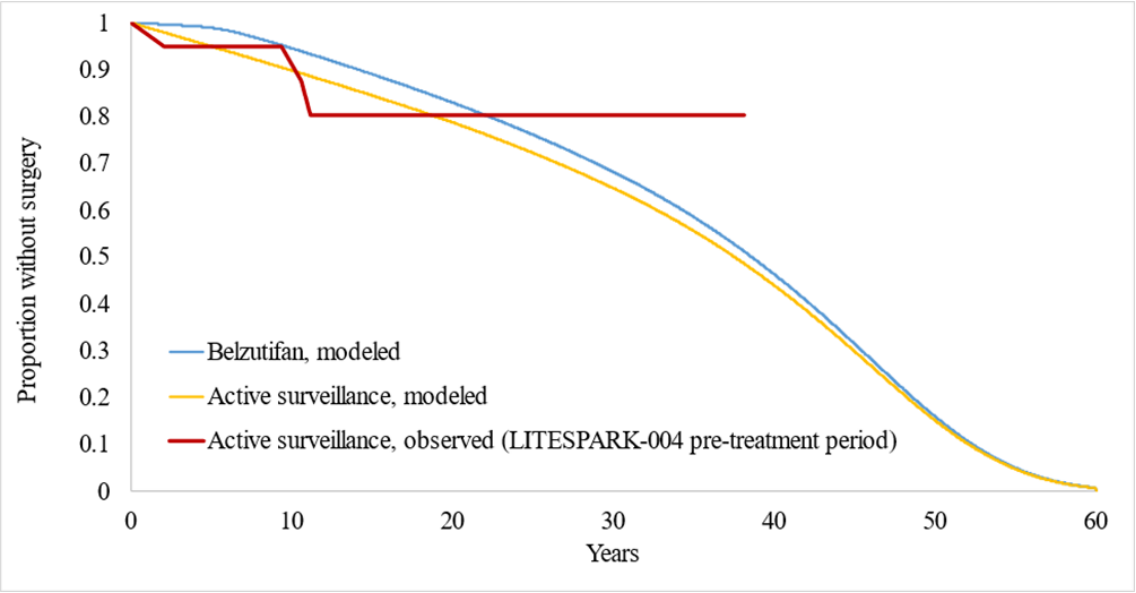
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.13: Validation of time to surgery, metastases, or death against MK-6482-004 trial data (Belzutifan arm) in the VHL pNET cohort



Based on Figure 20 in the CS.[5]
CS = company submission; pNET = pancreatic neuroendocrine tumours; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.14: Validation of time to surgery, metastases, or death against VHL Natural History Study (SoC arm) in the VHL pNET cohort†

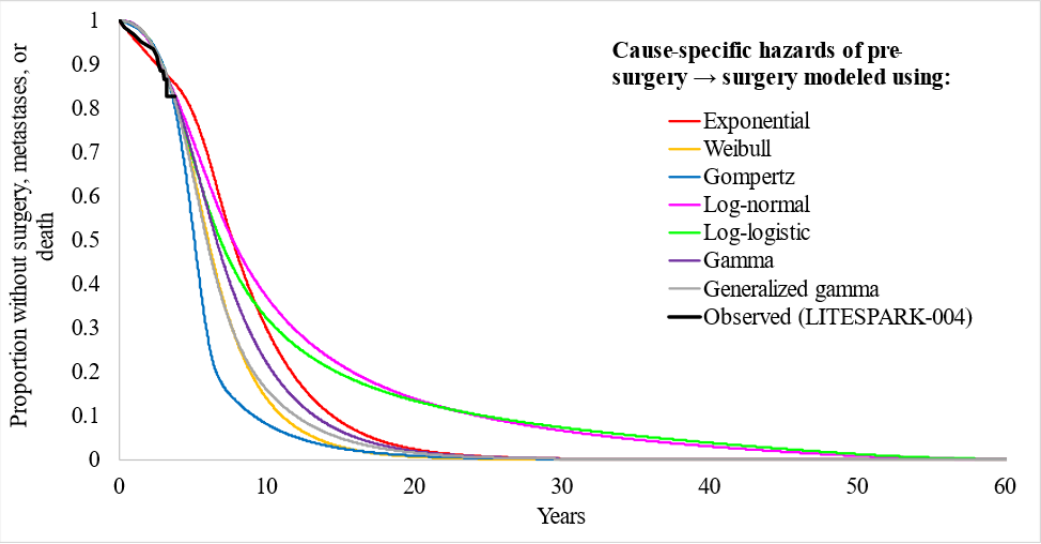


Based on Figure 20 in the CS.[5]
CS = company submission; pNET = pancreatic neuroendocrine tumours; SoC = standard of care; VHL = Von Hippel-Lindau LITESPARK-004 refers to MK-6482-004 trial.

In clarification question B6,[4] the EAG asked the company to provide more detailed survival analyses presented above and to justify the selection of the parametric extrapolations included in the economic model. The EAG was mostly satisfied with the company's approach, which followed the methodology described in NICE DSU TSD 14.[118] Plots of fitted versus observed time to surgery, metastases, or death are shown in Figures 5.15 to 5.20 for Belzutifan and SoC in the three cohorts of interest. The company

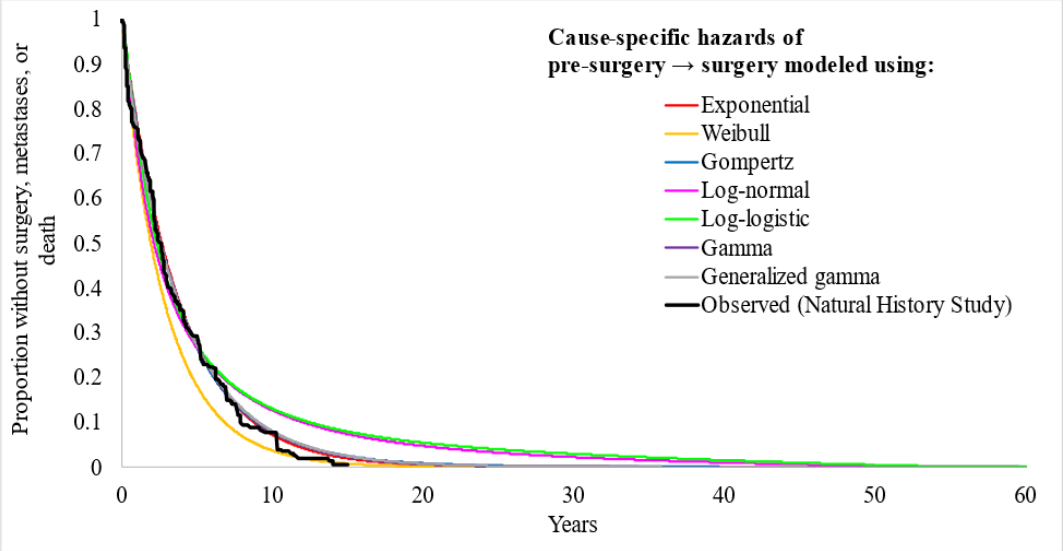
selected the exponential distribution for their base-case for the three cohorts. The EAG agrees with the choice of the exponential distribution as a pragmatic approach, but not because it is considered a suitable best candidate for both arms in the three cohorts. This is challenging to assess for the Belzutifan arm in general, given the lack of long-term data, and for both arms in the pNET cohort. Scenario analyses should be conducted to assess the impact of selecting different parametric extrapolations on the model results. It could be argued that due to different mechanism of action of Belzutifan and SoC (active surveillance), different type of parametric extrapolations would be allowed between the two treatment arms.

Figure 5.15: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL RCC cohort (Belzutifan arm)



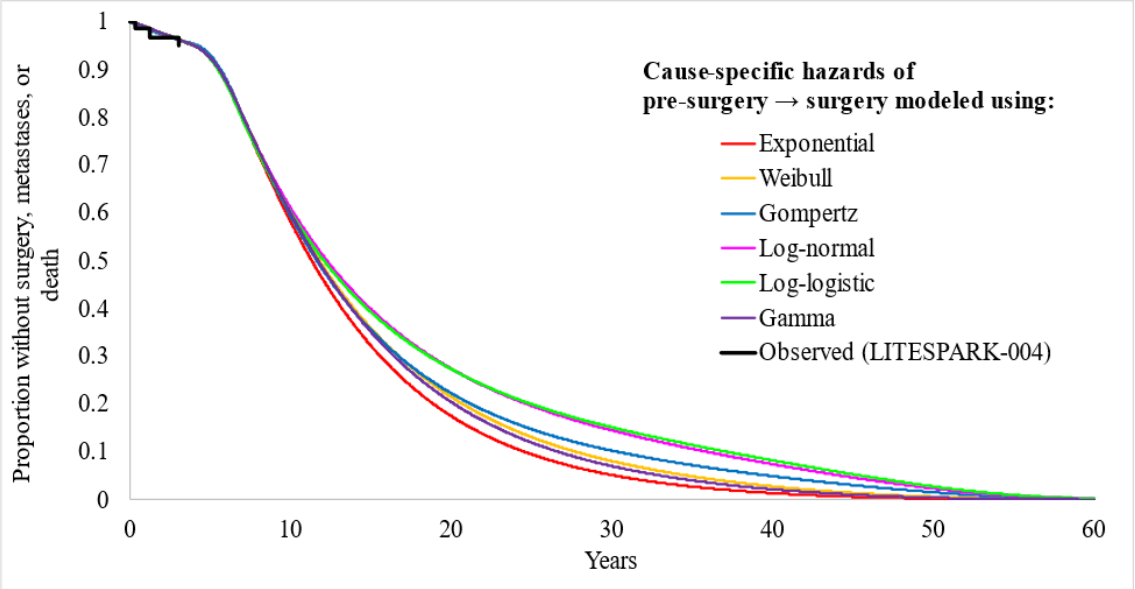
Based on Figure 1 in the clarification letter response.[4]
RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.16: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL RCC cohort (SoC arm)



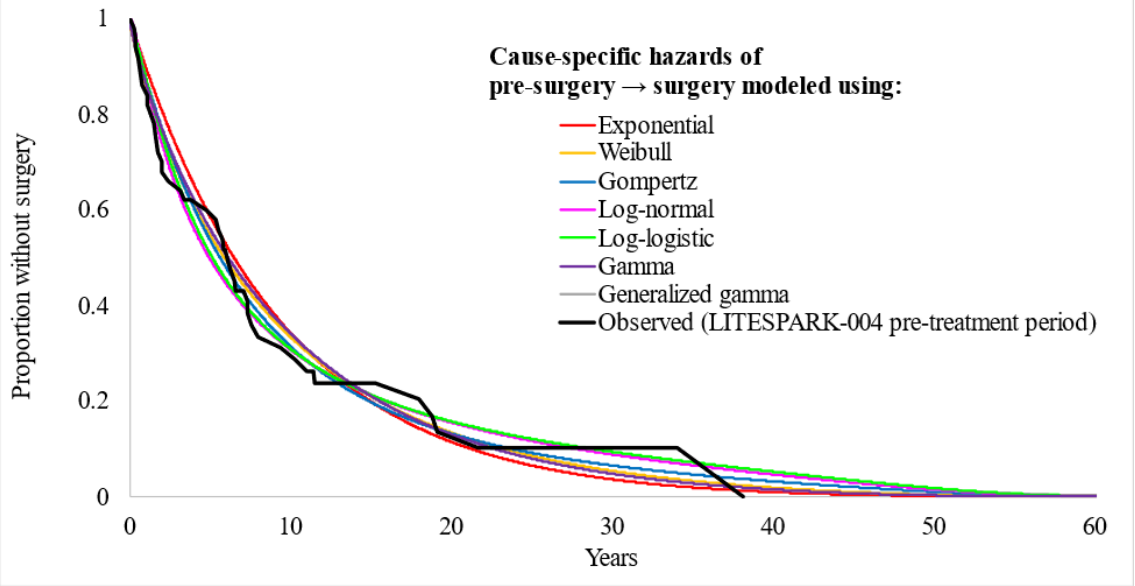
Based on Figure 1 in the clarification letter response.[4]
RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.17: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL CNS Hb cohort (Belzutifan arm)



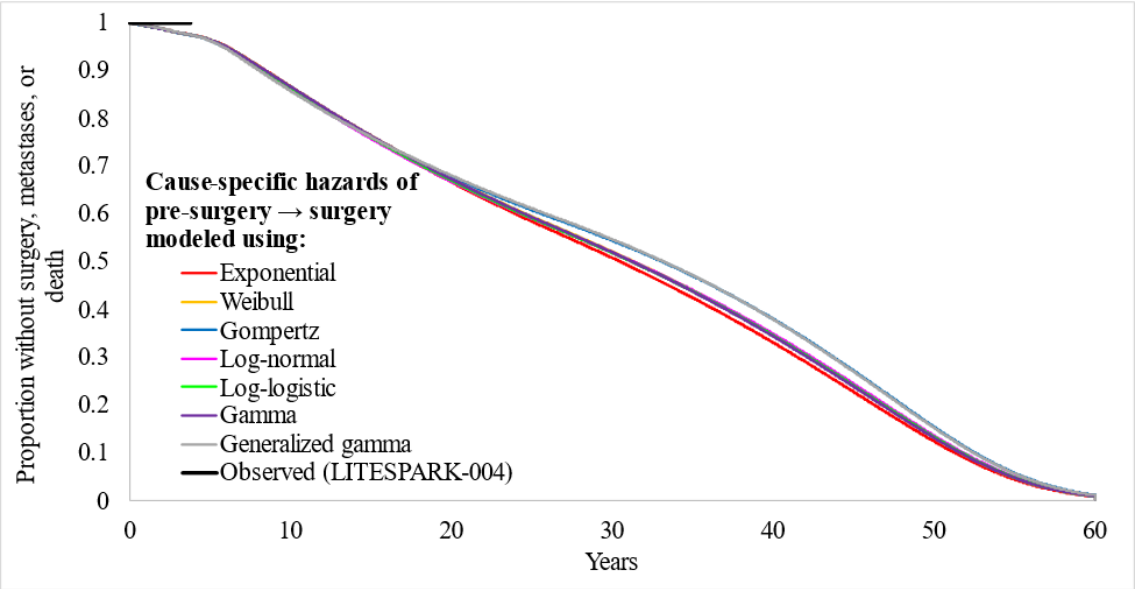
Based on Figure 1 in the clarification letter response.[4]
CNS Hb = central nervous system haemangioblastoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.18: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL CNS Hb cohort (SoC arm)



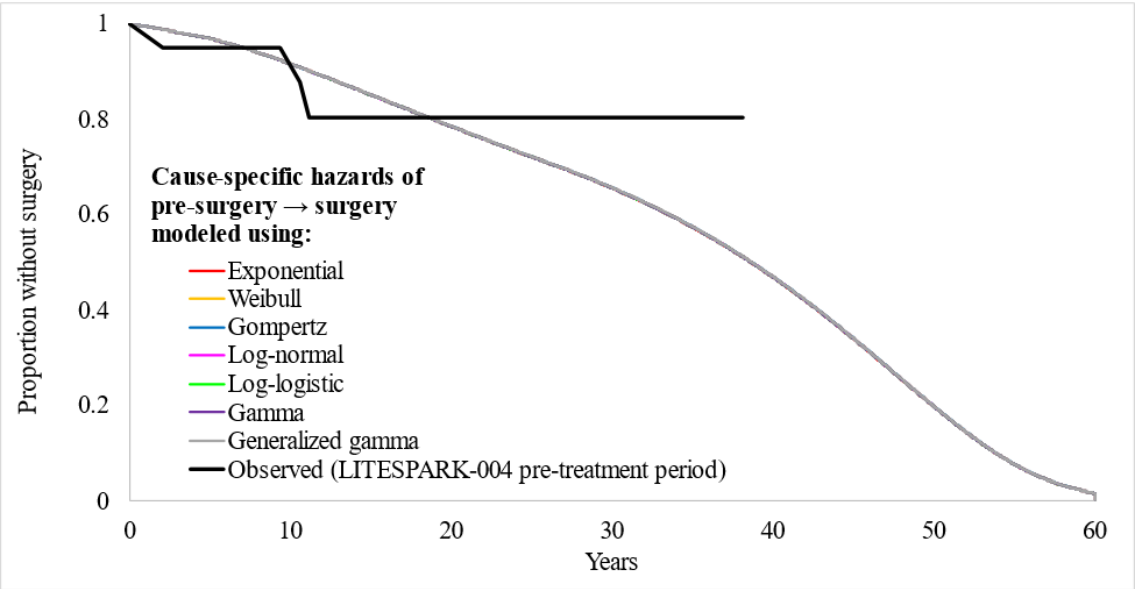
Based on Figure 1 in the clarification letter response.[4]
CNS Hb = central nervous system haemangioblastoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.
The cause-specific hazards of pre-surgery → metastases and pre-surgery → death is temporarily set to 0 when generating this figure, as the Kaplan-Meier curve from the MK-6482-004 pre-treatment period represents time from pre-surgery → surgery in the absence of any competing risks from pre-surgery → metastatic disease or pre-surgery → death.

Figure 5.19: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL pNET cohort (Belzutifan arm)



Based on Figure 1 in the clarification letter response.[4]
pNET = pancreatic neuroendocrine tumours; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.
There were no observed transitions from the pre-surgery state in MK-6482-004 as of the 01 April 2022 data cutoff date. Therefore, the seven candidate distributions were fitted to pre-surgery → surgery in the SoC arm, and a HR approach was applied in the Belzutifan arm.

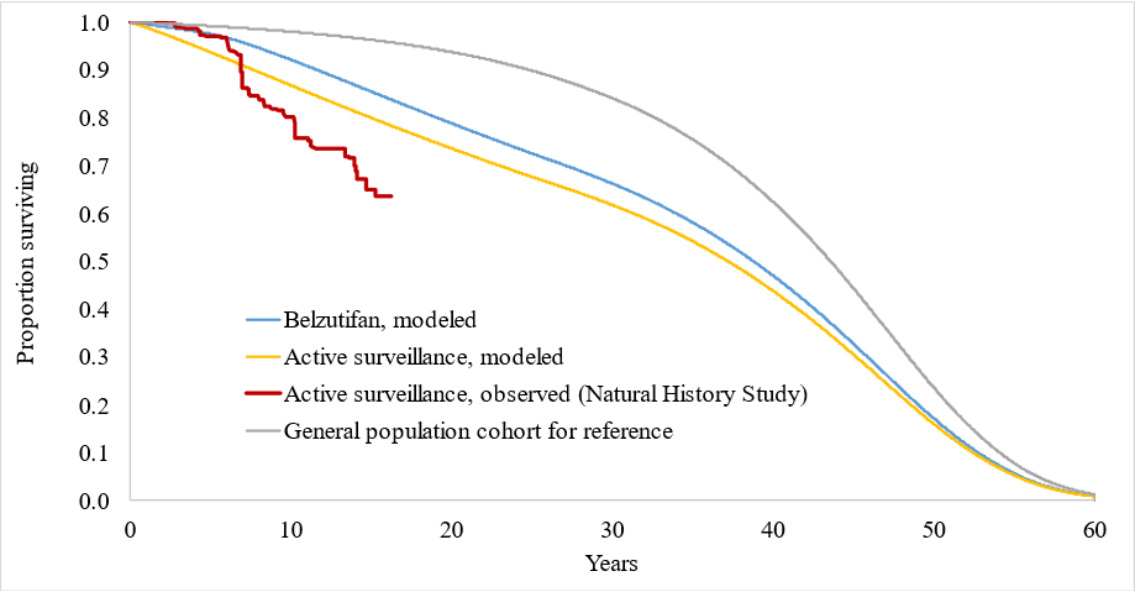
Figure 5.20: Fitted parametric models vs. observed time to surgery, metastases, or death in the VHL pNET cohort (SoC arm)



Based on Figure 1 in the clarification letter response.[4]
pNET = pancreatic neuroendocrine tumours; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.
The cause-specific hazards of pre-surgery → metastases and pre-surgery → death is temporarily set to 0 when generating this figure, as the Kaplan-Meier curve from the MK-6482-004 pre-treatment period represents time from pre-surgery → surgery in the absence of any competing risks from pre-surgery → metastatic disease or pre-surgery → death. Due to the small number of pNET surgeries observed during the pre-treatment period, all fitted distributions for pre-surgery → surgery appear similar to the exponential distribution, as shown above.

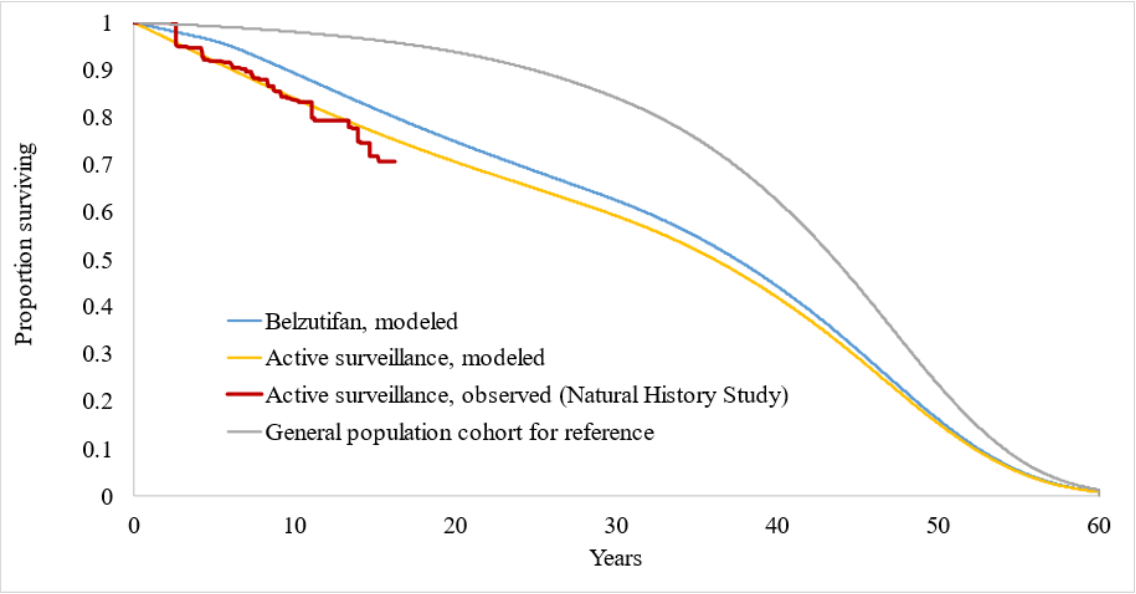
In a similar way, the company discussed in the CS validation efforts conducted on the long-term extrapolations for OS against the original sources (thus on the MK-6482-004 trial population). Also, in this case, the adjustments to account for real-world SoC and the assumption of immediate surgery were removed. Overall survival curves showing long-term extrapolations for Belzutifan and SoC in each VHL cohort are presented in Figures 5.21 to 5.23. Landmark estimates for OS are presented in Table 70 of the CS but they are not shown here.[5] The company explained that for the VHL RCC cohort, the modelled OS curve was plotted alongside the observed KM OS curve from the VHL Natural History Study cohort, after reweighting it to match key baseline characteristics of the MK-6482-004 trial population (see Figure 5.21). For the VHL CNS Hb and VHL pNET cohorts, the modelled OS curves were plotted alongside the observed KM OS curves from the VHL Natural History Study cohorts with a pre-index history of CNS Hb and pNET, respectively, after reweighting them to match key baseline characteristics in the corresponding subgroups of the MK-6482-004 trial population (see Figures 5.22 and 5.23, respectively). The company also explained that the modelled OS curves for SoC depend on all transition probabilities included in the model, including transition probabilities that were not estimated using VHL Natural History Study data (i.e., metastatic disease → death in all cohorts, and pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts), and that were not directly fitted to OS curves obtained from the VHL Natural History Study. Therefore, some divergence is expected between the observed and modelled OS curves for active surveillance shown in Figures 5.21 to 5.23. However, the company have not discussed how to interpret these divergences nor the implications for the validity of the modelled OS estimates. The company indicated that, as shown in Figure 5.21, the modelled OS for SoC in the VHL RCC cohort seems to clearly overestimate the long-term observed OS curve from the VHL Natural History Study. From this the company concluded that the modelled effectiveness results in the VHL RCC cohort may be conservative with respect to Belzutifan. The EAG agrees with the interpretation of the survival curves, but it is uncertain whether this and to what extent represents a conservative approach towards Belzutifan. As the company mentioned above, some divergence is expected between the observed and modelled OS curves, but the expected magnitude and impact of this divergence has not been discussed. Also, in the RCC cohort, the modelled OS seems to underestimate the short-term trial observations. As it happened with the modelled time to surgery, metastases, or death discussed above, the EAG considers that in general long-term assumptions carry more weight on the model results than short-term ones, but in conclusion the overall impact of a potential overestimation of modelled OS for SoC is unclear. Based on Figures 5.22 and 5.23, the company concluded that the modelled OS for SoC was better aligned with the observed OS from the VHL Natural History Study in the CNS Hb and pNET cohorts. The company explained that the larger difference observed in the VHL RCC cohort might be caused by the difference between the available treatments for advanced RCC included in the economic model compared to the treatments included in the VHL Natural History Study. In the model, these treatments were included based on NICE recommendations or because they are listed as a preferred or recommended first-line regimen according to the NCCN Guidelines. The company stated that this may not be reflective of the advanced RCC treatments in the VHL Natural History Study, and that for this reason, the modelled and observed OS are less likely to be aligned. For the VHL CNS Hb cohort this is not an issue since CNS Hb tumours cannot metastasise. For the VHL pNET cohort, the company mentioned that there have been fewer developments in treatments for metastatic disease recently and, therefore, those included in the model are more likely to be similar to those available to patients in the VHL Natural History Study. The EAG considers this explanation plausible, but it should be confirmed by clinical experts. Also, the implications for the model results are unclear.

Figure 5.21: Validation of overall survival against VHL Natural History Study (SoC arm) in the VHL RCC cohort



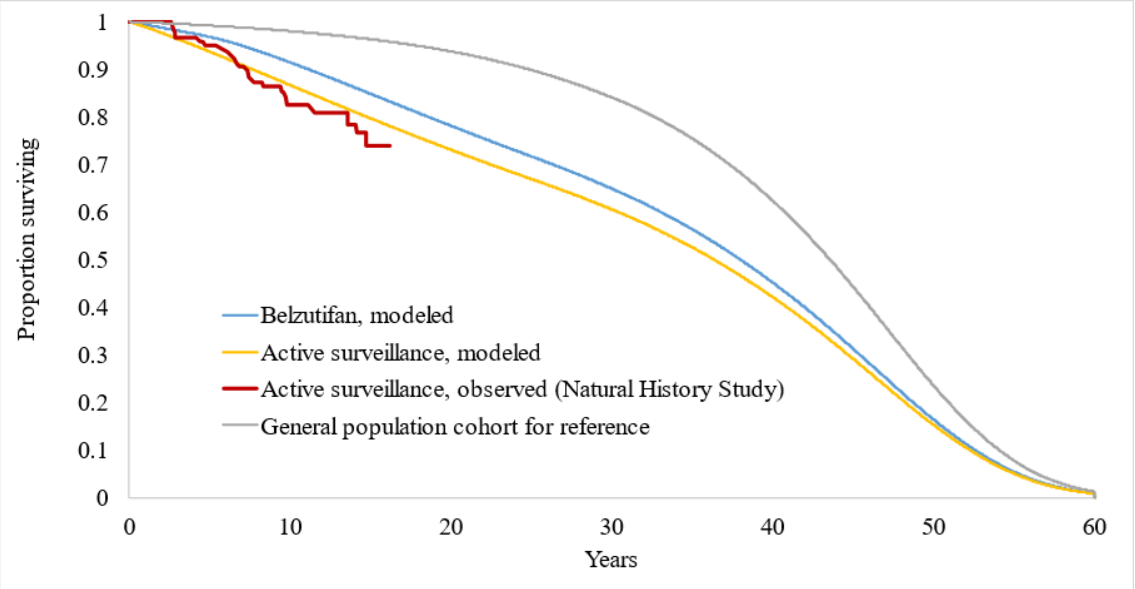
Based on Figure 21 in the CS.[5]
CS = company submission; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.22: Validation of overall survival against VHL Natural History Study (SoC arm) in the VHL CNS Hb cohort



Based on Figure 21 in the CS.[5]
CNS Hb = central nervous system Haemangioblastoma; CS = company submission; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

Figure 5.23: Validation of overall survival against VHL Natural History Study (SoC arm) in the VHL pNET cohort



Based on Figure 21 in the CS.[5]
CS = company submission; pNET = pancreatic neuroendocrine tumours; SoC = standard of care; VHL = Von Hippel-Lindau
LITESPARK-004 refers to MK-6482-004 trial.

5.3.3 Validation of the computerised model (technical verification)

5.3.3.1 External review

As explained in Section B.3.14 of the CS,[5] an early version of the model (thus, not the version used for this appraisal) was independently reviewed by external health economists at the [redacted]. However, no details on the validation efforts were reported in the CS.

5.3.3.2 Extreme value testing

No details about quality-control procedures for code verification were provided by the company. Therefore, it is unclear whether extreme value or other types of testing were performed on the model. These could have been conducted following the guidance of the Technical Verification (TECH-VER) tool for example.[119]

5.3.3.3 Testing of traces

Markov traces can be found in the model sheets named "Trace_TxReg1_RCC", "Trace_TxReg1_CNSHb", and "Trace_TxReg1_pNET" for each respective cohort. The model includes standard checks to test that the distribution of patients across health states always add up to 100%. No discussion about the face validity of the traces was provided by the company.

5.3.3.4 Unit testing

As mentioned above, the company did not provide details regarding the specific verification efforts conducted on the model. Therefore, it is unknown whether code verification included checks of the model results, calculations, data references, model interface, or Visual Basic for Applications code.

5.3.4 Operational validation (validation of model outcomes)

5.3.4.1 Face validity testing (model outcomes)

Although it is not explicitly mentioned in the CS or in the response to the clarification letter, the EAG is assuming that model results were presented to experts who provided some sort of validation.

5.3.4.2 Cross validation testing (model outcomes)

Comparisons with other technology appraisals

As mentioned above (validation of the conceptual model), previous NICE appraisals provide justification for some of the assumptions used in the economic analysis, but these mostly relate to the choice of input parameters, but not with the comparison of their results.

Comparisons with other models (not necessarily technology appraisals)

As mentioned above (validation of the conceptual model), this submission includes the first cost-effectiveness model of any treatment for VHL-associated RCC, pNET, and CNS Hb. Therefore, there is no evidence that can be used for cross-validation against other, independently developed economic models in the same indication.

5.3.4.3 Validation against outcomes using alternative input data

This type of validation was not explicitly reported by the company unless it was considered part of the scenario analyses.

5.3.4.4 Validation against empirical data

Comparison with empirical data used to develop the economic model (dependent validation)

As explained in Section B.3.14 of the CS,[5] the company indicated that external validity of the model was also assessed. Due to the lack of data to validate the outcomes for the population in the decision problem, the company assessed the modelled efficacy outcomes against the original sources that informed the efficacy inputs without the assumption of immediate surgery for the SoC arm or adjustments to reflect real-world SoC to allow for interpretable comparisons. These are discussed in Section 5.3.2.2. However, the EAG considers that this type of validation concerns the input parameters of the model rather than its outcomes.

Comparison with empirical data not used to develop the economic model (independent validation)

The company could not identify clinical trials or real-world evidence studies which could be used to externally validate the model outcomes reported in the current appraisal (Section B.3.14 of the CS).[5]

EAG comment:

The main concerns of the EAG relate to the many assumptions included in the model that can hardly be validated, which are due to the paucity of data. The EAG considers that this is inevitable with the current available evidence, which illustrates the large amount of uncertainty in this appraisal.

6. EVIDENCE ASSESSMENT GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

6.1.1 Explanation of the EAG adjustments

Table 6.1 summarises the key issues related to the CE categorised according to the sources of uncertainty as defined by Grimm et al. 2020:[120]

- Transparency (e.g., lack of clarity in presentation, description, or justification).
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case).
- Imprecision (e.g., particularly wide CIs, small sample sizes, or immaturity of data).
- Bias and indirectness (e.g., there is a mismatch between the DP and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered).
- Unavailability (e.g., lack of data or insight).

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the CE, whether it is reflected in the EAG exploratory analyses as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this report, the EAG was unable to define a new base-case. The EAG considered that the majority of the uncertainties identified in the CS cannot be resolved with the current evidence. The EAG believes that any alternative base-case scenario that could have been presented, would still be subjected to too many uncertainties and its results would thus be unreliable. Therefore, the EAG is afraid that a new base-case could give the wrong impression that it would be appropriate for the current DP. Instead, additional scenario analyses were explored by the EAG in order to assess the impact of some alternative assumptions on the current CE results. These scenarios included adjustments to the original company's base-case. These adjustments can generally be categorised following Kaltenthaler et al. 2016:[121]

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong).
- Fixing violations (FV) (correcting the model where the EAG considered that the NICE reference case, scope, or best practice had not been adhered to).
- Matters of judgement (MJ) (amending the model where the EAG considers that reasonable alternative assumptions are preferred).

6.1.1.1 Fixing errors

No errors were corrected by the EAG in the model provided by the company.

6.1.1.2 Fixing violations

1. Cost-effectiveness analyses should be based on subgroup-specific parameters, including QALY severity weighting.
   The EAG has no access to patient-level data to derive subgroup-specific input parameters. Only QALY severity weights were adjusted by the EAG.

6.1.1.3 Matters of judgement

2. Surgery rates observed in the MK-6482-004 trial (Belzutifan) might underestimate the surgery rates for the population in the DP.

The EAG doubled the transition probability from pre-surgery to surgery in the Belzutifan arm.

3. The percentages of immediate surgery in the SoC arm seem arbitrary and have a major impact on the model results.
   The EAG explored the scenario where patients are not receiving immediate surgery as part of SoC at all.
4. Doubling the perioperative mortality risk in the three cohorts seems arbitrary and has a major impact on the model results
   The EAG explored a scenario analysis in which the adjustment in the perioperative mortality risks was removed from the computations.
5. Increasing risks of risks of short- and long-term complications following surgery seems arbitrary and has a major impact on the model results.
   The EAG run two separate scenarios; one in which only adjustment of risks for end stage renal disease and/or dialysis in the VHL RCC cohort and secondary diabetes and immune-compromisation in the VHL pNET cohort, and cerebral vascular occlusion/stroke in the VHL pNET cohort were allowed in the model, and another one in which all adjustments were removed.
6. The comparator data might not be representative for the UK.
   The EAG explored a scenario in which the adjustment in the risk of surgery and metastasis based on Optum Clinformatics Data Mart Claims Study data is omitted.
7. Data to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable.
   Alternative parametric models for time to surgery, metastases, or death in all cohorts in both arms were explored by the EAG.
8. There is uncertainty in the derivation of the transition probabilities in the SoC arm.
   The EAG run a scenario in which the pre-surgery → surgery incidence rate in the SoC arm of the VHL CNS Hb and VHL pNET subgroups is doubled.
9. There is uncertainty in the implementation of time on treatment and treatment effect waning.
   Alternative parametric distributions for time on treatment and alternative duration of Belzutifan residual effect were explored by the EAG.
10. There is uncertainty in the derivation and implementation of HRQoL in the model.
    The EAG explored a scenario with an arbitrary 20% reduction in health utility values and another scenario incorporating the effect of time to response in the model.

6.1.2 EAG exploratory scenario analyses

The EAG conducted a series of scenario analyses to explore the impact of key assumptions and uncertainties within the CE analyses, focusing on the key issues described in Table 6.1. All EAG exploratory scenarios analyses are based on subgroup specific QALY weighting. A description of scenario analyses conducted by the EAG is provided below.

6.1.2.1 Scenario analyses set 1: mismatch between decision problem and trial population

The main concern of the EAG in this submission regarding the CE evidence relates to the mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses. The following scenarios were conducted to explore the impact of this source of uncertainty on the model results.

Surgery rates observed in the MK-6482-004 trial (Belzutifan) might underestimate the surgery rates for the population in the decision problem

Given the severity of the disease of those patients included in the DP, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial might underestimate the surgery rates for the population in the DP. The EAG doubled the transition probability from pre-surgery to surgery in the Belzutifan arm. However, the EAG would like to emphasise that in the absence of any other sources of evidence, the surgery rates values considered in this scenario was arbitrarily selected for explorative purposes only.

The percentages of immediate surgery in the SoC arm seem arbitrary and have a major impact on the model results

The assumed percentages of patients undergoing immediate surgery or receiving symptom management treatment in the SoC arm seem to be arbitrarily defined. Decreasing the percentages of patients undergoing immediate surgery in the SoC arm would increase the ICER. However, also in this case, the EAG would like to emphasise that any range of values considered in this scenario analysis would be arbitrarily selected for explorative purposes only. Therefore, the EAG decided to explore the scenario where patients are not receiving immediate surgery as part of SoC at all. This is thought to be a closer representation of the MK-6482-004 trial. As mentioned in Section 4.2.2 of this report, the EAG considers that the company's model is more suited to reflect the population recruited into MK-6482-004 trial (i.e., the initial MA VHL-associated RCC only) but cannot provide reliable estimates of the CE of Belzutifan compared to SoC in the population defined in the DP. Even though this scenario reflects a population that is different to that in the DP, the EAG considered that this scenario could still provide some useful insights to the Appraisal Committee.

Doubling the perioperative mortality risk in the three cohorts seems arbitrary and has a major impact on the model results

Perioperative mortality risks in each of the three cohorts were set equal between Belzutifan and SoC arms. For the VHL RCC and VHL CNS Hb cohorts, the perioperative mortality risk was 1.96% (1/51) and 1.82% (1/55), respectively, based on retrospective chart reviews, whereas for the VHL pNET cohort, the perioperative mortality risk was 1.7% (2/117), based on a multicentre international registry study. In the base-case analysis, the company doubled these risks to better reflect patients "for whom localised procedures are unsuitable or undesirable". The EAG considered this is an arbitrary adjustment and considering the mismatch between the population in the MK-6482-004 trial and the MHRA authorisation, it could be argued that all transition probabilities estimated from the MK-6482-004 trial should also be adjusted. A scenario analysis in which the adjustment in the perioperative mortality risks

was removed from the computations was explored in this scenario. This scenario, in addition, also removes all adjustments made to the transition probabilities and complication risks made by the company when attempting to better reflect the MHRA population.

Increasing risks of risks of short- and long-term complications following surgery seems arbitrary and has a major impact on the model results

The risks of short- and long-term complications following surgery were adjusted upwards in the company base-case compared to the estimates from the Optum Clinformatics Data Mart Claims Study to reflect patients "for whom localised procedures are unsuitable or undesirable". While the risks of most of the long-term surgery complications were doubled in the company base-case, for ESRD and/or dialysis in the VHL RCC cohort and secondary diabetes and immuno-compromisation in the VHL pNET cohort, the risks were respectively increased from 4% to 80%, from 20% to 100% and from 0% to 100% (see Table 4.16). Similarly, the risk of cerebral vascular occlusion/stroke was increased from 7.7.% to 85% of the patients in the VHL CNS Hb cohort. The company argued that these long-term metabolic consequences resulting from surgery were substantially adjusted upwards to capture the limited organ function following surgery in the licensed population. The EAG was concerned about the arbitrariness and validity of the adjustments implemented on these parameters. However, to distinguish the impact between the adjustments made for the other short- and long-term risk complications (doubling the risks) from the ones that may reflect immediate consequences of surgery resulting from partial/complete removal of the organs, the EAG run two separate scenarios; one in which only adjustments on the risks of these complications were allowed in the model, and another one in which all adjustments were removed.

6.1.2.2 Scenario analyses set 2: The comparator data might not be representative for the UK

The company noticed a difference in surgery rates for VHL-related tumours between the Optum Clinformatics Data Mart Claims database and the VHL Natural History Study. Following feedback from clinical experts, the company concluded that the cause of this difference is likely attributed to a less proactive surveillance system in the Optum Clinformatics Data Mart Claims database, indicating a lower level of disease control and, therefore, lower rates of surgeries and higher rates of metastatic disease than those observed in the VHL Natural History Study. Based on this, the company adjusted these rates to better reflect expectations regarding UK practice. The adjustment for VHL RCC cohort was implemented for both the surgery and metastases rates, whereas for both the VHL CNS Hb and VHL pNET cohorts adjustments were implemented only for the metastases rates. That was because the surgery transitions in the VHL CNS Hb and VHL pNET cohorts were modelled based on data from the pre-treatment period of the MK-6482-004 trial and, according to the company, those patients may not have received an elevated standard of care during this period as expected for patients in the VHL Natural History Study. The treatment benefit of Belzutifan compared to SoC was nevertheless maintained. Considering the different setting of the data collection (US versus UK) the EAG is uncertain on the validity of these assumptions. As mentioned in Section 4.2.2, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial reflect a lower bound of surgery rates for the population eligible to use Belzutifan in clinical practice. Therefore, lowering further the rates of surgeries would only reinforce the potential underestimation. The EAG considers it likely that more severe patients would be at greater risk of surgery and metastasis, and for this reason, in the scenario analysis removed the adjustment that was implemented in the company base-case analysis in the rates of surgeries and metastasis based on the Optum data. Furthermore, to reflect the impact of aligning adjustments between cohorts, the EAG presented a scenario in which the surgery rates were also adjusted for the VHL CNS Hb and VHL pNET cohorts at a similar relative reduction as the rates in the VHL RCC cohort.

6.1.2.3 Scenario analyses set 3: Data to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable

The company considered an exponential distribution to extrapolate the transition probability from pre-surgery to surgery in their base-case. This was selected for both the Belzutifan and SoC arms in the VHL RCC and CNS Hb cohorts. For the VHL pNET cohort, an exponential distribution was fitted to the SoC arm, and given the lack of observed events, for the Belzutifan arm an HR was applied. This HR was the one estimated for the pre-surgery to surgery transition in the VHL RCC cohort. In particular, the EAG considered the following scenarios:

- For the VHL RCC cohort in the Belzutifan arm we selected the curves providing the highest (log-logistic) and lowest (Gompertz) survival estimates according to Figure 5.15. For the SoC arm, all distributions seem to fit the data in a similar way according to Figure 5.16. The generalised Gamma was chosen for completeness.
- For the VHL CNS Hb cohort no extra scenarios were selected given that, based on Figures 5.17 and 5.18 all extrapolations are similar, and results are robust to changes in parametric distributions.
- For the pNET cohort different values HRs increased by a factor of 2, 5 and 10 were explored by the EAG.

6.1.2.4 Scenario analyses set 4: There is uncertainty in the derivation of the transition probabilities in the SoC arm

The company used the pre-treatment data from the MK-6482-004 trial to inform transitions from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm. The company acknowledged that this approach may be subject to biases due to the different data sources used to define transitions within the same cohort: the pre-treatment period of the MK-6482-004 trial and the VHL Natural History Study. The company also noted that there may also be implications when comparing these two cohorts with the VHL RCC cohort of the SoC arm for which transition probabilities were completely informed from the VHL Natural History Study. To illustrate the impact of using the pre-treatment period trial data, the company estimated the respective transition probabilities for the VHL RCC cohort using the pre-treatment data from MK-6482-004 trial instead of the VHL Natural History Study. When estimated using the pre-treatment period trial data, the rate of RCC surgeries decreased to more than half of the rate estimated from the VHL Natural History Study, confirming the EAG concerns on the appropriateness of using the MK-6482-004 pre-treatment period data to estimate the pre-surgery to surgery transitions for the VHL CNS Hb and VHL pNET cohorts. Based on the difference in the incidence rate for the VHL RCC cohort as estimated by the VHL Natural History Study and the pre-treatment period data from the MK-6482-004 trial (shown in Table 4.4), it could be expected that the pre-surgery → surgery incidence rate for the VHL CNS Hb and VHL pNET cohorts in the SoC arm might also be underestimated. To reflect the impact of this potential underestimation the EAG run a scenario in which the pre-surgery → surgery incidence rate in the SoC arm of the VHL CNS Hb and VHL pNET subgroups is doubled considering the rate of RCC surgeries was estimated to be more than double when using the VHL Natural History Study as compared to the pre-treatment period of the trial (shown in Table 4.6).

6.1.2.5 Scenario analyses set 5: There is uncertainty in the implementation of time on treatment and treatment effect waning

Different parametric models used to define ToT

The company considered a Gompertz distribution to extrapolate the time-to-treatment discontinuation in their base-case based on model fit evaluated through AIC/BIC statistics, visual inspection and clinical

plausibility. However, the EAG noticed that there is substantial uncertainty around the parametric distribution to model ToT. Therefore, to illustrate the impact of the parametric models used to extrapolate ToT, the EAG selected the curves providing the highest survival estimates (log-normal) according to Figure 4.2 and the curve with the second-best fit scores based on AIC/BIC values (Weibull), as the curve with the lowest survival estimates and lowest AIC/BIC values was used in the company base-case (Gompertz).

Different options for duration of residual benefit

The company set the duration of the residual benefit at 2.71 years, which could be seen as a reasonable choice for the base-case. However, it could also be argued that it could be different, especially for CNS Hb and pNET related tumours, for which no data are available. The survival analysis using data from the MK-6482-004 trial were estimated using a small number of observed events, imposing high uncertainty in the survival analysis and duration of Belzutifan treatment effect. For this reason, the EAG run scenario analyses varying the duration of the residual benefit also to lower values than the maximum duration of the trial (3.84 years) at 1.5 years and 3 years. Note that these values for the scenario analyses were selected arbitrarily for explorative purposes.

6.1.2.6 Scenario analyses set 6: There is uncertainty in the derivation and implementation of HRQoL in the model

Health state utility values decreased by 20%

As acknowledged by the company, the utility estimates from the patients in the VHL RW QoL Disease Burden Study are likely an overestimate of the utilities that would be obtained from the licensed population. Therefore, the EAG explored a scenario with an arbitrary 20% reduction of the health state utility values used in the base-case. This means that all utility values as presented in Table 4.18 were reduced with 20%, with exception of the complete response health state and PD for the VHL CNS Hb subgroup, since these values came from a different source and were assumed to be a good fit for this scenario.

Incorporate (median) time to treatment response to the QALY calculation

Patients in the Belzutifan arm obtain an immediate benefit in HRQoL compared to SoC from the first model cycle. It seems unrealistic that patient in the Belzutifan arm start benefiting from the treatment right from the beginning. As presented in Table 3.13 the median TTR was 11.1 months (range 2.7 to 30.5 months) for the VHL RCC cohort, 10.8 months (range 2.3 to 33.1) months for VHL CNS Hb cohort and 8.2 months (2.5 to 16.4 months) for the VHL pNET cohort. Therefore, including the time to response in the model is considered more appropriate by the EAG. The EAG did a scenario analysis incorporating this median time to treatment response by replacing the utility values of patients treated with Belzutifan by the utility for SoC corresponding with the right patient population (Table 4.18) for 48 cycles (11.1/12*52) in the VHL RCC cohort, 47 cycles in the VHL CNS Hb cohort and 36 cycles in the VHL pNET cohort. The cycle after the cycles required to incorporate the time to treatment response, the cohort specific utility value (see Table 4.18) reflecting the effect of the treatment with Belzutifan is used. Note that this is simply an illustrative scenario with the purpose to show the potential impact of including time to response in the model. A more sophisticated approach should be considered if TTR were available.

6.1.3 EAG subgroup analyses

No subgroup analyses (other than those defined by the population cohorts) were performed by the EAG.

Table 6.1: Overview of key issues related to the CE

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
There is a mismatch between the decision problem and evidence used to inform it in terms of population	2.1 4.2.2 4.2.3 4.2.4 4.2.6.5 4.2.6.7	Bias & indirectness Unavailability Transparency	Alternative transition probability from pre-surgery to surgery in the Belzutifan arm. Alternative percentages of patients undergoing immediate surgery in the SoC arm. Alternative assumptions on perioperative mortality risks, and other surgery-related complications.	+/-	No/Explored	Clinical effectiveness data on the decision problem population, allowing classification by primary tumour. Evidence supporting the definition of SoC (immediate surgery, risks associated to surgery, etc.). The potential "harm" for Belzutifan patients for not having immediate surgery and the potential "benefit" for SoC patients for having immediate surgery should be captured in the model.
The comparator data might not be representative for the UK	4.2.6.4	Bias & indirectness Unavailability Methods	Omit the adjustment in the risk of surgery and	+/-	No/Explored	Data from UK clinical practice.

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
			metastasis based on Optum data.			
Data to inform Belzutifan arm (MK-6482-004 trial) immature or unavailable	4.2.6 5.3.2	Imprecision Unavailability	Alternative fitted parametric models for time to surgery, metastases, or death in all cohorts in both arms	+/-	No/Explored	Long-term survival data for all three cohorts.
There is uncertainty in the derivation of the transition probabilities in the SoC arm	4.2.6.2	Transparency Imprecision Unavailability	Alternative transition probability from pre-surgery to surgery in the SoC arm.	+/-	No/Explored	Clinical effectiveness data on the decision problem population, allowing classification by primary tumour (SoC arm).
There is uncertainty in the implementation of time on treatment and treatment effect waning	4.2.6.7	Imprecision	Different parametric models used to define ToT. Different options for duration of residual benefit.	+/-	No/Explored	Long-term data
There is uncertainty in the derivation and implementation of HRQoL in the model	4.2.8.1	Bias Unavailability	Scenario with reduction in	+/-	No/Explored	Patients need to be classified in response

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
			health utility values. Incorporate the effect of time to response in the QALY calculations.			categories based on a clinical diagnosis by a physician (not self-reported). Requires additional evidence about the HRQoL of partial response patients. Required addition evidence about the UK VHL disease patient population.
Cost-effectiveness analyses should be based on subgroup-specific parameters (including QALY severity weighting)	4.2.3 4.2.10	Methods	Estimate input parameters using subgroup-specific data	+/-	No (subgroup specific QALY weighting used by EAG)	No, but input parameters that change per cohort should be re-estimated

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the EAG and '+' indicates that the EAG believes this issue likely induces bias in favour of the intervention versus at least one comparator
CE = cost effectiveness; EAG = Evidence Assessment Group; FE = fixing errors; FV = fixing violations; MJ = matters of judgement; ICER = incremental cost-effectiveness ratio

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the EAG

6.2.1 Scenario analyses set 1: mismatch between decision problem and trial population

6.2.1.1 Alternative surgery rates in the Belzutifan arm

In this additional scenario, the EAG doubled the transition probability from pre-surgery to surgery in the Belzutifan arm. Results are presented in Table 6.2. As expected, compared to the base-case in Table 5.1, the ICERs increased for all three cohorts, especially for the VHL RCC cohort (approximately £8,000 increased). In this scenario, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained. The lowest ICER was for the CNS Hb cohort which was £39,988 per QALY gained.

Table 6.2: Scenario with doubled transition probability from pre-surgery to surgery in Belzutifan – deterministic CE results (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	81,594	59,761
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	58,351	39,988
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	78,066	63,334

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SoC = standard of care; VHL = Von Hippel-Lindau

6.2.1.2 No immediate surgery in the comparator arm

The scenario where patients are not receiving immediate surgery as part of SoC is thought to be a closer representation of the MK-6482-004 trial. The deterministic results of this scenario are presented in Table 6.3. These indicated that Belzutifan was more costly and more effective than SoC in all cohorts. Compared to SoC, in the RCC cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER in the RCC cohort was £164,169 per QALY gained when no immediate surgery in the SoC arm is included. In the CNS Hb cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs compared to SoC. Therefore, the ICER in the CNS Hb cohort was £159,104 per QALY gained when no immediate surgery in the SoC arm is included. Finally, in the pNET cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER in the pNET cohort was £159,682 per QALY gained when no immediate surgery in the SoC arm is included. A PSA was also run as part of this scenario. However, its results are not shown here given the high ICERs, and that all PSA outcomes were in the NE quadrant of the CE plane,

but clearly above the commonly used threshold ICER of £30,000 per QALY gained. The PSA outcomes were nevertheless used to calculate severity weights for QALYs. Based on this distribution the severity adjusted ICERs are equal to £140,023, £136,078, and £138,511 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. For the three cohorts, with and without severity weighting, the probability that Belzutifan is cost effective compared to SoC at threshold ICER of £30,000 per QALY gained was [redacted]%.

Table 6.3: Scenario with no immediate surgery in SoC – deterministic CE results (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	164,129	140,023
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	159,104	136,078
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	159,682	138,511

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

As an attempt to gain insight into the effect of immediate surgery on the CE results, we present in Tables 6.4 to 6.6 disaggregated results for the comparison of SoC with and without immediate surgery. It can be observed in Table 6.4 that in terms of LYs, the option of SoC without no immediate surgery results in longer LYG. Even though the difference in total LYs is small in the three cohorts, these results do not seem completely rational since they imply that patients getting immediate surgery would have a shorter life expectancy. When looking at the QALY results in Table 6.5, the difference is even larger when patients do not have immediate surgery in SoC. The EAG understands that, in the model, patients undergoing surgery are at risk of experiencing surgical complications and this is implemented by applying disutilities to these complications. This difference in disutilities post-surgery is the main driver of the difference in QALYs with and without immediate surgery. It is unclear why patients should receive immediate surgery as last resort intervention, when patients do much better without it according to the model results. Again, this does not seem rational: if by not getting immediate surgery patients get better outcomes, why should patients undergo immediate surgery at all? Also, in terms of costs, the more patients going through immediate surgery the more the total costs in the SoC arm as shown in Table 6.6. These results reinforce the EAG's idea that the severity of the decision problem population has not been appropriately captured by the company's model.

Table 6.4: Disaggregated LYs results – SoC with and without no immediate surgery assumed

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)
Total life years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS and company electronic model,[5, 114]

CNS Hb = central nervous system hemangioblastoma; CS = company submission; LYs = life years; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

Table 6.5: Disaggregated QALYs results – SoC with and without no immediate surgery assumed

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Event-free after surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metastatic disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgical complication disutility for other tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Caregiver disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)
Age-related disutility	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS and company electronic model,[5, 114]
AE = adverse event; CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

Table 6.6: Disaggregated cost results (£) – SoC with and without no immediate surgery assumed

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Belzutifan treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Advanced treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs for primary tumour	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery and surgical complication costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes	VHL RCC cohort		VHL CNS Hb cohort		VHL pNET cohort	
	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)	SOC (no immediate surgery)	SOC (immediate surgery)
for other tumours						
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 156 in Appendix J of the CS and company electronic model,[5, 114]
AE = adverse event; CNS Hb = central nervous system hemangioblastoma; CS = company submission; pNET = pancreatic neuroendocrine tumours; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.1.3 Omit adjustment in the perioperative mortality risks

The results of the scenario in which the company's adjustment in the risk of perioperative mortality (increased by a factor of 2.0), and other adjustments made to the transition probabilities and complication risks made when attempting to better reflect the MHRA population, are omitted are presented in Table 6.7. Compared to the company base-case, the ICER in the VHL RCC cohort was approximately 2.2 times higher, while the impact for the VHL CNS Hb and the VHL pNET cohorts was slightly lower, but still substantial, namely 2.0 and 1.7 times higher in the VHL CNS Hb and the VHL pNET cohort, respectively. In this scenario, with and without severity weight, all ICERs were well above the commonly used threshold ICER of £30,000 per QALY gained. The lowest ICER was for the CNS Hb cohort which was £95,110 per QALY gained.

Table 6.7: Scenario removing adjustment in the perioperative mortality risks (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	160,629	139,455
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	115,926	95,110
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	129,453	107,964

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.1.4 Short- and long-term complications following surgery were set equal to the risks estimated from the Optum study

The results of the scenario in which all adjustments in the risk of short-and long-term complications are omitted, including the adjustment on the risk of perioperative mortality, are presented in Table 6.8. Compared to the company base-case, the ICER in the VHL RCC cohort was 2.3 times higher, while the impact for the VHL CNS Hb and the VHL pNET cohorts was slightly lower (2.1 times higher ICER in the VHL CNS Hb and 1.7 times higher in the VHL pNET cohort). Table 6.9 presents the results of the

scenario in which adjustments of certain risks that may reflect immediate consequences of surgery resulting from partial/complete removal of the organs were allowed. Compared to the company base-case, the ICER in the VHL RCC cohort increased by about £1,500 per QALY gained, while the impact was greater in the VHL CNS Hb and the VHL pNET cohorts, for which the ICERs increased by approximately £7,000 and £5,000 per QALY gained, respectively. In these two scenarios, with and without severity weight, all ICERs were still well above the commonly used threshold ICER of £30,000 per QALY gained. The lowest ICER was for the CNS Hb cohort in the scenario allowing only for adjustment of risks for ESRD and/or dialysis in the VHL RCC cohort and secondary diabetes and immune-compromisation in the VHL pNET cohort, and cerebral vascular occlusion/stroke in the VHL pNET cohort, which was £47,375 per QALY gained. These results indicate that a key driver in the adjustments of short- and long-term complications risks is the doubling in the risk of perioperative mortality. The impact of the adjustments implemented in the other short- and long-term complications risks is minor compared to the risks of perioperative mortality.

Table 6.8: Scenario removing all adjustments on the short- and long-term complications following surgery (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	166,133	141,668
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	117,145	96,435
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	131,043	109,919

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

Table 6.9: Scenario allowing only for adjustment of risks for ESRD and/or dialysis in the VHL RCC cohort and secondary diabetes and immune-compromisation in the VHL pNET cohort, and cerebral vascular occlusion/stroke in the VHL pNET cohort (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	74,881	55,623
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	64,124	47,375
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	82,773	67,530

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = Pancreatic Neuroendocrine Tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.2 Scenario analyses set 2: The comparator data might not be representative for the UK

The scenario results in which the adjustment in the risk of surgery and metastasis based on Optum data is omitted are presented in Table 6.10. Compared to the company base-case, the ICER in the VHL RCC cohort increased by about £6,000 per QALY gained for the VHL RCC cohort, while they decreased by about £7,000 and £22,000 per QALY gained for the VHL CNS Hb and the VHL pNET, respectively. The scenario results in which the adjustment in the risk of surgery implemented in the VHL RCC cohort based on Optum data is also implemented in the VHL CNS Hb and VHL pNET cohorts are presented in Table 6.11. Compared to the company base-case, the ICER in the VHL CNS Hb cohort decreased by about £7,000 per QALY gained, while it decreased by about £2,000 per QALY gained in the VHL pNET, cohort. Also, in these two scenarios, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained. The lowest ICER was for the CNS Hb cohort in the scenario including adjustment in the risk of surgery for the VHL CNS Hb and the VHL pNET cohorts equal to the respective adjustment in the VHL RCC cohort in both treatment arms based on the Optum data, which was £30,107 per QALY gained.

The adjustments implemented in the rates of surgeries and metastases are expected to influence outcomes in the opposite direction. That is because adjustments were attributed to an elevated standard of care in the VHL Natural History Study and the MK-6482-004 trial, meaning that in reality the company stated that they would expect lower rates of surgeries and higher rates of metastatic disease. Therefore, lower rates of surgeries would translate into an increased potential benefit for Belzutifan treatment, whereas a higher rate of metastatic disease could mean a lower potential benefit for Belzutifan treatment. For the RCC cohort, the downward adjustment of surgery rates seems to be surpassing the ICER increase from the upward adjustment of metastases rates resulting in a lower base-case ICER than the ICER presented below when the adjustment is omitted. For this cohort Belzutifan

treatment is therefore producing an increased health benefit with the implementation of the adjusted parameters. For the VHL CNS Hb and VHL pNET cohorts the upward adjustment to metastases rates in the base-case analysis of the company produced a higher ICER than if the adjustment is omitted (due to an increased metastases rates in the base-case calculations) which in turn indicates a reduced potential benefit for Belzutifan treatment. When implementing the relative reduction in the risk of VHL RCC surgeries on the surgery rates of the VHL CNS Hb and VHL pNET cohorts (in addition to adjustments on metastatic disease that were already implemented), the ICER of the VHL CNS Hb and VHL pNET cohorts improved as expected though without substantial changes compared to the company base-case.

Table 6.10: Scenario omitting adjustment in the i) pre-surgery → surgery, ii) pre-surgery → metastatic disease, and iii) event-free after surgery → metastatic disease transitions in both treatment arms based on the Optum data (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	75,814	56,293
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	49,901	34,174
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,768	50,336

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

Table 6.11: Including adjustment in the risk of surgery for the VHL CNS Hb and the VHL pNET cohorts equal to the respective adjustment in the VHL RCC cohort in both treatment arms based on the Optum data (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	51,996	35,800
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	77,202	62,440

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL CNS Hb cohort								

* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.3 Scenario analyses set 3: Alternative fitted parametric models for time to surgery, metastases, or death in all cohorts in both arms

Results for the VHL RCC cohort are presented in Table 6.12. These results indicate that results can be sensitive to the distribution chosen for the Belzutifan arm, illustrating in this way the uncertainty associated to the immaturity of the data (see e.g., Figure 5.15). When a Gompertz distribution was selected for the Belzutifan arm, the ICER increased in almost £20,000 per QALY gained with respect to the base-case ICER.

Results for CNS Hb cohort are not shown here because these were robust to changes in the selected distributions. However, the EAG considers that this does not mean that there is no uncertainty associated with these extrapolations. In fact, Belzutifan data are very immature, but the extrapolations are quite similar, as shown in Figure 5.17. This could give the wrong impression that collecting additional data for Belzutifan in the CNS Hb cohort is not a priority, whereas having data could potentially change all current survival data extrapolations.

A similar model behaviour was observed for the VHL pNET cohort since the model seems to be insensitive to changes in the HR pre-surgery to surgery estimated in the RCC cohort, which is the one used to model the treatment effect in the pNET cohort too. This, again, could be giving the wrong impression that there is no uncertainty associated to the VHL pNET cohort, yet this is the cohort with the most limited data, but its results are the most stable of the three cohorts. The model is nevertheless rather insensitive to changes in the HR. When the HR is increased by a factor of 2 (HR = 0.29) the ICER was £78,049 and, when increased by a factor of 5 (HR = 0.73), the ICER was £79,244. This raised some additional concerns about the validity of the results for the VHL pNET cohort. Therefore, the EAG explored a more extreme, and possibly unrealistic, scenario and considered an HR >1, by using a factor of 10. Using this factor, the corresponding HR was 1.47. Therefore, in this scenario SoC is expected to be more effective than Belzutifan. However, the ICER was £81,204, with 3.34 additional QALYs gained by Belzutifan. The EAG understands that the model is complex and that this is not the only effectiveness parameter driving the results. These results for example are still influenced by the assumption of immediate surgery. However, even when this is assumed to be 0% in the pNET cohort, the model still predicts 1.99 additional QALYs gained for Belzutifan, with an HR >1 (SoC is more effective in terms of the probability of surgery. Even though the ICER in this case was as high as £194,758, these results seem to be lacking face validity.

Table 6.12: Scenario with alternative fitted parametric models for time to surgery, metastases, or death in all cohorts in both arms (Belzutifan list price) – VHL RCC cohort

Scenario	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
Base-case (Belzutifan and SoC exponential)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	73,095	49,359
Belzutifan log-logistic								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	78,672	56,853
Belzutifan Gompertz								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	91,874	66,380
SOC Generalised Gamma								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	70,508	51,000

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.4 Scenario analyses set 4: Risk of pre-surgery → surgery in SoC arm doubled in the VHL CNS Hb and VHL pNET cohorts

The results of the scenario in which the risk of surgery in the SoC arm of the VHL CNS Hb and VHL pNET cohorts were doubled are presented in Table 6.13. Results can be sensitive to the risk of surgery in the VHL CNS Hb cohort but less sensitive for the VHL pNET cohort. When the risk of surgery was doubled, the ICER increased by almost £15,000 per QALY gained with respect to the base-case ICER in the VHL CNS Hb cohort, whereas the respective change in the VHL pNET cohort increased the ICER by £3,000 per QALY gained only.

Table 6.13: Scenario with risk of pre-surgery → surgery in SoC arm of the VHL CNS Hb and VHL pNET cohorts increased by 100% (Belzutifan list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	71,540	49,152
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL CNS Hb cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	79,565	64,577

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.5 Scenario analyses set 5: There is uncertainty in the implementation of time on treatment and treatment effect waning

6.2.5.1 Different parametric models used to define ToT

The results of the scenario using different parametric models for ToT are presented in Table 6.14. Compared to the company base-case, the ICER in the VHL RCC cohort increased by almost £20,000 and £63,000 when using the Weibull and log-normal respectively to extrapolate ToT. In the VHL CNS Hb cohort, the ICER increased by almost £15,000 and £59,000 when using the Weibull and log-normal distributions, respectively. In the VHL pNET cohort, the ICER increased by almost £19,000 and £60,000 when using the Weibull and log-normal distributions, respectively. These results indicate that results can be sensitive to the distribution chosen for ToT.

Table 6.14: Scenario with alternative fitted parametric models for ToT in all cohorts in both arms (Belzutifan list price)

Scenario	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort (Weibull)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	91,265	66,073
VHL CNS Hb cohort (Weibull)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	71,497	49,441
VHL pNET cohort (Weibull)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	96,471	78,588
VHL RCC cohort (log-normal)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	136,345	98,716
VHL CNS Hb cohort (log-normal)								
SOC	[redacted]	[redacted]	[redacted]					

Scenario	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort (Weibull)								
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	106,403	73,588
VHL pNET cohort (Log-normal)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	137,719	112,183

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; ToT = time on treatment; VHL = Von Hippel-Lindau

6.2.5.2 Different options for duration of residual benefit

The results of the scenario altering the residual health benefit are presented in Table 6.15. Compared to the company base-case, the ICER increased by almost £4,000 in the VHL RCC cohort, remained almost unaltered in the VHL CNS Hb cohort, while it also increased by about £5,000 in the VHL pNET cohort when reducing the residual benefit to 1.5 years. When increasing the residual benefit to 3 years the ICERs in all three cohorts improved by about £1,000 per QALY gained.

Table 6.15: Scenario with alternative duration for residual benefit (Belzutifan list price)

Scenario	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort (1.5 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	77,222	55,638
VHL CNS Hb cohort (1.5 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	59,688	40,779
VHL pNET cohort (1.5 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	82,585	67,056
VHL RCC cohort (3.0 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	72,101	52,035
VHL CNS Hb cohort (3.0 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	56,253	38,422

Scenario	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort (1.5 years)								
VHL pNET cohort (3.0 years)								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	76,455	62,008

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.6 Scenario analyses set 6: There is uncertainty in the derivation and implementation of HRQoL in the model

6.2.6.1 Health state utility values decreased by 20%

The 20% utility value reduction resulted, as expected, in all cohorts to a lower total QALY and consequently a lower incremental QALY as shown in Table 6.16. Since the costs remained the same, the ICER increases in all cohorts. For the VHL RCC cohort the ICER went up with £16,000, for the VHL-CHN Hb cohort the ICER went up with £16,500 and for the VHL pNET cohort the ICER went up with £7,700.

Table 6.16: Scenario with a 20% reduction of the utility values based on the VHL RW QoL Disease Burden

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	89,281	61,270
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	73,435	50,324
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	85,368	65,870

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.2.6.2 Incorporate (median) time to treatment response to the QALY calculation

As expected, delaying the treatment effect according to the cohort specific median time to treatment response resulted in slightly lower total QALYs for the cohorts of patients in the Belzutifan arm and consequently a slightly higher ICER (Table 6.17). For the VHL RCC cohort the ICER increased with £500. Similarly, for the VHL CNS Hb cohort the ICER increased with £500. In the VHL pNET cohort the ICER increased with £800.

Table 6.17: Scenario where cohort specific median time to treatment response is incorporated

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Severity Adjusted ICER – EAG*
VHL RCC cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	73,602	53,542
VHL CNS Hb cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	57,440	39,523
VHL pNET cohort								
SOC	[redacted]	[redacted]	[redacted]					
Belzutifan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	78,481	63,229

Based on company electronic model,[114] and severity weighting calculations in Versteegh et al. 2019.[112]
* EAG's severity adjusted ICERs based on distribution of QALY weights per cohort.
CNS Hb = central nervous system hemangioblastoma; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; pNET = pancreatic neuroendocrine tumours; QALY = quality-adjusted life year; RCC = renal cell carcinoma; SOC = standard of care; VHL = Von Hippel-Lindau

6.3 EAG's preferred assumptions

As mentioned in Section 6.1.1 of this report, the EAG was unable to define a new base-case. Based on the evidence presented by the company, the EAG considered that the majority of the uncertainties identified in the CS cannot be resolved. For that reason, the EAG believes that any alternative base-case scenario, based on "preferred" assumptions that could have been presented, would still be surrounded by too many uncertainties and its results would be unreliable. Therefore, the EAG is afraid that a new base-case could give the wrong impression that it would be appropriate for the current DP. Instead, additional scenario analyses were explored by the EAG in order to highlight key uncertainty areas and to assess the impact of some alternative assumptions on the company base-case.

6.4 Conclusions of the cost effectiveness section

The CS[5] and response to the clarification letter[4] provided sufficient details for the EAG to appraise the literature searches conducted to identify economic, HRQoL and cost data on Belzutifan for treating tumours associated with VHL disease. Searches were conducted in July 2020, with updates in July 2022. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases were searched. Overall, the EAG has no major concerns about the literature searches conducted.

The company's base-case partly complied with the NICE reference case. Deviations from the NICE reference case related to the source of data for measurement and valuations of changes in HRQoL. HRQoL data were collected from three different sources, namely the VHL RW QoL Disease Burden Study,[1] KEYNOTE-564 and Kiebert et al. 2001.[2] These three studies reported utility values of patients from a population different to the one defined in the decision problem, and only the VHL RW QoL Disease Burden Study was conducted for the VHL disease. Furthermore, it is unclear if the populations in these studies are representative for the UK population since the company was unable to provide a comparison to the UK patient characteristics. In addition, the same QALY weight was applied by the company for the three different patient cohorts included in the submission (VHL associated RCC, CNS Hb and pNET) where the evidence presented suggests otherwise.

The key issues highlighted by the EAG throughout this report (and summarised in Table 6.1) were the following:

1) There is a mismatch between the DP and evidence used to inform it in terms of population. This could be broadly categorised in two types of issues:
    - Type of primary tumour (cannot be explored by the EAG).
    - Severity of the patient population. This had the following associated issues:
        - No evidence seems to be available for the DP population (cannot be explored by the EAG).
        - Surgery rates observed in the MK-6482-004 trial (Belzutifan) might underestimate the surgery rates for the population in the DP (explored by the EAG, but not evidence based, just for illustrative purposes).
        - The percentages of immediate surgery in the SoC arm seem arbitrary and have a major impact on the model results (explored by the EAG, but not evidence based, just for illustrative purposes).
        - The potential "harm" for Belzutifan patients for not having immediate surgery is not captured in the model. The potential "benefit" for SoC patients for having immediate surgery is not captured in the model (cannot be explored by the EAG).
        - Doubling the perioperative mortality risk in the three cohorts seems arbitrary and has a major impact on the model results (explored by the EAG, but not evidence based, just for illustrative purposes).
        - Increasing risks of risks of short- and long-term complications following surgery seems arbitrary and has a major impact on the model results (explored by the EAG, but not evidence based, just for illustrative purposes).
2) The comparator data might not be representative for the UK. This was addressed by the company by making some adjustments. Issues concerning these adjustments are the following:
    - Adjustments on transition probabilities based on Optum database may bring additional uncertainty rather than reducing it (explored by the EAG, but not evidence based, just for illustrative purposes).
    - There are inherent uncertainties imposed by using data from a different clinical practice (the US versus the UK) that cannot be resolved (cannot be explored by the EAG).
3) Data to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable (explored by the EAG but non-quantifiable uncertainties remain).
4) There is uncertainty in the derivation of the transition probabilities in the SoC arm (partially explored by the EAG, but not evidence based, just for illustrative purposes).
5) There is uncertainty in the derivation of transition probabilities in metastatic disease in both arms (not explored by the EAG due to lack of time).

6) There is uncertainty in the implementation of time on treatment and treatment effect waning. Potential issues relate to assuming:
    - Alternative parametric distributions for time on treatment (explored by the EAG).
    - Alternative duration of residual effect (explored by the EAG, but not evidence based, just for illustrative purposes).
7) There is uncertainty in the derivation and implementation of HRQoL in the model. Potential issues relate to:
    - A mismatch between the DP and the evidence used to inform HRQoL in terms of population (explored by the EAG, but not evidence based, just for illustrative purposes).
    - A potential bias attributable to assuming an immediate HRQoL benefit for Belzutifan patients (cannot be explored by the EAG).

First, the main concern of the EAG in this submission regarding the CE evidence relates to the mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses. The population included in this submission DP was defined by the company as adult patients with VHL disease who require treatment for VHL-associated RCC, CNS Hb, or pNET for whom surgery is unsuitable or undesirable. The main sources of evidence used to inform the effectiveness parameters in the CE model were the MK-6482-004 trial for Belzutifan,[24] and the VHL Natural History Study and the Optum Clinformatics Data Mart claims study for SoC.[26, 33] None of these studies included patients "for whom surgery is unsuitable or undesirable". The company acknowledged that *some* patients in the MK-6482-004 trial may have had less severe disease relative to those in the DP/MA population. Despite this, the company concluded that the population assessed in the economic analyses is in line with the population in the DP for Belzutifan. The EAG does not agree with this conclusion and considers that this mismatch creates a baseline imbalance for treatment comparison within the evidence synthesis of the submission, for the reasons summarised below.

The company explained that the economic model was initially developed to assess VHL-associated RCC only (in line with the population recruited into the MK-6482-004 trial) and without restriction to those patients for whom localised procedures are unsuitable or undesirable. The company then adapted the model by incorporating VHL-associated CNS Hb and pNET patients (with no restriction yet around localised procedures), following Belzutifan FDA approval in August 2021. This adaptation already raised some concerns since in the model the health states correspond to surgery and metastases of the primary tumour type only, while the sources of evidence used to inform the effectiveness input parameters of the model do not distinguish whether the type of tumour is primary or not. Thus, while the rate of surgeries should be informed by surgeries on the primary tumour, the data used to inform surgery rates lack this information. This is a potential source of bias because it is unknown whether the trial results (e.g., surgery rates) are applicable to the model population. To mitigate this potential bias, the company conducted several adjustments to the surgery-rates of non-primary tumours. However, the EAG would like to emphasise that the incidence rates of *primary* tumour (those used to estimate transition probabilities) were directly sourced from the MK-6482-004 trial. In this trial, for the CNS Hb cohort, there were only two CNS Hb surgeries observed, both performed on the same patient, while for the pNET cohort no pNET-related surgeries were observed, indicating the lack of mature data for these two subgroups. Thus, the EAG considers that the impact on the model results of not being able to distinguish whether the type of tumour is primary or not in the available data is unknown, and cannot be resolved, unless a different model structure is built to approach the DP, or another source of data is used to inform the current model. Moreover, the company indicated that distinguishing by different tumour type combinations was not feasible and, therefore, conducting CEAs using subgroup-specific parameters was not possible. This is another limitation of the current CEAs.

A further and possibly more important issue regarding the mismatch between the populations is related to the narrow definition of the patient population in the DP by including "those patients who require therapy and for whom localised procedures are unsuitable or undesirable". Since the MK-6482-004 trial did not include the restriction of localised procedures being unsuitable or undesirable, the trial population is likely to be less severe than the population in the DP. In an attempt to accommodate this more severe population, the company made several adjustments to their model, that sometimes seemed arbitrary to the EAG. The main EAG concerns associated to these adjustments are summarised below:

- The company assumed that only one surgery was possible as a "last resort" intervention and that only patients under SoC would receive this intervention immediately at the model start. The EAG considers that this assumption is problematic because it is intended to represent a severely ill population for which no evidence has been presented.
- Data from the MK-6482-004 trial were still used as the main source of evidence for Belzutifan. Given the severity of the disease of those patients included in the DP, which implies in the economic model that they are in need of immediate surgery, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial, underestimate the surgery rates for the population in the decision problem that would use Belzutifan in clinical practice. For instance, in the MK-6482-004 trial, the ORR in the VHL RCC subgroup was 63.9%, with a median TTR of 11.1 months (for patients with a response). The EAG is uncertain whether these data are appropriately representing patients as severe as those included in the DP, who need immediate surgery. If patients need immediate surgery, the EAG wonders whether in daily practice these patients would be able to wait almost 1 year (median TTR in RCC subgroup) without surgery until a response to treatment is observed. This EAG concern is further strengthened when considering that 36.1% of the trial population did not respond at all to Belzutifan treatment in the MK-6482-004 trial. This concern could be clarified by clinical experts.
- Immediate surgery was only applied in the SoC arm. It was assumed that immediate surgery would result in loss of organ function in 90% of patients, with the remaining 10% receiving symptom management in the VHL RCC and VHL pNET cohorts. For the VHL CNS Hb cohort, 50% of patients would undergo immediate surgery with a risk of brain injury, while in the remaining 50% of patients, for whom tumour location would not allow for operation, it was assumed that patients would undergo symptom management but with the same risk of brain injury as with surgery. These percentages of patients undergoing immediate surgery seem to be arbitrary and have a major impact on the model results. The company acknowledged that these percentages were not informed by means of formal elicitation methods and suggested to estimate the impact of these parameters through scenario analyses. However, considering the lack of evidence for these assumptions the EAG considered the scenarios provided by the company too narrow to reflect the underlying uncertainty.
- The EAG considers that by allowing only patients in the SoC to undergo immediate surgery, the company might disproportionally be favouring the Belzutifan arm. Patients have a requirement for immediate surgery to treat their primary tumour of significant burden in the absence of Belzutifan as a treatment option. This does not preclude that patients eligible for Belzutifan are also in need of immediate surgery to treat their primary tumour, as treatment with Belzutifan does not imply an immediate treatment benefit. To further reinforce this point, the company in response to question B4c on the availability of the active surveillance for CNS Hb patients says that 'patients in the CNS Hb cohort can have active surveillance but not without experiencing significant sequelae associated with tumour burden which would otherwise be alleviated through localised procedures'.[4] It is unclear to the EAG, why "similar" patients in the Belzutifan arm waiting for a response to Belzutifan treatment do not experience 'significant sequelae' which could otherwise be avoided

through localised procedures. Moreover, this statement seems to imply that having surgery would result in an improvement with respect to not having surgery for CNS Hb patients. Therefore, patients in the Belzutifan arm, who do not get surgery, would not get this improvement until they start to respond to treatment. The company further mentioned that 'the population stipulated by the MHRA label "for whom localised procedures are unsuitable or undesirable" are patients experiencing either debilitating sequelae as a result of surgery or debilitating sequelae as a result of not undergoing needed surgery'.[4] This suggests that the sequelae related to 'not undergoing needed surgery' would also be experienced by Belzutifan patients, until and if some response is achieved. Based on these statements from the company, the EAG is unclear whether patients waiting until they achieve response to Belzutifan treatment, might be in a worse state than patients in SoC undergoing immediate surgery. For this reason, in question B4e the company were asked to clarify if patients in the Belzutifan arm would effectively suffer the harm entailed to not receiving immediate surgery. The company responded that 'it is logical that a patient would suffer some harm if needed surgery were not provided immediately. The harm in this case would be risk of metastatic disease due to tumour growth (for RCC and pNET) or symptomatic burden (in all cohorts but particularly in CNS Hb). Belzutifan works by shrinking tumours and therefore reducing the risk of these two types of "harm". This benefit is reflected in the economic model through the transitions within the health states as informed by the trial evidence for Belzutifan'.[4] The EAG does not agree with this response and does not think this potential harm is completely captured in the model as patients in the trial did not need immediate surgery, meaning that their tumours were probably smaller than in the company's defined SoC arm and, therefore, the risk of metastatic disease would be smaller as well.

- Allowing a higher proportion of patients in the SoC arm to undergo surgery when they are in need of immediate surgery would be expected to lead to some treatment benefit compared to not receiving surgery. However, the EAG noticed that in the economic model when reducing the proportion of patients in the SoC arm undergoing immediate surgery the total QALYs and LYs gained in the SoC arm are higher, with the results being strongly influenced by more QALYs and LYs gained in the pre-surgery health state. This translates to patients being in a better position by not having surgery despite the immediate need, which seems to be contradictory to current clinical practice.
- The company doubled the perioperative mortality risk in the three cohorts to reflect the severity of patients in need of "last resort" surgery. The EAG considers this an arbitrary adjustment. In response to clarification question B9, the company justified this increase arguing that patients with more severe manifestations of VHL disease for whom surgery is unsuitable or undesirable would also experience a significantly increased risk of perioperative mortality risk. The EAG is of the opinion that these assumptions should be better justified, since considering the mismatch between the population in the trial and the MHRA authorisation, one should also adjust all transition probabilities estimated from the MK-6482-004 trial. Furthermore, the company refers to clinical expert opinion for the decision to make these adjustments, but no formal reference is provided. In response to clarification question C1, the company stated that "we are unable to provide documentation of this as it contains confidential information and we have not sought permission from participating experts".[4] Therefore, the EAG cannot properly assess the validity of these parameters.
- For similar reasons, the majority of the risks of short- and long-term complications following surgery for the VHL RCC cohort were doubled when considering the MHRA label population compared to the risks estimated from the Optum study. In response to clarification question B9, the company indicated that this increase is justified given that patients with more severe manifestations

of VHL disease, for whom surgery is unsuitable or undesirable, would also experience a significantly increased risk of surgery-related complications. As above, the company refers to clinical expert opinion for the decision to make these adjustments, but no formal reference is provided. Therefore, the EAG cannot assess the validity of these parameters either.

Second, the comparator data might not be representative of the UK. This was addressed by the company by making some adjustments. The company adjusted the transition probabilities i) pre-surgery → surgery, ii) pre-surgery → metastatic disease, and iii) event-free after surgery → metastatic disease in both treatment arms using data from the Optum database.[33] Adjustments to surgery and metastases rates were only conducted for the VHL RCC cohort, whilst adjustments to metastases rates only were made for the VHL CNS Hb and VHL pNET cohorts. The underlying assumption was that patients in the VHL Natural History Study and in the MK-6482-004 trial were subject to "high-quality" SoC. While maintaining the treatment benefit of Belzutifan compared to SoC, the surgery rates were further lowered compared to those observed in the MK-6482-004 trial. As mentioned above, the EAG considers it likely that the surgery rates observed in the MK-6482-004 trial, underestimate the surgery rates for the population in the DP that would use Belzutifan in clinical practice. This additional adjustment would only reinforce that underestimation. All in all, the EAG considers it likely that more severe patients would be at greater risk of surgery and metastasis. It could also be that Belzutifan is more effective in more severe patients, but with the current evidence it is unknown if the treatment effect would be the same as the one modelled (based on the MK-6482-004 trial) in a more severe population, so the underlying adjustments are not actually resolving uncertainties, rather than imposing additional assumptions. Furthermore, the EAG also considers that there are inherent uncertainties in the estimation of these transition probabilities, imposed by using data from a different clinical practice (the US versus the UK), which cannot be resolved by the implementation of additional adjustments, as these assumptions on adjustment parameters would likely be uncertain as well.

Third, the EAG is concerned that the data used to inform effectiveness in the Belzutifan arm (MK-6482-004 trial) are either immature or unavailable. Transition probabilities from the pre-surgery health state in the Belzutifan arm were estimated using data from the MK-6482-004 trial. As mentioned above, the population in the trial does not match the population in the DP and, in addition, these were either based on a small number of observed events or derived from other assumptions. This indicates that data from the MK-6482-004 trial (also for the pre-treatment period), and consequently, the survival analyses conducted by the company are subject to great uncertainty. This uncertainty is also applicable to the SoC arm for the VHL pNET and VHL CNS Hb cohorts, for which the pre-treatment period data from the MK-6482-004 trial were used to estimate pre-surgery → surgery transitions. Since no metastatic events prior to or following surgery were observed in the MK-6482-004 trial, the transition probabilities from pre-surgery → metastatic disease in the three cohorts in the Belzutifan arm were estimated by using the HR of pre-surgery → surgery estimated when comparing Belzutifan versus SoC. This assumption implies that the treatment effect of Belzutifan on the risk of surgery is equal to the treatment effect on the risk of metastases. According to the company, this assumption is clinically plausible given that Belzutifan is expected to reduce both the risks of surgeries and metastatic disease by reducing the tumour size and/or inhibiting their growth, but it is unclear to the EAG why it would be expected that the two effects would be exactly of the same magnitude. Regarding the derivation of transition probabilities from the event-free after surgery health state in the Belzutifan arm the EAG is concerned that the transition probability from event-free after surgery → metastatic disease in the VHL RCC cohort in the Belzutifan arm was estimated using the HR of pre-surgery → surgery for Belzutifan versus SoC, whereas for the VHL CNS Hb and VHL pNET cohorts in both arms, the transition probabilities for event-free after surgery → metastatic disease were set equal to the transition probability for pre-

surgery → metastatic disease. It is unclear to the EAG what the evidence basis for these assumptions is and if those have been validated by clinical experts.

Fourth, there is uncertainty in the derivation of the transition probabilities in the SoC arm. To estimate the transition probability from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm, the company used a retrospective analysis of the MK-6482-004 trial based on data from the pre-treatment period of the trial. That is because patients with CNS Hb or pNET tumours in the VHL Natural History Study could not be identified on the patient-level index date.[4] The CS states that for this analysis patient-level data were selected "(looking backwards) to the most recent primary tumour surgery prior to Belzutifan initiation in patients with VHL CNS Hb and VHL pNET tumours in the MK-6482-004 trial".[5] As the primary tumour surgery in these subgroups could be a surgery either due to RCC tumour, pNET or CNS Hb, it could be argued why the company used the subgroups to build the economic model at all. Focusing on VHL pNET or VHL CNS Hb patients would only make sense if the primary tumour would be VHL pNET and VHL CNS Hb, respectively. Considering the company's definition of the primary tumour throughout the submission (the tumour that defines treatment decision), the previous does not seem to be the case. Therefore, it remains a question to the EAG, why did the company decide to model the impact of Belzutifan treatment by using three different cohorts as previously highlighted above. Also, in response to clarification question B7b, the company stated that using the pre-treatment data from the MK-6482-004 trial to inform transitions from pre-surgery → surgery in the VHL CNS Hb and VHL pNET cohorts in the SoC arm, entails potential biases due to using two different data sources to define transitions within the same cohort: the pre-treatment period of the MK-6482-004 trial and the VHL Natural History Study.[4] These potential biases can be attributed to different treatment options, disease management and pathways. Furthermore, the company acknowledged that there may also be implications when comparing these two cohorts with the VHL RCC cohort of the SoC arm for which transition probabilities were completely informed from the VHL Natural History Study. In that perspective, to illustrate the impact of using the pre-treatment period trial data, the EAG requested the company to provide estimates of the respective transition probabilities for the VHL RCC cohort using the pre-treatment data from MK-6482-004 trial instead of the VHL Natural History Study (clarification question B7c).[4] The company noted that under this scenario, the incidence rate and distribution of non-RCC surgeries for SoC in the VHL RCC cohort should be also informed from the MK-6482-004 pre-treatment data. According to the company's statement, "*the inputs were not dramatically different under these data sources; when using pre-treatment period data, the rate of pre-surgery → first RCC surgery decreased, while the incidence of non-RCC surgeries increased. Additionally, the ICER improves when using this data source*".[4] The EAG does not agree with the company's perspective on the "inputs not being dramatically different", as the rate of RCC surgeries decreased to more than half of the rate estimated from the VHL Natural History Study, which can be seen as a substantial decrease. The impact of this change is also reflected on the ICER which dropped by more than 40% compared to the base-case ICER. The EAG confirmed that this change in ICER is driven by the change in the incidence rate of pre-surgery → first RCC surgery, and not by the change in the incidence of non-RCC surgeries. This analysis is further reflecting the EAG concerns on the appropriateness of using the MK-6482-004 pre-treatment period data to estimate the pre-surgery to surgery transitions for the VHL CNS Hb and VHL pNET cohorts and the potential impact these parameters may have in the currently presented company's base-case analysis. Based on the difference in the incidence rate for the VHL RCC cohort as estimated by the VHL Natural History Study and the pre-treatment period data from the MK-6482-004 trial, it could be expected that the pre-surgery → surgery incidence rate for the VHL CNS Hb and VHL pNET cohorts in the SoC arm might also be underestimated based on the VHL Natural History Study.

Fifth, there is also uncertainty in the derivation of the transition probabilities for the metastatic disease health state in both treatment arms. The company relied on many assumptions to model the transition probabilities from metastatic disease to death. However, the EAG was unable to verify all of them due to tie constraints. It is unclear though why many treatment options for metastatic disease are included but then some of the market shares are set at 0%. For the VHL RCC cohort the market shares of avelumab/axitinib and pembrolizumab/lenvatinib are set to in Table 62 of the CS.[5] For the VHL pNET cohort, the market shares for streptozocin / 5-fluorouracil, streptozocin / doxorubicin, temozolomide / capecitabine, everolimus, sunitinib, interferon a2B and no active treatment are set to 0%, while these have been included in the NMA analysis. The EAG wonders whether this was performed to allow connections for treatment effectiveness in the NMA network and if a simpler approach would have been more transparent in this case. Also, AEs associated to the treatments in the metastatic disease health state should have also been included in the model, although usually the impact of AEs in model results is not major.

Sixth, there is uncertainty in the implementation of ToT and treatment effect waning in the model. There is uncertainty in the long-term extrapolations of the parametric distribution selected to model ToT. This uncertainty has also been acknowledged by the company in response to clarification B13.[4] Despite this, the company presented only one additional scenario analysis using the second best more fit based on AIC and BIC values (the Weibull distribution). The duration of the residual benefit assumed for Belzutifan is also uncertain. The 2.71-year period assumed in the base-case could be seen as a reasonable choice, but it could also be argued that it could be different, especially for CNS Hb and pNET related tumours, for which no data are available. Therefore, also in this case, scenario analyses are appropriate to assess the impact of this assumption on the model results. In response to clarification question B14, the company stated that applying treatment waning before the maximum observed trial period (3.84 years) would lead to a mismatch between "observed versus predicted curves for time to surgery, metastatic disease, or death in the Belzutifan arm" as these data account for patient discontinuation (response in question B14).[4] The company noted that "it would therefore be inappropriate to consider a treatment effect waning assumption of no residual benefit (i.e. before the maximum follow-up period of the trial is complete)".[4] However, survival data from the MK-6482-004 trial were only used to estimate transitions in the pre-surgery → surgery transitions for VHL RCC and VHL CNS Hb patients, which were also estimated using a small number of observed events. Other transitions were derived from other assumptions such as HR-based approaches. This indicates that data from the MK-6482-004 trial (also for the pre-treatment period), and consequently, the survival analyses conducted by the company are subject to great uncertainty as mentioned above. This in turn, points towards a great uncertainty also for the duration of the Belzutifan treatment effect.

Finally, there are two main concerns of the EAG regarding the implementation of HRQoL in the model. First, the company have used response-adjusted health state utilities computed as a weighted average of utility values by response level observed in each treatment arm. This approach seems reasonable but as it is currently implemented in the model, patients in the Belzutifan arm obtain an immediate benefit in HRQoL compared to SoC from the first model cycle. This seems unrealistic. The EAG understands that at the treatment decision point patients are in a poor health condition, including those who receive Belzutifan. It seems therefore unlikely that before Belzutifan starts showing some effect patients could experience any type of benefit. This also implies that another potential source of bias comes from having distinct baseline response distributions between treatment arms. At model start, patients are not equal in the Belzutifan and the SoC arms, and in fact SoC patients in the model are more severe than patients in Belzutifan. The EAG considers that this issue could have been resolved by including the timing of response in the model. However, with the available data, the EAG is unable to make such a change to the economic model in a proper way. The company have also acknowledged that the licensed population

(the population in the DP) may have worse utility scores than those used in the model because the survey in the VHL RW QoL Disease Burden Study was not limited to only those patients who require therapy and for whom localised procedures are unsuitable or undesirable. Thus, the VHL RW QoL Disease Burden Study is not fully generalisable to the MHRA license population. As a result, the utility estimates from the patients in the VHL RW QoL Disease Burden Study are likely an overestimate of the utilities that would be obtained from the licensed population. The EAG wonders whether the company could have "adjusted" these utilities to better reflect the relevant patient population as they did with other parameters such as risks of surgery-related complications. However, even if the licensed population is expected to have worse utility scores, it does not mean that the effect of Belzutifan on HRQoL is underestimated in the model as the company claims. It could be the other way around, even if it is in an indirect way, e.g., some model assumptions could lead to an overestimation of Belzutifan HRQoL, for example when in the model it is assumed an immediate HRQoL effect associated to Belzutifan as explained above.

The company's base-case results indicated that Belzutifan was more costly and more effective than SoC in all cohorts. Compared to SoC, in the RCC cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs and the ICER was £73,095 per QALY gained. In the CNS Hb cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs compared to SoC, and the ICER was £56,933 per QALY gained. Finally, in the pNET cohort Belzutifan accrued [redacted] incremental QALYs at [redacted] additional costs and the ICER in the pNET cohort was £77,649 per QALY gained. When accounting for disease severity, the company assumed a QALY weight of 1.7 for all VHL cohorts, which results on ICERs equal to £42,997, £33,490, and £45,676 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. The EAG does not agree with the company's rationale for selecting a QALY weight of 1.7 for all cohorts and applied severity adjusted QALYs based on the weight likelihood derived from Versteegh et al. 2019.[112] Based on this weight distribution, the severity adjusted ICERs were equal to £49,359, £33,976 and £64,311 per QALY gained for the RCC, CNS Hb and pNET cohorts, respectively. The average PSA were in line with the deterministic ones. For the three cohorts [redacted]. At the common thresholds of £20,000 and £30,000 per QALY gained, the estimated probability that Belzutifan is a cost-effective alternative to SoC was [redacted] for the three cohorts. When accounting for disease severity, the company's implementation of the severity weights was incorrect according to the EAG. First, in the CS, it was assumed the same severity weight of 1.7 for all cohort and all PSA iterations. This does not match with the model implementation where the severity weight varies per PSA iteration. Therefore, the severity adjusted ICERs in the CS are incorrect. Furthermore, in the model, the severity weights were calculated for the RCC cohort but applied to the CNS Hb and pNET cohorts too. This is also incorrect. The largest impact is observed in the pNET cohort, because the company assumed a weight of 1.7, but for this cohort the majority of the PSA outcomes resulted in a weight of 1.2. As a consequence, the severity adjusted ICER calculated by the EAG is approximately £20,000 larger than the severity adjusted ICER calculated by the company. When severity weighting was not considered, at the common thresholds of £20,000 and £30,000 per QALY gained, the estimated probability that Belzutifan is a cost-effective alternative to SoC was [redacted] for the three cohorts. When accounting for disease severity, the severity adjusted probability of being cost effective at the threshold of £30,000 per QALY gained, was equal to [redacted]% for the RCC and pNET cohorts, and [redacted]% for the CNS Hb cohort. The company presented in total the results of 14 scenario analyses to assess the robustness of the model results to changes in some modelling assumptions. These included exploring alternative long-term assumptions regarding Belzutifan treatment effect, changing utilities, considering alternative pre-surgery rates for the VHL RCC cohort

in the Belzutifan arm, not adjusting surgery and metastases rates to account for real-world SoC, not applying relative dose intensity in the calculations, including indirect costs, considering shorter model time horizons or lower discount rates. The modelling assumptions explored by the company that had the greatest effect on the ICER were related to the utility in the non-metastatic health states, the proportion of patients to receive immediate surgery in the SoC arm and the removal of treatment effect waning. The EAG was concerned that sometimes there was no clear rationale for conducting scenarios and that many assumptions in the model were not tested (or not shown) by the company. Therefore, it is unclear the impact of other modelling assumptions.

The scenario analyses conducted by the EAG were mostly explorative given the lack of other sources of evidence and many of the alternative assumptions explored were arbitrarily selected. Results nevertheless indicated that the ICER in the current model was sensitive to several assumptions. As expected, the proportion of patients receiving immediate surgery in the SoC arm had a major impact on the results. Also, by comparing SoC with and without immediate surgery the EAG's idea that the severity of the DP population has not been appropriately captured by the company's model was reinforced. Results were also sensitive to changes in perioperative mortality risks and in risks of risks of short- and long-term complications following surgery. This seemed to be arbitrarily defined by the company and had a major impact on the model results. Results indicated that the distribution chosen to model time to surgery for the Belzutifan arm also had major impact on the ICER, illustrating in this way the uncertainty associated to the immaturity of the data. Special care needs to be taken when assessing the results for the pNET cohort since these might be lacking face validity. Alternative assumptions on time on treatment, duration of the Belzutifan residual treatment effect or the choice of utilities, should also be considered as sources of relevant uncertainty. It is notable that in all scenarios explored by the EAG, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained. Only in a couple of them, and for the VHL CNS Hb cohort only, the ICER was close to £30,000.

The EAG acknowledges the difficulty of representing the population in the DP with the current evidence, which is mostly derived from the MK-6482-004 trial for Belzutifan. Considering the above, the current model structure and the available data, the EAG is unable to change the model in a straightforward manner to account for patients requiring immediate surgery in the Belzutifan arm as assumed in SoC. In its current form, the company's model might be considered appropriate to reflect the company's initially sought MA VHL-associated RCC only and for the population recruited into MK-6482-004 trial, but cannot provide reliable estimates of the CE of Belzutifan compared to SoC in the population defined in the DP given the model's inability to properly represent/capture "those patients who require therapy and for whom localised procedures are unsuitable or undesirable".

7. REFERENCES

[1] Merck Sharp & Dohme. *VHL-related disease impact on quality of life and work productivity in patients with renal cell carcinoma, central nervous system hemangioblastoma, or pancreatic neuroendocrine tumor manifestation [MSD Data On File]. PDF supplied by Company*: Merck Sharp & Dohme, 2022

[2] Kiebert GM, Green C, Murphy C, Mitchell JD, O'Brien M, Burrell A, et al. Patients' health-related quality of life and utilities associated with different stages of amyotrophic lateral sclerosis. J Neurol Sci 2001; 191(1-2):87-93

[3] Rini BI, Plimack ER, Stus V, Waddell T, Gafanov R, Pouliot F, et al. Pembrolizumab (pembro) plus axitinib (axi) versus sunitinib as first-line therapy for advanced clear cell renal cell carcinoma (ccRCC): Results from 42-month follow-up of KEYNOTE-426. J Clin Oncol 2021; 39(15_suppl):4500-4500

[4] National Institute for Health and Care Excellence. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): response to request for clarification from the ERG*. London: NICE, 2023. 116p.

[5] Merck Sharp & Dohme (UK) Ltd. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): submission to National Institute of Health and Care Excellence. single technology appraisal (STA): document B - company evidence submission (v3.0 - 2 June 2023)*: Merck Sharp & Dohme (UK) Ltd., 2023. 431p.

[6] Medicines & Healthcare Products Regulatory Agency. *Welireg 40 mg film-coated tablets. Summary of Product Characteristics*: MHRA, 2022 [Internet] Available from: https://products.mhra.gov.uk/search/?search=welireg&page=1&doc=Spc&rerouteType=0

[7] National Institute for Health and Care Excellence. *Belzutifan for treating tumours associated with von Hippel-Lindau disease [ID3932]: final scope [Internet]*. London: NICE, 2023 [accessed 14.03.23] Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10817/documents

[8] Maher ER, Adlard J, Barwell J, Brady AF, Brennan P, Cook J, et al. Evaluation of tumour surveillance protocols and outcomes in von Hippel-Lindau disease in a national health service. Br J Cancer 2022; Epub ahead of print

[9] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. J Clin Epidemiol 2016 [accessed 26.4.23]; 75:40-6

[10] Canadian Agency for Drugs and Technologies in Health. *PRESS - peer review of electronic search strategies: 2015 guideline explanation and elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 26.4.23] Available from: https://www.cadth.ca/resources/finding-evidence/press

[11] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 26.4.23]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[12] National Cancer Institute (NCI). About cancer: diagnosis and staging [Internet]. National Institutes of Health (NIH), n.d. [accessed 11.7.23]. Available from: https://www.cancer.gov/about-cancer/diagnosis-staging

[13] Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012; 28(2):138-44

[14] Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997; 350(9074):326-9

[15] Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Marshall C, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*: Cochrane, 2022. Available from: http://www.training.cochrane.org/handbook

[16] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 4.3.22] Available from: https://training.cochrane.org/handbook

[17] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019; 366:l4898

[18] Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016; 355:i4919

[19] Joanna Briggs Institute. The Joanna Briggs Institute critical appraisal tools for use in JBI systematic reviews: checklist for quasi-experimental studies [Internet]. Adelaide: Joanna Briggs Institute, 2017 [accessed 13.7.23]. Available from: https://jbi.global/critical-appraisal-tools

[20] Jonasch E, Donskov F, Iliopoulos O, Rathmell WK, Narayan VK, Maughan BL, et al. Belzutifan for renal cell carcinoma in von Hippel-Lindau disease. N Engl J Med 2021; 385(22):2036-2046

[21] Merck Sharp & Dohme. *HTA advisory board summary slides [MSD Data on File]*: MSD, 2022

[22] Sano M, Aizawa Y, Katsumata Y, Nishiyama N, Takatsuki S, Kamitsuji S, et al. Evaluation of differences in automated QT/QTc measurements between Fukuda Denshi and Nihon Koden systems. PLoS One 2014; 9(9):e106947

[23] Cancer Research UK. TNM staging for breast cancer [Internet]. Cancer Research UK, 2023 [accessed 17.7.23]. Available from: https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-grades/tnm-staging

[24] Merck Sharp & Dohme LLC. *MK-6482-004 clinical study report [Data on file] [as cited in the CSJ*: Merck Sharp & Dohme, 2022

[25] Merck Sharp & Dohme LLC. Belzutifan/MK-6482 for the treatment of advanced pheochromocytoma/paraganglioma (PPGL), pancreatic neuroendocrine tumor (pNET), or von Hippel-Lindau (VHL) disease-associated tumors (MK-6482-015). NCT04924075. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2022 [accessed 14.3.23]. Available from: https://clinicaltrials.gov/ct2/show/NCT04924075

[26] Merck Sharp & Dohme (UK) Ltd. *Von Hippel-Lindau Natural History Study - Final Study Report v1.0. [MSD Data On File]*: Merck Sharp & Dohme, 2021

[27] Merck Sharp & Dohme (UK) Ltd. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): submission to National Institute of Health and Care Excellence. single technology appraisal (STA): Appendix R. Analyses of pre-treatment period data from the MK-6482-*

004 trial to estimate primary and non-primary VHL-related surgery rates under active surveillance in the CNS Hb and pNET cohorts: Merck Sharp & Dohme (UK) Ltd., 2023. 4p.

[28] National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual. Process and methods [PMG36] [Internet]*. London: National Institute for Health and Care Excellence, 2022 [accessed 8.2.22]. 181p. Available from: https://www.nice.org.uk/process/pmg36

[29] Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. *NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE*. Sheffield: Decision Support Unit, ScHARR, 2015. 82p. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf

[30] Faria R, Hernandez Alava M, Manca A, Wailoo AJ. *NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness for technology appraisal: methods for comparative individual patient data*. Sheffield: Decision Support Unit, ScHARR, 2015. 86p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD17-DSU-Observational-data-FINAL.pdf

[31] National Institute for Health and Care Excellence. *NICE health technology evaluation topic selection: the manual [PMG37]*. London: National Institute for Health and Care Excellence, 2022 Available from: https://www.nice.org.uk/process/pmg37/chapter/about-this-guide

[32] Hernandez-Alava M, Wailoo A, Pudney S. Methods for mapping between the EQ 5D 5L and the 3L. 2017;

[33] Jonasch E, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Epidemiology and economic burden of von Hippel-Lindau disease-associated renal cell carcinoma in the United States. Clin Genitourin Cancer 2023; 21(2):238-247

[34] Lonser RR, Butman JA, Huntoon K, Asthagiri AR, Wu T, Bakhtian KD, et al. Prospective natural history study of central nervous system hemangioblastomas in von Hippel-Lindau disease. J Neurosurg 2014; 120(5):1055-62

[35] National Institute for Health and Care Excellence. *Pembrolizumab for adjuvant treatment of renal cell carcinoma [ID3810] - [TA830] - Committee Papers*: National Institute for Health and Care Excellence, 2022 Available from: https://www.nice.org.uk/guidance/ta830/documents/committee-papers

[36] Pavel M, Öberg K, Falconi M, Krenning EP, Sundin A, Perren A, et al. Gastroenteropancreatic neuroendocrine neoplasms: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2020; 31(7):844-860

[37] National Institute for Health and Care Excellence. *Pembrolizumab for adjuvant treatment of renal cell carcinoma [TA830] - Final Appraisal document*. London: National Institute for Health and Care Excellence, 2022 Available from: https://www.nice.org.uk/guidance/ta830/documents/final-appraisal-determination-document

[38] Bensimon AG, Zhong Y, Swami U, Briggs A, Young J, Feng Y, et al. Cost-effectiveness of pembrolizumab with axitinib as first-line treatment for advanced renal cell carcinoma. Curr Med Res Opin 2020; 36(9):1507-1517

[39] Motzer R, Alekseev B, Rha S-Y, Porta C, Eto M, Powles T, et al. Lenvatinib plus pembrolizumab or everolimus for advanced renal cell carcinoma. N Engl J Med 2021; 384(14):1289-1300

[40] Sun W, Lipsitz S, Catalano P, Mailliard JA, Haller DG, Eastern Cooperative Oncology Group. Phase II/III study of doxorubicin with fluorouracil compared with streptozocin with fluorouracil or dacarbazine in the treatment of advanced carcinoid tumors: Eastern Cooperative Oncology Group Study E1281. J Clin Oncol 2005; 23(22):4897-904

[41] Faivre S, Niccoli P, Castellano D, Valle JW, Hammel P, Raoul JL, et al. Sunitinib in pancreatic neuroendocrine tumors: updated progression-free survival and final overall survival from a phase III randomized study. Ann Oncol 2017; 28(2):339-343

[42] Kaderli RM, Spanjol M, Kollar A, Butikofer L, Gloy V, Dumont RA, et al. Therapeutic options for neuroendocrine tumors: a systematic review and network meta-analysis. JAMA Oncol 2019; 5(4):480-489

[43] Riaz IB, He H, Ryu AJ, Siddiqi R, Naqvi SAA, Yao Y, et al. A Living, Interactive Systematic Review and Network Meta-analysis of First-line Treatment of Metastatic Renal Cell Carcinoma. Eur Urol 2021; 80(6):712-723

[44] Johnson A, Sudarshan S, Liu J, Linehan WM, Pinto PA, Bratslavsky GC. Feasibility and outcomes of repeat partial nephrectomy. J Urol 2008; 180(1):89-93

[45] Lonser RR, Weil RJ, Wanebo JE, DeVroom HL, Oldfield EH. Surgical management of spinal cord hemangioblastomas in patients with von Hippel-Lindau disease. J Neurosurg 2003; 98(1):106-16

[46] Krauss T, Ferrara AM, Links TP, Wellner U, Bancos I, Kvachenyuk A, et al. Preventive medicine of von Hippel-Lindau disease-associated pancreatic neuroendocrine tumors. Endocr Relat Cancer 2018; 25(9):783-793

[47] Wang L, Bensimon AG, Sundaram M, Xu R, Lai Y, Liu Y, et al. Burden of surgeries and surgical complications in patients with von Hippel Lindau (VHL) disease before and after treatment with belzutifan. J Clin Oncol 2023; 41(6_suppl):733-733

[48] Schuhmacher P, Kim E, Hahn F, Sekula P, Jilg CA, Leiber C, et al. Growth characteristics and therapeutic decision markers in von Hippel-Lindau disease patients with renal cell carcinoma. Orphanet J Rare Dis 2019; 14(1):235

[49] Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health 2010; 13(5):509-18

[50] Efthymiadou O, Mossman J, Kanavos P. Health related quality of life aspects not captured by EQ-5D-5L: results from an international survey of patients. Health Policy 2019; 123(2):159-165

[51] Siller S, Szelenyi A, Herlitz L, Tonn JC, Zausinger S. Spinal cord hemangioblastomas: significance of intraoperative neurophysiological monitoring for resection and long-term outcome. J Neurosurg Spine 2017; 26(4):483-493

[52] Rochette C, Baumstarck K, Canoni-Zattara H, Abdullah AE, Figarella-Branger D, Pertuit M, et al. Psychological impact of von Hippel-Lindau genetic screening in patients with a previous history of hemangioblastoma of the central nervous system. J Psychosoc Oncol 2018; 36(5):624-634

[53] Rasmussen A, Alonso E, Ochoa A, De Biase I, Familiar I, Yescas P, et al. Uptake of genetic testing and long-term tumor surveillance in von Hippel-Lindau disease. BMC Med Genet 2010; 11:4

[54] Fraser L, Watts S, Cargill A, Sutton S, Hodgson S. Study comparing two types of screening provision for people with von Hippel-Lindau disease. Fam Cancer 2007; 6(1):103-11

[55] Gurusamy KS, Riviere D, van Laarhoven CJH, Besselink M, Abu-Hilal M, Davidson BR, et al. Cost-effectiveness of laparoscopic versus open distal pancreatectomy for pancreatic cancer. PLoS One 2017; 12(12):e0189631

[56] Lee AJ, Morgan CL, Conway P, Currie CJ. Characterisation and comparison of health-related quality of life for patients with renal failure. Curr Med Res Opin 2005; 21(11):1777-83

[57] Simianu VV, Gaertner WB, Kuntz K, Kwaan MR, Lowry AC, Madoff RD, et al. Cost-effectiveness evaluation of laparoscopic versus robotic minimally invasive colectomy. Ann Surg 2020; 272(2):334-341

[58] Gandhi SK, Jensen MM, Fox KM, Smolen L, Olsson AG, Paulsson TC. Cost-effectiveness of rosuvastatin in comparison with generic atorvastatin and simvastatin in a Swedish population at high risk of cardiovascular events. Clinicoecon Outcomes Res 2012; 4:1-11

[59] National Institute for Health and Care Excellence. *Neuroendocrine tumours (metastatic, unresectable, progressive) - everolimus and sunitinib [ID858]*. London: National Institute for Health and Care Excellence, 2017 Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10024

[60] Nisula S, Vaara ST, Kaukonen KM, Reinikainen M, Koivisto SP, Inkinen O, et al. Six-month survival and quality of life of intensive care patients with acute kidney injury. Crit Care 2013; 17(5):R250

[61] Zargar M, McFarlane T, Chan KKW, Wong WWLC. Cost-effectiveness of nivolumab in recurrent metastatic head and neck squamous cell carcinoma. Oncologist 2018; 23(2):225-233

[62] Poole CD, Chambers C, Allsopp R, Currie CJ. Quality of life and health-related utility analysis of adults with moderate and severe atopic dermatitis treated with tacrolimus ointment vs. topical corticosteroids. J Eur Acad Dermatol Venereol 2010; 24(6):674-8

[63] Stevenson SM, Danzig MR, Ghandour RA, Deibert CM, Decastro GJ, Benson MC, et al. Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer. Urol Oncol 2014; 32(8):1172-7

[64] Sankar A, Dixon PR, Sivanathan L, Memtsoudis SG, de Almeida JR, Singh M. Cost-effectiveness analysis of preoperative screening strategies for obstructive sleep apnea among patients undergoing elective inpatient surgery. Anesthesiology 2020; 133(4):787-800

[65] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins JC. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes 2008; 6:84

[66] Wang CY, Pham PN, Kim S, Lingineni K, Schmidt S, Diaby V, et al. Predicting cost-effectiveness of generic vs. brand dabigatran using pharmacometric estimates among patients with atrial fibrillation in the United States. Clin Transl Sci 2020; 13(2):352-361

[67] Memeh K, Ruhle B, Vaghaiwalla T, Kaplan E, Keutgen X, Angelos P. Thyroidectomy for euthyroid patients with Hashimoto thyroiditis and persisting symptoms: a cost-effectiveness analysis. Surgery 2021; 169(1):7-13

[68] Verkleij ML, Heijnsdijk EAM, Bussé AML, Carr G, Goedegebure A, Mackey AR, et al. Cost-effectiveness of neonatal hearing screening programs: a micro-simulation modeling analysis. Ear Hear 2021; 42(4):909-916

[69] National Institute for Health and Care Excellence. *Ocrelizumab for treating relapsing–remitting multiple sclerosis - Technology appraisal guidance [TA533]*. London: National Institute for Health and Care Excellence, 2018 Available from: https://www.nice.org.uk/guidance/ta533

[70] Turner D, Adams E, Boulton M, Harrison S, Khan N, Rose P, et al. Partners and close family members of long-term cancer survivors: health status, psychosocial well-being and unmet supportive care needs. Psychooncology 2013; 22(1):12-9

[71] Sjolander C, Rolander B, Järhult J, Mårtensson J, Ahlstrom G. Health-related quality of life in family members of patients with an advanced cancer diagnosis: a one-year prospective study. Health Qual Life Outcomes 2012; 10:89

[72] Song JI, Shin DW, Choi JY, Kang J, Baek YJ, Mo HN, et al. Quality of life and mental health in the bereaved family members of patients with terminal cancer. Psychooncology 2012; 21(11):1158-66

[73] NHS. *National Cost Collection: 2020/21 Report [Excel spreadsheet supplied by Company]*. London: NHS England and NHS Improvement, 2022 [accessed 14.3.23]

[74] BNF. *British National Formulary*: National Institute for Health and Care Excellence, 2022 Available from: https://bnf.nice.org.uk/

[75] Jones KC, Burns A. *Unit costs of health and social care 2021*. University of Kent: Personal Social Services Research Unit, 2021 Available from: https://kar.kent.ac.uk/id/eprint/92342

[76] Jonasch E, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Healthcare resource utilization (HRU) and costs among patients with von Hippel-Lindau disease (VHL)-associated renal cell carcinoma (RCC): a retrospective administrative claims analysis. J Clin Oncol 2022; 40(6_suppl):305-305

[77] Sundaram M, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Real-world treatment patterns in von Hippel-Lindau (VHL) disease-associated renal cell carcinoma (RCC): costs of tumor reduction procedures and their complications. J Clin Oncol 2022; 40(16_suppl):4539-4539

[78] National Institute for Health and Care Excellence. *Pembrolizumab for adjuvant treatment of completely resected stage 3 melanoma [TA766]*. London: National Institute for Health and Care Excellence, 2022 Available from: https://www.nice.org.uk/guidance/ta766

[79] NHS. Drugs and pharmaceutical electronic market information tool (eMIT). 2022. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[80] National Institute for Health and Care Excellence. *Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma [TA169]*. London: National Institute for Health and Care Excellence, 2009 Available from: https://www.nice.org.uk/guidance/ta169

[81] National Institute for Health and Care Excellence. *Pazopanib for the first-line treatment of advanced renal cell carcinoma - Technology appraisal guidance [TA215]*. London: National Institute for Health and Care Excellence, 2013 Available from: https://www.nice.org.uk/guidance/ta215

[82] Merck Sharp & Dohme. *MSD KEYNOTE-426 Clinical Study Report. [MSD Data On File]*: Merck Sharp & Dohme, 2018

[83] National Institute for Health and Care Excellence. *Tivozanib for treating advanced renal cell carcinoma - Technology appraisal guidance [TA512]*. London: National Institute for Health and Care Excellence, 2018 Available from: https://www.nice.org.uk/guidance/ta512

[84] National Institute for Health and Care Excellence. *Cabozantinib for untreated advanced renal cell carcinoma - Technology appraisal guidance [TA542]*. London: National Institute for Health and Care Excellence, 2018 Available from: https://www.nice.org.uk/guidance/TA542

[85] Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, et al. Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. N Engl J Med 2019; 380(12):1103-1115

[86] National Institute for Health and Care Excellence. *Nivolumab for previously treated advanced renal cell carcinoma - Technology appraisal guidance [TA417]*. London: National Institute for Health and Care Excellence, 2016 Available from: https://www.nice.org.uk/guidance/ta417

[87] National Institute for Health and Care Excellence. *Axitinib for treating advanced renal cell carcinoma after failure of prior systemic treatment [TA333]*. London: National Institute for Health and Care Excellence, 2015 Available from: https://www.nice.org.uk/guidance/ta333

[88] National Institute for Health and Care Excellence. *Cabozantinib for previously treated advanced renal cell carcinoma - Technology appraisal guidance [TA463]*. London: National Institute for Health and Care Excellence, 2017 Available from: https://www.nice.org.uk/guidance/ta463

[89] Motzer RJ, Hutson TE, Glen H, Michaelson MD, Molina A, Eisen T, et al. Lenvatinib, everolimus, and the combination in patients with metastatic renal cell carcinoma: a randomised, phase 2, open-label, multicentre trial. Lancet Oncol 2015; 16(15):1473-1482

[90] National Institute for Health and Care Excellence. *Everolimus for the second-line treatment of advanced renal cell carcinoma - Technology appraisal guidance [TA219]*. London: National Institute for Health and Care Excellence, 2011 Available from: https://www.nice.org.uk/guidance/ta219

[91] National Institute for Health and Care Excellence. *Everolimus for advanced renal cell carcinoma after previous treatment - Technology appraisal guidance [TA432]*. London: National Institute for Health and Care Excellence, 2017 Available from: https://www.nice.org.uk/guidance/ta432

[92] Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med 2007; 356(22):2271-81

[93] Strosberg JR, Fine RL, Choi J, Nasir A, Coppola D, Chen DT, et al. First-line chemotherapy with capecitabine and temozolomide in patients with metastatic pancreatic endocrine carcinomas. Cancer 2011; 117(2):268-75

[94] Iwasa S, Morizane C, Okusaka T, Ueno H, Ikeda M, Kondo S, et al. Cisplatin and etoposide as first-line chemotherapy for poorly differentiated neuroendocrine carcinoma of the hepatobiliary tract and pancreas. Jpn J Clin Oncol 2010; 40(4):313-8

[95] Hentic O, Hammel P, Couvelard A, Rebours V, Zappa M, Palazzo M, et al. FOLFIRI regimen: an effective second-line chemotherapy after failure of etoposide-platinum combination in patients with neuroendocrine carcinomas grade 3. Endocr Relat Cancer 2012; 19(6):751-7

[96] Faure M, Niccoli P, Autret A, Cavaglione G, Mineur L, Raoul JLC. Systemic chemotherapy with FOLFOX in metastatic grade 1/2 neuroendocrine cancer. Mol Clin Oncol 2017; 6(1):44-48

[97] Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 2015; 373(19):1803-13

[98] Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier S, et al. Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. Lancet Oncol 2013; 14(6):552-62

[99] Motzer RJ, Escudier B, Powles T, Scheffold C, Choueiri TK. Long-term follow-up of overall survival for cabozantinib versus everolimus in advanced renal cell carcinoma. Br J Cancer 2018; 118(9):1176-1178

[100] Sternberg CN, Hawkins RE, Wagstaff J, Salman P, Mardiak J, Barrios CH, et al. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. Eur J Cancer 2013; 49(6):1287-96

[101] Hutson TE, Escudier B, Esteban E, Bjarnason GA, Lim HY, Pittman KB, et al. Randomized phase III trial of temsirolimus versus sorafenib as second-line therapy after sunitinib in patients with metastatic renal cell carcinoma. J Clin Oncol 2014; 32(8):760-7

[102] Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, Ou S-S, et al. Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. J Clin Oncol 2008; 26(33):5422-5428

[103] Maher ER, Neumann HP, Richard SC. von Hippel-Lindau disease: a clinical and scientific review. Eur J Hum Genet 2011; 19(6):617-23

[104] Kanno H, Kobayashi N, Nakanowatari SC. Pathological and clinical features and management of central nervous system hemangioblastomas in von Hippel-Lindau disease. J Kidney Cancer VHL 2014; 1(4):46-55

[105] National Institute for Health and Care Excellence. *Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer - Technology appraisal guidance [TA476]*. London: National Institute for Health and Care Excellence, 2017 Available from: https://www.nice.org.uk/guidance/ta476

[106] Georghiou T, Bardsley M. *Exploring the cost of care at the end of life*: Nuffield Trust, 2014 Available from: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life

[107] Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews BC. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrol Dial Transplant 2012; 27 Suppl 3(Suppl 3):iii73-80

[108] Coronini-Cronberg S, Appleby J, Thompson JC. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. J R Soc Med 2013; 106(7):278-87

[109] Patel A, Berdunov V, Quayyum Z, King D, Knapp M, Wittenberg RC. Estimated societal costs of stroke in the UK based on a discrete event simulation. Age Ageing 2020; 49(2):270-276

[110] Fineberg NA, Haddad PM, Carpenter L, Gannon B, Sharpe R, Young AH, et al. The size, burden and cost of disorders of the brain in the UK. J Psychopharmacol 2013; 27(9):761-70

[111] Moore A, Young CA, Hughes DA. Health utilities and costs for motor neurone disease. Value Health 2019; 22(11):1257-1265

[112] Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-adjusted probability of being cost effective. Pharmacoeconomics 2019; 37(9):1155-1163

[113] iDBC - iMTA Disease Burden Calculator. [accessed 17.7.23]. Available from: https://imtamodels.shinyapps.io/iDBCv2_1/

[114] Merck Sharp & Dohme (UK) Ltd. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): submission to National Institute of Health and Care Excellence. single technology appraisal (STA): CEA model (v3.0 - 6 June 2023)*: Merck Sharp & Dohme (UK) Ltd., 2023

[115] Merck Sharp & Dohme (UK) Ltd. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): submission to National Institute of Health and Care Excellence. single technology appraisal (STA): document B - company evidence submission (v2.0 - 19 May 2023)*: Merck Sharp & Dohme (UK) Ltd., 2023. 431p.

[116] Merck Sharp & Dohme (UK) Ltd. *Belzutifan for treating tumours associated with von Hippel-Lindau disease (ID3932): submission to National Institute of Health and Care Excellence. single technology appraisal (STA): document B - company evidence submission (v1.0 - 13 March 2023)*: Merck Sharp & Dohme (UK) Ltd., 2023. 438p.

[117] Vemer P, Corro Ramos I, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics 2016; 34(4):349-61

[118] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.* Sheffield: Decision Support Unit, ScHARR, 2011. 52p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[119] Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. Pharmacoeconomics 2019; 37(11):1391-1408

[120] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. Pharmacoeconomics 2020; 38(2):205-216

[121] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. Health Technol Assess 2016; 20(26):1-48

ERG report – factual accuracy check and confidential information check

[ID3932] Belzutifan for treating tumours associated with von Hippel-Lindau disease

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
In section 1.3 Table 1.3 Key Issue 2, the following statement is made regarding the expected effect on the cost-effectiveness estimates "None if the DP population is adapted to better align with the MK-6482 study, notwithstanding any change due to addressing Key Issue 1."	This text should be removed.	Such comment from the EAG is out of scope as the decision problem population is agreed and finalised prior to the company submission.	Not a factual inaccuracy. Also, please note that the DP population is changeable by the committee.
In section 1.4 Table 1.4, with regard to the clinical effectiveness SLR presented by the company and the 25 records initially included in the SLR but subsequently excluded for not reporting relevant information on belzutifan, it is stated that "It is not	This text should be removed, or it should be clarified that the details of these 25 records were provided in the company submission.	The full bibliographic details of these 25 records are provided in Appendix D Table 110 of the company submission that would allow for this to be clarified. While these 25 records were noted in the EAG's clarification letter, additional details on the information reported in them were not requested by the EAG in that letter. Therefore, this statement erroneously mis-characterises this issue as "not clear".	Not a factual inaccuracy.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
clear whether the remaining 25 records could have contributed comparator data"			
In section 1.4 Table 1.4, with regard to the clinical effectiveness SLR presented by the company it is stated that "it was not clear whether the search strategy and study selection criteria were designed to identify all relevant interventional, non-interventional and natural history studies"	This text should be removed.	The SLR search strategy and selection criteria presented in Appendix D of the company submission, which are described by the EAG in section 3.1.1 of the report to be "well structured, transparent and reproducible", clearly state that the following study types would be included in the SLR: • Randomised controlled trials (RCTs) • Controlled clinical trials • Non-randomized clinical trials, including single-arm prospective interventional trials • Prospective and retrospective cohort studies These clearly encompass and would necessitate the inclusion of any relevant interventional and non-interventional trial studies and so there is no ambiguity with this. Accordingly, the VHL Natural History Study is a retrospective non-interventional study and would have been identified and included in the SLR based on these searches and criteria if it had been a published study at the time of the SLR search. General "natural history studies" of relevance to this appraisal would also be prospective or retrospective	Not a factual inaccuracy.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		cohort studies and so would also be included in the SLR. Furthermore, the full bibliographic details of each of the 5 studies in total excluded at the full-text screening phase on the basis of study design are provided in Appendix D Table 111 for transparency and therefore how the SLR included/excluded studies based on study design (including in terms of "relevant interventional, non-interventional and natural history studies") can be evaluated in a clear manner. Therefore, this statement erroneously mis-characterises this issue as "not clear".	
In section 1.4 Table 1.4, with regard to the clinical effectiveness SLR presented by the company it is stated that as an alternative approach suggested by the EAG it is stated "Provide greater transparency in relation to the existing SLR methods so that the approaches used to identify comparator data can be fully evaluated."	This text should be removed.	The SLR search strategy and selection criteria presented in Appendix D of the company submission are described by the EAG in section 3.1.1 of the report to be "well structured, transparent and reproducible" and are fully transparent and reproducible as described in the comments above, and so this statement is contradictory and incorrect.	Not a factual inaccuracy. The description taken from the EAG report "*well structured, transparent and reproducible*" refers only to the search strategy and not the study selection criteria.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
In section 1.4, Table 1.5, with regard to the row describing the issue, there are several factual inaccuracies regarding which data source has been used to inform which transitions. There appears to be a misunderstanding on which transitions were informed by the pre-treatment phase of the MK-6482-004 trial and why the Optum study could not be used for an ITC.	This issue should be removed as the description of the issue presents comments on the data source that are factually inaccurate.	In the description of the issue the following statement is made: "the pre-treatment phase of MK-6482-004 to inform rates of pre-surgery->metastatic disease and pre-surgery->death" which is incorrect. The pre-treatment phase was used to inform rates of **pre-surgery → surgery in the VHL CNS Hb and pNET cohorts**. Additionally, the text states "the company indicated the superiority of the pre-treatment phase of MK-6482-004 as a source of comparator data and it is unclear why this data source was not used to estimate all outcomes including TTS". This statement is incorrect and misrepresents the use of the data source. The superiority of the pre-treatment phase was only mentioned in relation to the lack of ability to use the VHL Natural History Study for the CNS Hb and pNET cohorts. Furthermore, the last part of the statement implies that this data source was not used to estimate TTS, which is factually inaccurate, since it was used to estimate TTS for the CNS Hb and pNET cohorts. Additionally, in response to B7c) of the CL, a scenario for the RCC cohort using this data source is also presented. The response to B7c) of the CL also details clearly why the pre-treatment phase could not be used to generate TTM or TTD, namely that patients were alive and metastases-free for the entirety of the pre-treatment period follow-up, as specified by the eligibility criteria of the MK-6482-004 trial.	Table 1.5 has been amended accordingly, as well as Section 3.3.2 and 3.6.3.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		This row of the CL also states, “It is also unclear why the Optum Clinformatics Data Mart Claims Study was not used in the ITC given its potentially greater applicability to UK clinical practice.” This statement does not take into account the explanation given in response to A16e) & A32 in the CL where the lack of matching variables in the Optum study was cited as the reason an ITC could not be conducted using this data source.	
In section 1.5, Table 1.7, regarding point a) made in the description of the issue “The model distinguishes three cohorts per tumour type, which is not possible with the current evidence.”	This text should be rewritten to say “The model distinguishes three cohorts **by** tumour type, which is not possible with the current evidence.”	The current text implies that there are 9 cohorts assessed, 3 for each tumour type.	This has been amended as suggested by the company.
In section 1.5, Table 1.9 regarding the statement “Data in the MK-6482-004 trial (Belzutifan) is extremely immature for the three cohorts” in the description of the issue. Table 1.11 also refers to “data immaturity” and these two should be classified as one issue.	The text should be removed and clarified in terms of number of events occurred.	We would suggest that the phrases here represent “cancer-like” thinking. VHL is a chronic disease that is characterized by both benign and malignant tumour manifestations. In chronic disease terms, the data are not immature with a mean follow-up of 3.18 years at the 01-APR-2022 data cut-off. Any statements referring to data immaturity would be better termed around number of surgery events that have occurred within the follow-up period if the issue is to be raised. Furthermore, key issue 8 and key issue	Not a factual inaccuracy. Key Issue 8 (Table 1.9) refers to immaturity in the sense that there is uncertainty in the long-term extrapolations of treatment effectiveness in general. Key issue 10 (Table 1.11) is specific to the implementation of time on treatment and the

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		10 requesting for longer-term data should be characterized as one issue.	residual benefit of belzutifan. While there is overlap with Key Issue 8, the EAG preferred to present them separately to emphasize that one referred to treatment effectiveness and the other one to time of treatment and residual benefit.
In section 1.5, Table 1.10 regarding a mistake made in the statement "Pre-treatment data from the MK-6482-004 trial were used to inform transitions from pre-surgery → surgery in the VHL RCC and VHL pNET cohorts in the SoC arm".	This text should be corrected to say "Pre-treatment data from the MK-6482-004 trial were used to inform transitions from pre-surgery → surgery in the VHL **CNS Hb** and VHL pNET cohorts in the SoC arm".	The pre-treatment period of the MK-6482-004 trial was used to inform pre-surgery → surgery for the CNS Hb and pNET cohorts. In the RCC cohort, this transition was informed by the VHL Natural History Study.	This has been amended as suggested by the company.
In section 1.5, Table 1.13 Key issue 12 is raised "Model is built to estimate Belzutifan cost effectiveness compared to SoC in three different subgroups of patients. However, subgroup-specific parameters were not used in the model."	This issue should be removed as it implies that the subgroups are mutually exclusive.	The suggestion that a cost-effectiveness analyses should be conducted with subgroup specific parameters implies that the subgroups are mutually exclusive which is not the case. By the inclusion criteria for the trial, all patients had RCC and a subset of these had CNS Hb and/or pNET. Given the nature of the dataset, these groups are not mutually exclusive and have significant overlap.	Not a factual inaccuracy. The EAG understands how the "subgroups" were defined and issues about the subgroup definition were also raised by the EAG. Given that the evidence presented suggests that the

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			clinical effectiveness or the disease severity may be different per subgroup, the EAG considers it to be more appropriate to use subgroup-specific parameters in this submission.
In section 1.7, the following statement is made "It is remarkable that in all scenarios explored by the EAG, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained."	The text should be reworded to state "It is **notable** that in all scenarios explored by the EAG, with and without severity weight, all ICERs were above the commonly used threshold ICER of £30,000 per QALY gained."	There is substantial uncaptured value in the cost-effectiveness analyses which is detailed in section B.3.13 of the company submission. We find the term "remarkable" inappropriate and the use of "notable" more appropriate.	This has been amended as suggested by the company.
In section 1.7, the following statement is made "…the company's model might be considered appropriate to reflect the initial marketing authorisation VHL-associated renal cell carcinoma (RCC) only…" which misrepresents that	This text should be reworded to state "…the company's model might be considered appropriate to reflect **the company's initially sought** marketing authorisation in VHL-	The current wording implies that a marketing authorization was initially granted in VHL-RCC alone which is not the case, there was no initial marketing authorization, only the final marketing authorization. The company initially sought marketing authorization for VHL-associated RCC, but it was granted (from the outset) for all three tumour types.	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
a marketing authorization was granted for RCC alone which is not the case.	associated renal cell carcinoma (RCC) only…".		
In section 2, Table 2.1, in the EAG comment on Population, it is stated that "Within their response, the company acknowledged the implication that study patients may have had less severe disease relative to those in the DP/MA population"	This text should be corrected to state "Within their response, the company acknowledged the implication that **some** study patients may have had less severe disease relative to those in the DP/MA population"	Correction to accurately reflect what was stated in the company's responses to the EAG clarification questions.	Comment amended.
In section 2, Table 2.1, in the EAG comment on Outcomes it is stated "The choice of outcomes appears to be driven by what was available in the MK-6482-004 study and does not fully address the NICE Final Scope".	This text should be removed.	While the company stated that OS and HRQoL were not collected as part of the MK-6482-004 study. These outcomes were considered in the company submission as explicitly stated in the same table: "OS and HRQoL are considered in the cost-effectiveness analyses, derived from sources other than the MK-6482-004 study." Therefore, the outcomes considered in the company submission align with the NICE final scope and the statement that this is not the case is erroneous.	Not a factual inaccuracy.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
In section 2.3 on page 39, the statement “This suggests that the role of active surveillance in the management of patients with VHL-associated RCC, CNS Hb or pNETs is unclear.” which is inaccurate.	This text should be removed.	The role of active surveillance in the management of VHL patients is not unclear; the uncertainty is surrounding the treatment options for SoC patients “who require therapy” but face a “Hobson’s choice” of treatment options.	Not a factual inaccuracy.
In section 2.4 the EAG comment includes the text “The EAG note the discrepancy between the outcomes listed in the NICE Final Scope and the DP table”.	This text should be removed.	As described in the comment above with regard to section 2, Table 2.1, in the EAG comment on Outcomes, there is no actual discrepancy.	Not a factual inaccuracy.
In section 3.1.1 as part of the EAG’s comments it is stated that “The EAG noted that no search terms relating to natural history studies were included in the study design filters in the Embase search strategy for clinical effectiveness” and that “It is unclear whether any other potentially relevant	This text should be removed.	As described in the comments above, the “well structured, transparent and reproducible” search strategies presented in Appendix D of the company submission would have identified any published studies analogous to the unpublished VHL Natural History Study, which is a retrospective non-interventional study for which search terms are included in the Embase search strategy presented in Appendix D Table 106 (e.g. the “Retrospective study/” string, among others), and so this statement erroneously mis-characterises this issue as “unclear”.	Comment amended

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
studies may have been missed by not including additional search terms for other natural history studies in the field."			
In section 3.1.1 as part of the EAG's comments it is stated that "The EAG noted that no search terms relating to natural history studies were included in the study design filters in the Embase search strategy for clinical effectiveness (Table 105; Appendix D), however the VHL natural history study was identified (clarification question A.2)."	This text should be removed.	This text mis-represents this situation as the EAG's clarification question A.2 related to the VHL Natural History Study *specifically*: "A 2. No search terms relating to the VHL Natural History Study are included in the study design filters for clinical effectiveness. Please explain why terms for the VHL Natural History Study are not included in the study design filters for clinical effectiveness in Appendix D, given that the main comparator study is a natural history study. Given the above, please explain how the VHL Natural History Study was identified." And not with regard to "natural history studies" *in general*.	Comment amended
In section 3.1.2 as part of the EAG's comments it is stated that "The EAG noted that the SLR excluded case series	This text should be removed.	The VHL Natural History Study is a retrospective non-interventional study (as detailed in the final study report for this study provided to the EAG as reference #46) and not a case series.	Not a factual inaccuracy.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
which is the design of the VHL Natural History study"			
In section 3.1.2 as part of the EAG's comments regarding the two non-English studies that were excluded from the SLR it is stated that "the company did not provide specific details of the excluded studies and therefore it was not possible for the EAG to determine the impact of these omissions on clinical effectiveness estimates"	This text should be removed, or corrected to "specific details on these two excluded studies were not sought from the company and the EAG did not determine the impact of these omissions on clinical effectiveness estimates"	The current wording of this statement mis-represents the situation, MSD were and are willing to provide the details of these two studies if these had been requested, these were only not provided because no request was made for them by the EAG.	Not a factual inaccuracy.
In section 3.1.3 as part of the EAG's comments with regard to the approach for data extraction used for the SLR of clinical effectiveness, it is stated that "the CS (Appendix Q)[5] included less detail on the main source of comparator data (the	This text should be removed.	This statement appears to be included here erroneously as the VHL Natural History Study was not identified or included in the SLR (indeed this would not be possible due to it being an unpublished study), and so any data extraction of this study would not be applicable/relevant in the context of the SLR. Furthermore, in terms of information provided on the VHL Natural History Study as part of the company submission in general, the full final study report for this study was provided to the EAG as reference #46 of the	Not a factual inaccuracy. Data extraction of all studies contributing data to clinical effectiveness estimates is part of recommended good practice in systematic literature reviews.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
VHL Natural History study), particularly in relation to outcomes."		company submission which provides full detailed information on the study.	
In section 3.4.2 as part of the EAG's comments regarding the MAIC, the MAIC is termed a "so-called" MAIC.	The "so-called" text should be removed.	This text erroneously mis-characterises the MAIC presented in the company submission as somehow not an MAIC. While the MAIC presented in the company submission was not performed via the academically ideal approach where IPD from both the VHL Natural History Study and the MK-6482-004 study are used, a MAIC that uses IPD from one study to match baseline summary characteristics reported from another study is still unambiguously an MAIC: • "Matching-adjusted indirect comparisons (MAICs) use IPD from trials of one treatment to match baseline summary statistics reported from trials of another treatment. After matching, by using an approach similar to propensity score weighting, treatment outcomes are compared across balanced trial populations." (1).	Not a factual inaccuracy: this term was used to reflect what the EAG knew pre-clarification letter. Elsewhere, this term is not used.
In section 4.2.2 on page 126 there is a typographical error in the statement "the transitions form *[sic]* the pre-surgery to the surgery health state".	The text "form" should be corrected to "from".	Typographical error	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
In section 4.2.3 on page 126 the statement "The population included in the economic analyses was defined by the company as adult patients (18 years or older) with VHL disease who require treatment for VHL-associated RCC, VHL-associated CNS Hb, or VHL associated pNET for whom surgery is unsuitable or undesirable" is worded slightly differently than the marketing authorisation.	The text should be amended to read "The population included in the economic analyses was defined by the company as adult patients (18 years or older) with VHL disease who require **therapy** for VHL-associated RCC, VHL-associated CNS Hb, or VHL associated pNET **and** for whom **localised procedures are** unsuitable or undesirable"	Lack of clarity in the text and worded slightly differently to DP/MA population.	This has been amended as suggested by the company.
In section 4.2.3 Table 4.4, the key baseline characteristics used in the economic model reports "Female/males (%)" for which the term used is inaccurate.	This should be corrected to "**Females (%)**" to reflect the value reported.	Lack of clarity in the text.	This has been amended as suggested by the company.
In section 4.2.6 on page 136, there is a lack of clarity in the statement in the statement "it was only	This text should be corrected to read "it was only used to inform the surgery	Lack of clarity in the text.	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
used to inform the surgery rates when the treatment effect derived from the Belzutifan treatment in the Belzutifan plus SoC comparator waned".	rates when the treatment effect derived from the Belzutifan treatment in the Belzutifan **arm** waned".		
In section 4.2.6.1 in Table 4.5 under the VHL pNET cohort in the Belzutifan row the incorrect data source "SoC rate, adjusted for reduced death cases attributable to VHL CNS Hb" is reported for the event-free after surgery → death transition.	This text should be removed and replaced with "**Assumed equal to pre-surgery → death**".	Incorrect data source reported.	This has been amended as suggested by the company.
In section 4.2.6.2 on page 137 point a) of the EAG comment reads "a) To estimate the transition probability from pre-surgery → surgery in the VHL RCC and VHL pNET cohorts in the SoC arm, the company used a retrospective analysis of the MK-6482-004 trial	The text should be corrected to read: "a) To estimate the transition probability from pre-surgery → surgery in the VHL **CNS Hb** and VHL pNET cohorts in the SoC arm, the company used a	The pre-treatment period of the MK-6482-004 trial was used to inform pre-surgery → surgery for the CNS Hb and pNET cohorts. In the RCC cohort, this transition was informed by the VHL Natural History Study.	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
based on data from the pre-treatment period of the trial". On the same page in point b) "…the company stated that using the pre-treatment data from the MK-6482-004 trial to inform transitions from pre-surgery → surgery in the VHL RCC and VHL pNET cohorts in the SoC arm…". Both of these statements incorrectly identify the RCC cohort instead of the CNS Hb cohort as being informed by the pre-treatment data from the MK-6482-004 trial.	retrospective analysis of the MK-6482-004 trial based on data from the pre-treatment period of the trial". "…the company stated that using the pre-treatment data from the MK-6482-004 trial to inform transitions from pre-surgery → surgery in the VHL **CNS Hb** and VHL pNET cohorts in the SoC arm…"		
In section 4.2.6.2 on page 137 point a) of the EAG comment states "Focusing on VHL pNET or VHL CNS Hb patients would only make sense if the primary tumour would be VHL pNET and VHL	This statement should be removed.	In response to B2c) of the CL, the 'primary tumour' is defined as the one driving treatment decisions. As we cannot determine in patients with more than one tumour manifestation which is the 'primary tumour' within the available dataset, patients are modelled through each cohort for their respective tumour manifestation. For example, a patient with an RCC + pNET tumour, they would be modelled in the RCC	Not a factual inaccuracy. As mentioned above, the EAG understands how the "subgroups" were defined. This text points out that the EAG considers that the subgroup definition should be

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
CNS Hb, respectively. Considering the company's definition of the primary tumour throughout the submission (the tumour that defines treatment decision), the previous does not seem to be the case." This appears to show misunderstanding in how the cohorts were used and defined in the model.		cohort (as though RCC was the primary tumour) and also in the pNET cohort (as though pNET was the primary tumour). The comment made by the EAG suggests that a patient in the pNET cohort would not have pNET as the primary tumour, this is inaccurate. As we do not know which is the primary tumour, the patients are modelled as though any one of the three tumour manifestations they have present (RCC/CNS Hb/pNET) could be the primary tumour.	based on the primary tumour, even though this information is not available to the company. This highlights another limitation/area of uncertainty within this submission.
In section 4.2.6.2 on page 139 the following statement is made: "It is unclear to the EAG if these estimates were based on the MAIC analysis or on a naïve comparison between Belzutifan and SoC treatments."	This statement should be removed.	We would like to apologise if this was not made clear in the CS. We can confirm that all parameters informed by the VHL Natural History were based on the MAIC analysis and no naïve comparison was performed. For pre-surgery → metastatic disease and pre-surgery → death the same MAIC methods were used for the CNS Hb and pNET cohorts as was done for the RCC cohort. Please see Table 44 of the CS for a summary of clinical parameters sourced from the VHL Natural History Study (using a MAIC analysis) and the pre-treatment period data from the MK-6482-004 trial.	Not a factual inaccuracy, since at the time of writing the EAG report, this was indeed unclear. However, the EAG report has been updated based on the additional clarification provided by the company here. This sentence was included in a paragraph illustrating an EAG concern, which now has been resolved. Therefore, the whole paragraph has been

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			removed from the EAG report (pages 138-139 and 252).
In section 4.2.6.3 on page 143 the following inaccurate statement is made in parentheses "reference for the NMA not provided within the CS or reference pack".	This statement should be removed.	References were provided within the CS and reference pack. The NMA references are reference no. 57 (Riaz 2021) and no. 61 (Kaderli 2019).	This has been amended as suggested by the company.
In section 4.2.8.1 on page 153 a typographical error is made in the statement "Twenty-one out of the 2,022 patients were from the UK".	This should be corrected to read "Twenty-one out of the **220** patients were from the UK".	Typographical error	This has been amended as suggested by the company.
In section 4.8.2.5 in the final row of Table 4.22, the complication details of endolymphatic sac tumour surgery is mistakenly copied for the complication details of retinal Hb surgery.	The following complication should be reported for complications of retinal Hb surgery as per Table 78 of the CS: "Vitreous heamorrhage; -0.223; Assumed equal to vision loss disutility derived from Ament et al. (2018)	Typographical error	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
	(neurological complication: visual loss at 2 months)"		
In section 4.8.2.5 on page 162 at the end of EAG comment point a) the statement "The EAG considers that this issue could have been resolved by including the timing of response in the model. However, with the available data, the EAG is unable to make such a change to the economic model.", which is inaccurate.	The statement should be amended to read: "The EAG considers that this issue could have been resolved by including the timing of response in the model **and this was explored by the EAG in scenario analyses set 6. The company explored extreme stress testing of this scenario in response to B21c) of the CL**."	In the EAG scenario analyses set 6, the EAG were able to incorporate median TTR into the QALY calculation to explore the impact of a delayed response. Therefore, it is factually inaccurate to state that the EAG were unable to make such changes. Furthermore, in response to B21c) of the CL, we explore the impact of the limitation of fixed proportions at each response level. The EAG's statement does not explain that exploratory analyses on this assumption were conducted by both the company and the EAG and therefore misrepresents how this assumption has been explored	The text has been amended as follows: "The EAG considers that this issue might have been resolved by including time to treatment response in the model and by linking the objective response level to time to response to calculate utility values in the pre-surgery, surgery, and event-free after surgery states. The EAG explored a scenario in which a fixed cut-off at the median time to treatment response was included to the QALY calculation (please refer to Scenario analyses set 6 in Section 6.1.2.6). The company also explored the impact of using fixed proportions at each response level in response to clarification question B21c.[4] However, these scenarios

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			were only exploratory and their results should be interpreted with caution. With the available data, the EAG was unable to remove Belzutifan's immediate effect on the response level which was used to estimate the weighted utility values as used in the economic model."
In section 4.2.10 page 186 to 189 details the QALY shortfall analysis and an alternative analysis method is put forward by the EAG that is not in line with the NICE reference case.	The QALY shortfall analysis conducted by the EAG should be in line with the NICE reference case and use UK QALY shortfall calculator tool as has been done in other submissions.	We note that there are issues around the EAG's method of calculating the severity modifier. Firstly, the EAG use an online tool that is based in the Netherlands and although this tool allows UK specific utility data, it uses a UK value set by derived by Heijink et al. (2011) to estimate general population QALYs. This is inconsistent with the NICE reference case which recommends the Hernández Alava et al. 2017 mapping function, using the 'EEPRU dataset' (Hernández Alava et al. 2020) which was done in the company submission. Secondly, a UK based free online tool has been developed by Schneider et al. (2022) (a collaboration between University of York, University of Sheffield and Lumanity) using the NICE reference case value set and can be accessed via this link https://shiny.york.ac.uk/shortfall/. This tool is in line with	Not a factual inaccuracy. The EAG considers that the discussion on the severity of the condition should not be focused on whether one specific tool or another should have been used. We consider this irrelevant. The main EAG issue with the CS relates to the fact that the same severity weights were used for all three subgroups. The EAG believes this is incorrect. Using the severity-adjusted probability of being cost effective is *one possible way* to address this issue.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		the NICE reference case and produces results consistent with those presented in the company submission. This tool has also been used in other EAG assessments conducted by KSR for other NICE appraisals (NICE TA911, ID4036). Lastly, the approach of using a "likelihood of applicable QALY weight" based on PSA results is mentioned nowhere in the NICE reference case on QALY shortfall analysis and appears to simply be an artefact of the Netherlands-based tool used by the EAG. The section on Decision modifiers in the NICE Manual does not mention use of PSA results to produce such weighting of QALY weights. More specifically, the EAG calculated a QALY weight of 1.7 with 55.9% likelihood and 1.2 with 44.1% likelihood effectively equating to a QALY weight of 1.48. This would not be permitted under the NICE Manual as it is clear that based on an absolute shortfall and/or proportional shortfall calculation it is **either** the 1.0, 1.2 or 1.7 QALY weight that can apply. Furthermore, the manual states that modifiers that cannot be included in the estimated QALYs can be taken into account qualitatively through committee discussion or quantitatively through QALY weighting.	We would like to add some clarification points below: Firstly, the Disease Burden Calculator (iDBC) tool used by the EAG is also a free online tool used to estimate disease burden, and to the best of our knowledge it was the first online tool developed for that purpose. The iDBC tool was developed by the Institute for Medical Technology Assessment (iMTA) of the Erasmus University Rotterdam (EUR). This iMTA/EUR team has been partnering with KSR and acting as EAG for many years. The EAG considers it inappropriate to refer to the iDBC tool as "based in the Netherlands", as opposed to "UK based" since this might suggest that the results of the iDBC tool are less valid. The iDBC tool can be applied to a wide range of countries, including the UK. Therefore, the results from the iDBC are

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			specific for the UK setting and considered appropriate for NICE submissions. Secondly, the EAG is aware of the QALY Shortfall Calculator tool (Scheider et al, 2022) mentioned by the company and that one of the differences with respect to the iDBC tool, *but not the only one*, is the use of Hernández Alava instead of Heijink for the UK value set. However, there is no sense in debating the platform with which the absolute and proportional shortfall is to be calculated, as the uncertainty can equally be calculated with the QALY Shortfall Calculator. This would only require taking the 'remaining QALYs without disease' calculated with the tool as input in the calculation of absolute and proportional shortfall for each run of the PSA. While the EAG agrees with the company that using Hernández Alava would be in

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			line with the NICE reference-case, we would like to invite the company to conduct the severity analyses with both tools and check whether results are substantially different. As an example, we compared the company's PSA results obtained with the iDBC and the QALY Shortfall Calculator tools. Note that the only difference would be in the number of QALYs without the disease. These would be 18.02 with the iDBC tool and 18.15 with the QALY Shortfall Calculator tool using Hernandez-Alava (as reported in the company's model). The results for the RCC cohort are as follows: RCC cohort with iDBC tool: 42.7% for weight 1.2 and 57.3% for weight 1.7. Weighted ICER £51,116. RCC cohort with QALY Shortfall Calculator tool: 42.2% for weight 1.2 and

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			57.8% for weight 1.7. Weighted ICER £51,027. Also, one of the main reasons why the QALY Shortfall Calculator tool has also been previously used in other EAG assessments conducted by KSR was that the iDBC tool was being updated to include more recent life tables and discounting, the latter to conform with the latest NICE methods. Finally, the EAG would like to clarify and emphasise that the likelihood of applicable QALY weight is not simply an "artefact of the Netherlands-based tool". The methodology included in the iDBC tool follows the 2019 publication in Pharmacoeconomics of the Severity-Adjusted Probability of Being Cost Effective (DOI: 10.1007/s40273-019-00810-8). Therefore, it is not an artefact, but a methodology published after a peer-

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			reviewed process. This method works under the assumption that the proportion and absolute number of QALYs lost is uncertain, and that this uncertainty is reflected in the PSA. Regardless of which tool is used, the EAG considers that a more fair assessment of proportional and absolute shortfall is to account for this uncertainty, as there may be submissions that happen to have a deterministic QALY loss results in a QALY multiplier group that is not fitting to the entire sample. While it is understandable that this might cause some resistance with submissions where the deterministic QALY happens to correspond with the upper QALY multiplier, there will equally be cases conceivable where the opposite occurs. Both can be dealt with equally by accounting for the uncertainty in the QALY loss predicted by

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			the model. Furthermore, the fact that the updated NICE methods manual does not mention a particular methodology, in this case the severity-adjusted probability of being cost effective, should not prevent the EAG for using it. We would also like to clarify that, in the example provided by the company, interpreting the severity-adjusted probability of being cost effective as effectively equating to a QALY weight of 1.48 is indeed incorrect, since as the company correctly indicate, there are only three possible weights. The severity-adjusted probability of being cost effective should be interpreted in relation to the cost-effectiveness thresholds.
In section 4.2.10 on page 187 there is a typographical error “This shows for example that, for the RCC cohort, even	This should be corrected to read “This shows for example that, for the RCC cohort, even	Typographical error	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
though a weight *[sic]* point estimate is 1.7, there is a 44.1% that the applicable QALY weight is 1.2, which may have a substantial impact on the severity adjusted results."	though a **weighted** point estimate is 1.7, there is a 44.1% that the applicable QALY weight is 1.2, which may have a substantial impact on the severity adjusted results."		
In section 4.2.10 on page 189 the statement "note that the so-called VHL-GB MA population has not been included in Table 4.41 since it is unclear how the QALY shortfall was calculated for this population." is inaccurate.	This statement should be removed and the QALY shortfall analysis for the VHL-GB MA population should be included.	In page 2 of the document titled "NICE ID3932 Belzutifan further information request [CIC] v1.0" it is clarified that this severity weight is calculated as a weighted average. The "Base-Case QALY Shortfall" sheet in the economic model details the number of patients with each tumour manifestation in the MK-6482-004 trial as the basis for the weighting to produce the QALY weight for the MA population. The CE model breaks down the GB MA population into non-mutually exclusive cohorts. A x1.7 QALY weight for the MA population was estimated from a weighted average of the SOC QALYs for each cohort. In the absence of UK specific data on the relative proportions of each cohort, the weights were derived from the number of patients with each tumour manifestation in the MK-6482-004 trial. Specifically, the "Remaining QALYs with disease (under SOC)" were weighted by these proportions and then the absolute and proportional shortfall subsequently calculated,	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		corresponding to a x1.7 QALY weight. The results of this QALY shortfall analysis are presented in Table 96 of the CS.	
In section 4.2.10 at the top of page 190 the statement "The EAG understands that in this case, the "specific population" are the three or four cohorts shown in Table 4.40, since the absolute and proportional QALY shortfall are estimated using four different subsets of data." misrepresents the results.	This statement should be reworded to reflect that there are **three** cohorts and then a weighted MA population is drawn from these cohorts.	The current statement implies that there are more than three cohorts/tumour manifestations that are being considered. There are only three with the final row of the table representing a weighted average of the three cohorts, not a separate cohort.	This has been amended as suggested by the company.
In section 4.2.10 at the top of page 190 the following statement is made "Also, the VHL-GB MA population seems to be irrelevant since it was not included in the company's model."	This statement should be removed as it misrepresents the cost-effectiveness analysis.	It appears that presenting a single ICER for the VHL-GB MA population, although not previously explicitly requested, would alleviate this concern raised by the EAG that the GB MA population was not included in the model. To mitigate this concern raised by the EAG, a combined weighted ICER is presented in Table 1 below, both with and without the severity weight included in the calculation. The interpretation of the combined weighted ICER is limited by the fact that all patients in the MK-6482-004 trial had RCC, and by the	The EAG would like to apologise for this misunderstanding. The EAG considers that this population irrelevant and that the severity weightings should be calculated per subgroup separately as explained in the EAG report.

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td>fact that no data is available on the proportions of VHL patients in the UK who are affected by these three individual tumour sites as primary tumours. However, it is consistent with our approach in calculating a single QALY weight for the full VHL-GB MA population. This has also been included in the cost-effectiveness model titled “NICE ID3932 STA Submission CEA v4.0 (CIC)” provided with this response.

Table 1 Summary base case ICER including combined weighted ICER for GB MA population
<table>
<tr><th>Cohort</th><th>Number of patients with manifestation in MK-6482-004</th><th>ICER (£/QALY)</th><th>Severity modifier adjusted ICER</th></tr>
<tr><td>VHL GB marketing authorisation population (weighted cohort, assuming no overlap)</td><td>133</td><td>£66,256</td><td>£39,133</td></tr>
<tr><td>VHL-RCC</td><td>61</td><td>£73,095</td><td>£42,997</td></tr>
<tr><td>VHL-CNS Hb</td><td>50</td><td>£56,933</td><td>£33,490</td></tr>
<tr><td>VHL-pNET</td><td>22</td><td>£77,649</td><td>£45,676</td></tr>
</table></td><td>The EAG report has been amended to illustrate this point.</td></tr>
</table>

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
In section 4.2.10 in the final paragraph on page 190 the statement "The error is that the QALY weights are calculated for the RCC cohort, but then these weights are applied to all three cohorts, which is incorrect" which is factually inaccurate.	This statement should be removed.	In page 2 of the document titled "NICE ID3932 Belzutifan further information request [CIC] v1.0" it is clarified that this QALY weight is calculated as a weighted average for the MA population in each PSA iteration and is not based on the RCC cohort alone.	This has been amended with the explanation provided by the company, but not deleted. The EAG still considers that this is not the correct approach and that the severity weights should have been calculated for each VHL cohort separately.
In section 5.2.3 on page 203 there is a typographical error in the statement "…whereas the EAG used a mixed *[sic]* of weights…".	This statement should read "…whereas the EAG used a **mix** of weights…".	Typographical error	This has been amended as suggested by the company.
In section 5.3.2.1 on page 209 there is a typographical error in the statement "…namely patients who "requires *[sic]* therapy"…".	This statement should read "…namely patients who "**require** therapy"…".	Typographical error	This has been amended as suggested by the company.
In section 5.3.2.1 on page 211 there is a typographical error in the statement "…a spontaneous reduction in size of VHL-related tumours is highly unlikely	The statement should read "…a spontaneous reduction in size of VHL-related tumours is highly unlikely ~~in~~	Typographical error	This has been amended as suggested by the company.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
in *[sic]* without an active VHL treatment…".	without an active VHL treatment…".		
In section 6.2.1.3 in Table 6.7, the results presented here appear to be incorrect and this scenario is unable to be replicated by the company.	This should be corrected with the adjustment factor amended correctly.	We are unable to replicate this scenario and produce the same results as presented by the EAG. The results estimated by the company produce much lower ICERs for this scenario than those presented by the EAG. To remove the perioperative mortality risk adjustment in the cost-effectiveness model, in the "Effectiveness" sheet cells P69 to U69 should be set to 1.00 (instead of 2.00). This gives the unadjusted ICERs of £74,881 for the RCC cohort, £64,124 for the CNS Hb cohort and £82,773 for the pNET cohort (combined weighted ICER for the MA population being £71,248).	The company is correct that the description of the scenario is inaccurate. The changes made in this scenario were the following: Change 1: Dropdown list in "Effectiveness O59" set to "No". Change 2: the risk from first surgery to death were doubled (Effectiveness H162, H192 and H221). This was meant to undo the company's adjustment of the "perioperative mortality risks by a factor of 2.0 grounded on clinical expert opinion in an attempt to better reflect the MHRA population which defines Belzutifan treatment appropriate for patients "for whom localised procedures are unsuitable or undesirable", described in page 145 of the EAG report.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			Therefore, in this scenario we did change more than just the perioperative mortality risk adjustment in “Effectiveness – row 69”. This has been clarified in the EAG report. Furthermore, after seeing the company’s next comment below, the EAG understands that making change 1 above would be incorrect. Instead, as explained by the company below, the following changes should be made: In the “Effectiveness” sheet, change: • Cell P64:Q64 from 80% to 4.0% • Cell P65:Q65 from 20% to 24.4% • Cell R66:S66 from 85% to 7.7% • Cell T67:U67 from 100% to 20.0% • Cell T68:U68 from 100% to 0.0% • Cells P69 to U69 from 2.00 to 1.00

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			Therefore, this scenario has been amended and the results changed in the EAG report.
In section 6.2.1.3 in Table 6.8, the results presented here are to be incorrect as this scenario has been incorrectly applied.	This should be corrected with the adjustment factors amended correctly.	The results for this scenario are produced by removing all adjustments to transition probabilities and complication risks to align with the GB population which is done in the cost-effectiveness model in the "Specifications" sheet cell L81 or the "Effectiveness" sheet cell Q59. However, adjustment in this way also changes the model structure permitting three surgeries rather than only one, by also removing the HR adjustment of event-free after surgery → surgery implemented in cells P63 to U63 in the "Effectiveness" sheet of the model. The model structure restricting to only one surgery based on the DP population was not an issue raised by the EAG and it appears that this additional adjustment is a mistake. To implement the scenario presented by the EAG as intended, removal of all adjustments in the risk of short- and long-term complications and perioperative mortality risk, and leaving the model structure to still only allow for one surgery, the following changes should be made to the model: In the "Effectiveness" sheet, change: • Cell P64:Q64 from 80% to 4.0% • Cell P65:Q65 from 20% to 24.4%	We thank the company for this additional clarification. We were not aware that this change would also change the structure of the model to allow more than one surgery. This scenario has been amended as the company suggested.

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
		• Cell R66:S66 from 85% to 7.7% • Cell T67:U67 from 100% to 20.0% • Cell T68:U68 from 100% to 0.0% • Cells P69 to U69 from 2.00 to 1.00 Implementing this scenario as described above produces the unadjusted ICERs of £163,133 for the RCC cohort, £117,145 for the CNS Hb cohort and £131,043 for the pNET cohort (combined weighted ICER for the MA population being £137,729).	
The EAG report does not document any EAG engagements with clinicians to test/validate assumptions.	Provide clarity on EAG engagements with clinicians and if any advice was sought to validate assumptions used in the economic analysis.	The EAG report refer to certain company assumptions as “arbitrary”; however, there does not appear to have been any engagement by the EAG with clinical experts with experience in treatment of VHL documented in the report to validate or test the company’s assumptions.	Not a factual inaccuracy. The EAG’s considered some of the company’s assumptions as arbitrary due to the lack of justification for them. The EAG can confirm that it did not engage with clinicians to validate assumptions. As mentioned for example on page 27 of the EAG report “The scenario analyses conducted by the EAG were mostly explorative given the lack of other sources of evidence and many of the alternative assumptions

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
			explored were arbitrarily selected." Please note that lack of clinical expert engagement for the EAG's exploratory assumptions cannot be a reason for not providing adequate justification, ideally evidence based, for the company's main assumptions.

References

1. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-7.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **28 September 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	Carl Selya-Hammer
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Stakeholder
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	N/A (we are the company)
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry	None

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Company opening comment	No	In response to Technical Engagement, MSD has submitted a Patient Access Scheme (PAS) [redacted]. The impact of this is an ICER range between [redacted] or a plausible ICER range of [redacted] per QALY gained, all with severity-modifier adjustment. The company acknowledges the challenges in this technology appraisal due to the single arm trial design, the rarity of the disease and the complex and heterogenous nature of the patients eligible for treatment. However, the clinical trial data are compelling, the ICER estimates are plausible, there is substantial value not captured in the economic model. Belzutifan is therefore plausibly cost-effective and should be recommended to the Cancer Drugs Fund following the first NICE appraisal committee meeting. In line with the new NICE Manual, paragraph 6.2.34, this technology appraisal is for both a rare disease and an innovative technology. More specifically, this is a highly heterogenous ultra-rare disease, thereby adding complexity to data collection. As such we kindly request the committee accept the inherent uncertainty in the data package, on the basis that the decision risk for the NHS is low due to patient numbers and the availability of a Patient Access Scheme (PAS). For patients, this disease is incredibly risky, the impact of not recommending this treatment is a propagation of this risk. We note both aspects: risk to NHS and patients can be mitigated by a rapid Cancer Drugs Fund (CDF) recommendation. The company accepts some limitations in the economic model. However, we do not accept these prevent a positive CDF recommendation, we note also substantial burden/benefit that the economic model in its current form was not able to capture.

		We note also paragraph 6.2.33, *'… the committee should take into account the likelihood of decision error and its consequence for patients and the NHS. There should be an explicit reference to the potential benefits and risks to patients based on the level of decision uncertainty and whether this can or cannot be mitigated. The committee should also consider the risks to the NHS of using the technology, based on the most plausible ICER and the impact of adopting the technology on NHS resources.'* Please see Appendix 1, where this is discussed in more detail. A complicating factor of this appraisal is that rarity is accompanied by heterogeneity in disease trajectory. Modelling the complexities of the disease, with a limited dataset, has been challenging and we believe the best place for this technology is in the CDF where additional data collection would immediately help to address the uncertainties. The PAS will mitigate some uncertainty; however, we must highlight that there is still significant uncaptured value of belzutifan resulting in an underestimate of its cost-effectiveness versus standard of care (SoC). The results presented within this document now reflect the submitted PAS and are therefore marked commercial-in-confidence. [redacted] The company's aim is achieving a recommendation for use in the CDF after the first Appraisal Committee Meeting (ACM). Whilst we acknowledge that this can only be determined after the committee discussion has taken place, we have demonstrated the plausibility of the cost-effectiveness of belzutifan and are confident that the additional uncertainties will best be resolved through data collection via the CDF. Importantly, no additional data from the clinical trial or otherwise will become available between the first and second ACM and therefore there is limited value in a second ACM. This appraisal process has been delayed far beyond conventional timelines ultimately delaying patient access to this first-licensed treatment for VHL.
Key issue 1: Implication of differences between intervention and	Yes	*We have grouped key issues 1, 2 and 6 into one response as they all refer to the misalignment between the decision problem population versus the clinical trial population and how this impacts the economic modelling.* **How the final marketing authorisation was reached**

comparator populations given interpretation of the MA that standard of care for most patients is immediate surgery Key issue 2: Misalignment between the decision problem and MK-6482-004 study populations; and between the latter and the UK target population Key issue 6: There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses		• Belzutifan was the first product filed for regulatory approval in the UK via the Orbis route, ie a regulatory application linking the MHRA process with the FDA regulatory process • An application was originally made to the MHRA for an indication aligned to that of the MK-6482-004 study population i.e., in adult patients with VHL disease associated renal cell carcinoma, not requiring immediate surgery. • The MHRA requested MSD provide further evidence of belzutifan's benefit over current local RCC management strategies and procedures i.e., ablative procedures. • MSD provided evidence and proposed a narrowing of the indication to "adult patients with VHL disease who require therapy for the associated renal cell carcinoma, not requiring immediate surgery" which MSD defined as patients for whom surgical intervention is inevitable in the foreseeable future and/or surgery is not a preferred option for these patients. • At the same time, as a result of the FDA conditional marketing authorisation, MSD also proposed expanding the indication to include "adult patients with VHL disease who require therapy for associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), • MSD suggested that the broadened indication, along with the addition of "who require therapy" to the indication wording would provide access to belzutifan for more patients who are in need of a non-surgical/non-procedural treatment option. • The MHRA additionally restricted to patients "for whom localized procedures are unsuitable or undesirable" and did not include the "not requiring immediate surgery" wording. The MHRA therefore granted a conditional marketing authorisation, the final indication i**s** "Welireg (belzutifan) is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable"

		Generalisability of the MK-6482-004 study to the decision problem (and the indication wording) In Key Issue 2, the EAG raise the issue of generalisability of the findings of the MK-6482-004 study to the decision problem population and the trial and the UK target population. The decision problem is, as required by NICE processes, aligned with the regulatory indication wording. The MHRA indicated population is the same as the UK target population. As explained above, the regulatory indication wording reflects a slightly different population to that in the MK-6482-004 study. VHL is an ultra-rare and highly heterogenous disease, consequently there is inevitable complexity in the progression of VHL disease that needs to be addressed. RCC is very common in patients with VHL disease such that patients with VHL disease-associated CNS-Hb or VHL disease-associated pNETs are highly likely to also present with RCC tumours. Therefore, the results of the MK-6482-004 study can be considered representative of patients in UK clinical practice who present with CNS-Hb and/or pNETs. Clinical experts also confirm that the population included in the MK-6482-004 study is representative of the VHL disease population in the UK. Furthermore, clinical experts have provided feedback indicating they are confident in being able to clearly identify belzutifan-eligible patients as per the indication wording. The EAG have raised concerns that at least some patients recruited to the MK-6482-004 study had less severe disease compared to those in the decision problem population, given that the study excluded patients who required immediate surgery or had inadequate organ function. Such exclusions are a common requirement of clinical trial protocol in a novel product. It would be unethical to deny surgery to patients with immediate need for surgery or administer a systemic treatment to patients with inadequate organ function in a trial for a not-yet-approved product. Hence, some patients with less severe disease in the clinical trial is an artefact of clinical trial requirements. There is no evidence that a slightly “sicker” patient would not benefit from same treatment effect with belzutifan. Clinical experts are clear on the severity of disease manifestations in the decision problem population and they define these patients as being at the “end of the road” in terms of their remaining treatment options. They have provided feedback that belzutifan-eligible patients would suffer loss of organ function or functional impairment on their next surgery, despite their urgent need for this surgery (see expert elicitation in Appendix 2). They also affirm the generalisability of the trial data to the UK target population.

		Defining SoC in the decision problem population In Key Issue 1, the EAG raise the concern that the populations implied by the initiation of immediate surgery for the comparator arm and no immediate surgery for belzutifan appear to be incompatible. By definition, as highlighted by clinical experts, patients "who require therapy" are no longer suitable for active surveillance and require an intervention. In the absence of belzutifan, in current SoC, this intervention is surgery for the UK target population. By that notion, it is technically appropriate for all patients in the comparator arm to have surgery. The model reflects the treatment point, which is a simplifying and transparent approach reflecting immediate surgery as they have reached the point when they require therapy. We have engaged further with clinical experts after the EAG raised the concern around the plausibility of delaying surgery in the comparator arm in the cost-effectiveness model. The experts were strongly against this, as it goes against gold-standard practice (see Appendix 2). We note this is different to a patient's lived reality. It is unlikely VHL patients would go to immediate surgery, but any 'delay' between treatment decision and treatment is an artefact of practicalities and NHS scheduling that would be difficult to model accurately and not the focus of the cost-effectiveness analysis examined in this appraisal. The EAG propose that belzutifan-treated patients should also receive immediate surgery to reduce any "immediate severe harm". No belzutifan-treated patients receive immediate surgery in the economic analysis given they have recourse to an effective therapy. Furthermore, results from the MK-6482-004 trial indicate that belzutifan's onset of efficacy is rapid, detectable by the first scan at 12 weeks. Figure 1 (below) demonstrates that between the baseline visit and the first scan following belzutifan treatment initiation [redacted] patients saw a reduction in the total sum of RCC target lesions diameters. The aim of both surgery and belzutifan is to prevent progression to metastatic disease and/or relieve symptomatic burden. Surgery does this via resection of the tumour; belzutifan does this via shrinking the tumour which results indicate occurs very rapidly after treatment initiation. The suggestion that belzutifan should be given alongside or immediately after a surgical intervention would render belzutifan an adjuvant therapy and is not the indication explored in the MK-6482-004 trial nor stated in the GB marketing authorisation. In fact, the opposite is stated in the marketing authorisation (for whom localised procedures are unsuitable or undesirable). The EAG also explore no immediate surgery for the comparator arm essentially rendering the comparator arm active surveillance. For a target population "who require therapy", this is an implausible scenario.

Figure 1 Spider plot – Percentage Change in Total Sum of RCC Target Lesion Diameters from Baseline in Scan before and After Treatment – Investigator Assessment (Efficacy Analysis Set)

Date of Data Cut-off: 01APR2022
Note: Dots indicate scan assessments

This figure presents the size of RCC tumours based on scan taken at times before and after initiation of treatment with belzutifan in the MK-6482-004 study. Tumour sizes are expressed in terms of the percentage difference in diameter relative to what it was at the pint of initiation of treatment (time 0). Patient data from the period prior to treatment initiation with belzutifan show progressive increase in the size of tumours over time. In contrast, following initiation of treatment with belzutifan the density of points with a negative change in diameter (i.e. indicating decreased diameter) demonstrates that the substantial majority of RCC tumours

		reduced in size, and that this happened very soon after initiation of treatment with belzutifan and was visible in the majority of cases by the first scan following treatment initiation (at 12 weeks following treatment initiation). **Modelling the decision problem population** Key Issue 6 raises the concern around the modelling of the decision problem population given the misalignment to the sources of evidence used to inform these analyses. This is an issue only if the effect size of the treatment benefit of belzutifan (in particular in terms of objective response rate) in the population of the MK-6482-004 study differs to that in the population specified in the GB MA indication and decision problem. As described in the section "How the final marketing authorisation was reached", the marketing authorisation was initially sought for VHL-associated RCC only in line with the MK-6482-004 trial. The model was initially developed in line with this and then expanded to include VHL-associated CNS Hb and pNET once the final marketing authorisation was granted. The model distinguishes between tumour types to reflect the relevant outcomes for different types of tumour manifestation e.g. an RCC surgery leading to dialysis versus a pNET surgery leading to type 3c diabetes. These cohorts are not discrete subgroups, they are different manifestations of the same disease and patients rarely only have one tumour type. From the MK-6482-004 trial, at least 90% of patients had a concurrent tumour manifestation in addition to RCC. As deemed by the regulator, the MK-6482-004 study provided sufficient evidence to support the efficacy of belzutifan in CNS Hb & pNET in addition to RCC, in the absence of any other data, the same data source is considered appropriate evidence to inform economic modelling in these tumour types. Key Issue 1 raises the concern around the proportion of patients receiving immediate surgery in the comparator arm. Formal elicitation methods were not used to determine the proportions receiving immediate surgery in the SoC arm; however, per the wording of the label and as informed by expert opinion (described in the "Defining SoC in the decision problem" section above), *all* patients require surgery in the absence of belzutifan. However, we conservatively allow a 10% allowance for active surveillance and this allowance is increased in scenario analyses and presented in Table 5. Key Issue 6 also considers the adjustments made to perioperative mortality and surgical complications risks. Clinical expert opinion informs that although the original source data used to estimate these risks are specific

		to VHL disease, they do not reflect the severity of the decision problem population. Clinicians suggest that, due to the rarity of disease, they only see a handful of patients fitting the target population in practice. This makes it challenging to draw a definitive conclusion on the magnitude of increase on these risks; however, when presented with those used in the company base-case experts have validated that they are reasonable (see Appendix 2). The company base case explores doubling the perioperative mortality and surgical complications risks which is reflected in both arms of the model. Clinical experts also inform that certain surgical complications are metabolic consequences of surgery, an inevitable outcome of "end of the road" surgery rather than a risk only for some patients (see Appendix 2). Below, we have also provided a scenario analysis on a x1.5 magnitude increase for information presented in scenario 2 of Table 5. We are providing this scenario even though the company considers them an underestimate and not reflective of the target population. **Uncaptured benefits in the economic model** The company acknowledges the challenges in the economic modelling of VHL. Whilst we have made every effort to consider the myriad aspects of VHL disease, VHL is highly heterogenous and due to data gaps and model limitations, there are aspects of patients' experience in VHL that have not been captured. The EAG's exploratory scenarios suggest that the modelling approach features optimistic assumptions, leading to more favourable ICERs. In fact, given the aspects of VHL disease not captured in the economic model the ICERs likely underestimate the true cost-effectiveness of belzutifan. A critical uncaptured element in the economic model is accurate assessment of the impact of multi-system tumours and the value of multi-system treatment. Currently a simplifying and transparent assumption is used that undervalues the real impact. For many patients VHL disease is characterised by spontaneous new tumour growth in multiple organs throughout a patient's lifetime, often termed a 'snowballing' effect. Belzutifan has shown the potential to have a receding impact on this snowballing. Very simply, VHL is more than the sum of its tumour manifestations. Moreover, VHL disease often results in multiple concurrent tumours in different organs and evidence from the clinical trial demonstrates belzutifan shrinks multiple tumours simultaneously therefore avoiding the need for multiple surgeries. The currently available data and the resultant cohort model structure do not consider the impact of multiple simultaneous tumours that are "more than the sum of their parts". That is to say that only considering multiple tumour manifestations additively does not capture the full impact of multi-systemic

		disease. Hence, belzutifan's impact on these simultaneous tumour manifestations is unable to be captured in the model and results in a conservative ICER estimate. Another critical element that is poorly captured in the economic model is the impact on the family of the current burden of disease: multiple people in the same family affected by the disease with wildly different and unpredictable presentations, currently with very limited treatment options. And the positive impact to whole families' clinical outcomes and broader HRQoL that the availability of a novel treatment will have. As stated in our opening statement, rapid data collection in the CDF will help to address the uncertainties in modelling the decision problem population. Particularly relating to patient baseline characteristics including prior VHL-related interventions to help further define the surgery unsuitable/undesirable population and belzutifan efficacy in this population, measured by both response and time to event(s), treatment duration and subsequent therapies. Moreover, we have provided a PAS discount to strengthen cost-effectiveness and acknowledge uncertainties.
Key issue 3: Potential risk of study selection bias resulting in possible omission of relevant comparator studies	Yes	**With regard to the 26 records that were initially included in the clinical effectiveness systematic literature review:** The full bibliographic details of the 26 records identified in the clinical effectiveness literature systematic review were provided in Appendix D Table 110 of the company submission, we acknowledge that this could have been pointed out more clearly and more details on these studies could have been provided. Detailed information extracted from these studies have now been provided in Appendix 3 of this response form. From this information it is evident that: • No information on the clinical effectiveness of belzutifan in the indication under assessment in addition to what is provided by the (single-arm) MK-6482-004 study is available. Therefore, no trial-based indirect treatment comparison belzutifan with other regimens is possible. • None of the studies that investigated other (non-belzutifan) treatment regimens provide data that is representative of the overall standard of care in current UK clinical practice for the specific indication

		under assessment/target population for belzutifan. Specifically, they either report information on treatments not used in UK clinical practice for the indication/target population (e.g. sunitinib, pazopanib, etc.), report data focused only on a specific treatment that alone would not be representative of overall standard of care in UK clinical practice (e.g. image-guided ablation, linear accelerator-based radiosurgery, etc.), or were conducted in countries where the clinical management landscape is unlikely to be representative of that in the UK (e.g. China, etc.). Therefore, matching-adjusted indirect comparison using data from any of these "comparator" studies will not yield results that are representative of the overall standard of care in current UK clinical practice for the specific indication/target population under assessment as part of this appraisal. Consequently, a large retrospective observational study was commissioned by MSD to provide information for the comparator arm prior to the final marketing authorisation, this was the VHL Natural History Study that is described in section B.2.9 and Appendix Q of the company submission (and also reference #46 in the company submission, confidential MSD data on file, document submitted during the clarification stage (1)). **With regard to the non-English records excluded from the clinical effectiveness systematic literature review:** Section 3.1.2 of the EAG report noted that the company did not provide specific details of the two studies excluded from the clinical effectiveness systematic literature review due to those not being English language publications at the abstract screening stage, and therefore it was not possible for the EAG to determine the impact of these omissions on clinical effectiveness estimates. MSD would have happily addressed this had it been requested at clarification stage. The English-language abstracts for these two records are provided in Appendix 4 of this response form. From these it is evident that neither record provides information on belzutifan, nor would they present data that is representative of the overall standard of care in current UK clinical practice for the specific indication/target population under assessment as part of this appraisal: • One record is for a review of the epidemiology, treatment, and prognosis of 19 VHL disease-associated CNS hemangioblastoma patients at one centre in Lille, France.

		• The other record is for a review of the factors related to tumour recurrence or development of new tumours in 35 patients with hemangioblastomas of the brain (including 21 patients with VHL disease) treated specifically with leksell gamma knife at one centre in Shanghai, China. **With regard to whether it was clear that the search strategy and study selection criteria were designed to identify all relevant interventional, non-interventional and natural history studies:** The EAG states in section 3.1.1 of their report that "searches were well structured, transparent and reproducible, and a good range of subject indexing terms (MeSH/EMTREE) and free text was used". The selection criteria presented in Appendix D of the company submission explicitly state that the following study types would be included in the SLR: • Randomised controlled trials (RCTs) • Controlled clinical trials • Non-randomized clinical trials, including single-arm prospective interventional trials • Prospective and retrospective cohort studies These clearly encompass, and would necessitate the inclusion of, any relevant interventional and non-interventional trial studies and so there is no ambiguity in this regard. Accordingly, the VHL Natural History Study is a retrospective non-interventional study and would have been identified and included in the SLR based on these searches and criteria if it had been a published study at the time of the SLR search. General "natural history studies" of relevance to this appraisal would also be prospective or retrospective cohort studies, and so would also be included in the SLR. Furthermore, the full bibliographic details of each of the 5 studies in total excluded at the full-text screening phase on the basis of study design are provided for transparency in Appendix D Table 111 of the company submission, and therefore how the SLR included/excluded studies based on study design (including in terms of "relevant interventional, non-interventional and natural history studies") can be evaluated in a clear and transparent manner.

Key issue 4: lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator	Yes	The MAIC comparing belzutifan to SOC using data from the treatment period of the MK-6482-004 study and the VHL Natural History Study was the only adjusted comparison performed and presented in the clinical effectiveness section. While other sources of clinical effectiveness data were available and mentioned in the company submission cost-effectiveness section, it was not appropriate to include these in MAIC for comparisons versus belzutifan for the reasons discussed below. These sources were: • Data from a retrospective analysis of the pre-treatment phase of the MK-6482-004 study: The MAIC is required to provide transition rates for use in the economic model, for endpoints of interest i.e. rates of transition pre-surgery → 1st surgery/metastasis/death. Analyses for certain endpoints are not available in the pre-treatment data, for example, "pre-surgery → death" or "pre-surgery → metastasis" in all 61 patients will be right censored since they are all necessarily alive and metastasis-free at the beginning of the treatment period/end of the pre-treatment period. Hence, only data on the rates of pre-surgery → 1st surgery were calculated from pre-treatment period data and these data are used in the economic model. No ITC/MAIC (i.e. applying matching or reweighting) comparing belzutifan to SOC is required when using the data from this source. The patient population of the retrospective analysis of the pre-treatment phase of the MK-6482-004 is fully comparable to, in fact is the population, of the MK-6482-004 study period. In the CNS-Hb and pNET populations, an important limitation of the VHL Natural History Study data was the inability to identify patients who had CNS-Hb and pNET tumours at the patient-level index date. Consequently, for these populations, the best available data source for pre-surgery → first surgery was the surgery event data collected for MK-6482-004 trial participants during the pre-treatment period of the MK-6482-004 study, as patients' CNS-Hb and pNET tumour status was identifiable at the baseline visit of the MK-6482-04 study. • Data from the Optum Clinformatics Data Mart Claims Study:

		This study was designed to generate key model inputs such as surgical complication risks (as well as to address other objectives unrelated to the economic model). The study quantified the prevalence, healthcare resource utilisation (HRU), and healthcare costs of VHL-RCC in the United States (US) via a claims-based algorithm that identified patients with VHL-RCC from a real-world database. For the purposes of deriving inputs for the economic model, data on the rates of event-free after 1st surgery → next surgery, event-free after 1st surgery → metastatic disease, and event-free after 1st surgery → death in patients with certain VHL-disease associated tumours were selected to be calculated from this study, which were not clinical efficacy outcomes measured in the MK-6482-004 study. This study was not designed for comparability with the MK-6482-004 study. The cohort from this study cannot be directly repurposed for use to compare against data on patients treated with belzutifan from the MK-6482-004 study since the index date definition and related eligibility criteria of this Optum study were not designed specifically to be aligned with those in the MK-6482-004 study. Furthermore, as the outcomes calculated from this study (i.e. rates of events after 1st surgery) are not the same as the outcomes collected from the MK-6482-004 study (i.e. rates of events from pre-surgery), it is not appropriate or feasible to compare/combine the data from these two studies in ITC/MAIC or for the construction of an external control arm. Additional significant limitations associated with using the Optum database to construct such an external control arm for MAIC include: o The very limited availability of matching variables (e.g. baseline renal tumour size would not be available). o The limited ability to measure the number of prior surgeries (as this would be dependent on each patient's length of continuous enrolment in the database). o The lack of a VHL ICD code for patient identification (a VHL ICD code was introduced recently, but its use in claims databases is still limited), which means that a VHL disease identification algorithm would have to be developed, the validation of which would have to rely on physician notes which may not be fully accurate due to the lack of genetic testing data, as well as claims data which do not necessarily contain patients' full medical histories.

		Based on the points noted above, the adjusted/re-weighted results from the VHL Natural History are expected to yield model parameters for a control arm that are well-matched to the MK-6482-004 study in terms of patient baseline characteristics, more so than any control arm that could be constructed using the Optum data. The methodologies associated with these two data sources are described in section B.3.3 of the company submission, with further details on the pre-treatment phase of the MK-6482-004 study available from the Wang 2023 publication (2), and further details on the Optum study available from the Jonasch 2022 publication (reference #41 in the company submission) (3). The methodology of the VHL Natural History Study and how the matching/reweighting to align with the MK-6482-004 study population was performed is described in section B.2.9 and Appendix Q of the company submission, with further details provided in the final study report of the VHL Natural Study (reference #46 in the company submission, confidential MSD data on file, document submitted during the clarification stage) (1). The transition rate data that were calculated from all three data sources (the MK-6482-004 study, the VHL Natural History Study, and the Optum Clinformatics Data Mart Claims Study) are also compiled and can be compared in Appendix 5.
Key issue 5: Limitations in the indirect treatment comparison hinder the assessment of the effectiveness of belzutifan compared to standard of care	No	Additional details on the justifications for the three aspects of the ITC methodology highlighted by the EAG are provided below: **Method of adjustment for confounding** IPD from the VHL Natural History Study and IPD from MK-6482-004 were stored, managed, and analysed separately (at IQVIA and MSD, respectively). Due to compliance reasons IPD from both studies could not be held and analysed by a third party. Consequently, adjustment methods requiring pooled IPD from both sources, such as traditional inverse probability of treatment weighting (IPTW) (i.e., propensity score weighting based on pooled IPD), could not be conducted. Instead, statisticians at MSD analysed MK-6482-004 IPD to estimate parameter inputs needed for the belzutifan model arm, while the VHL Natural History Study and the derivation of population-matched (to the MK-6482-004 study) results from it was performed by the research group at IQVIA, who had access to IPD for

		the VHL Natural History Study cohort but only summary-level results from MK-6482-004. IQVIA did not provide any comparison of matched/reweighted results of the VHL Natural History Study versus the results of the MK-6482-004 study. At the direction of MSD, IQVIA obtained MAIC-based propensity score weights for each target population of interest, and then used the reweighted sample to estimate an analogous set of parameter inputs for the SOC model arm. It was thus feasible to conduct parallel IPD analyses of each source individually, with close coordination between the two organisations to ensure consistency of statistical approaches and study variable definitions. While this deviates from the academic ideal as described in NICE DSU TSD 18, this is nonetheless a very methodologically rigorous approach and the best that is possible while working within these practical and compliance-related limitations. **Choice of characteristics used in propensity score reweighting** The purpose of the MAIC was primarily to inform inputs for the cost-effectiveness analyses, at the design stage of these analyses, a clinical panel that included two expert physicians, one from the Department of Genitourinary Medical Oncology of the University of Texas MD Anderson Cancer Center and one from MSD, was arranged to validate the conceptual model and other aspects of the model development plan. To inform the set of baseline variables to use for matching in the MAIC, expert input was sought from these experts regarding baseline characteristics that are likely to be prognostic of transition probabilities starting from the pre-surgery state, or that may modify the effect of belzutifan on these transition probabilities. In the VHL-RCC cohort, the MAIC adjusted for the covariates that MSD's consulted clinical experts identified as the most important prognostic factors or effect modifiers for RCC surgery. The set of matched covariates included size of the largest renal tumour at baseline, the predominant risk factor for renal surgery (given that size of the largest renal tumour dictates the need for renal surgery). This covariate is also a strong, well-established risk factor for metastases; in fact, the 3 cm size threshold that warrants renal surgery originates from data that showed a high risk of metastases among patients with renal tumours exceeding 3 cm. The inclusion of further baseline covariates in addition to size of the largest renal tumour and other matched covariates (age, sex, number of prior renal surgeries) is unlikely to have had a meaningful impact on the effect estimates. Other variables that could not be matched included: VHL type and VHL gene alteration type (which

		experts suggested only as potentially relevant) and number of measured tumours (less directly relevant than largest renal tumour size for predicting renal surgery and metastases). As with any non-randomised comparative effectiveness analysis, there is a risk of residual confounding from unobserved between-group differences in patient characteristics. However, prior to matching, the largest renal tumour size and renal surgery burden at baseline were lower in the VHL Natural History Study cohort than the MK-6482-004 cohort, suggesting that the direction of any residual confounding may be in favour of the active surveillance arm. Moreover, the relative effect size of belzutifan vs. active surveillance on the pre-surgery → surgery rate (85% reduction) was too large to be reasonably attributed to confounding. This relative effect size of belzutifan vs. active surveillance (as implied by the MAIC) is also corroborated by a study that compared observed surgery rates before vs. after patients initiated belzutifan in MK-6482-004 (2). **Choice of outcomes for which adjusted analyses were performed** Adjusted analyses were performed for all outcomes that could inform the rate of relevant state transitions used in the cost -effectiveness model for which data were available. The outcomes for which such analyses were possible and results derived (from the MK-6482-004 study and VHL Natural History Study as well as others for which data are available) are shown in Appendix 5, and include cause-specific hazards of pre-surgery → metastatic disease and pre-surgery → death for all three cohorts.
Key issue 7: The comparator data might not be representative for the UK	No	The company submission acknowledged that that both the MK-6482-004 trial and VHL Natural History Study patients likely received elevated care compared to routine UK practice. This was addressed using data from the Optum study, a real-world evidence study which analysed treatment patterns and healthcare resource use associated with VHL using claims data. The EAG has suggested this adjustment be removed, which would improve the cost effectiveness of belzutifan, as shown in Table 5. Clinical experts validated the use of this data source to adjust surgery and metastases rates in the economic model and these adjustments were made to both arms of the model. Furthermore, removal of the Optum study adjustment improves the ICER so has no material impact on decision making. The EAG note inherent uncertainties resulting from using data originating from US clinical practice rather than using UK data. The company acknowledges this is a limitation, however, use of international data is common in technology appraisals, particularly in ultra rare diseases, and does not preclude UK HTA bodies from

		making access decisions. Again, a rapid CDF recommendation would allow data collection to begin to address these concerns. **The EAG's noted inability to explore all uncertainties due to lack of data** In the note on the alternative approach suggested by the EAG, it is stated that "Not all uncertainties mentioned could be explored by the EAG due to lack of data" in key issues 6, 7 and 9. VHL is a rare disease, and the decision problem population is a subpopulation of this rare disease. The NICE manual regarding structured decision making (paragraph 6.2.34) states: *"The committee will be mindful that there are certain technologies or populations for which evidence generation is particularly difficult because they are:* ***rare disease****, for use in a population that is predominantly children (under 18 years old),* ***innovative*** *and complex technologies. In these specific circumstances, the committee may be able to make recommendations accepting a higher degree of uncertainty. The committee will consider how the nature of the condition or technology(s) affects the ability to generate high-quality evidence before applying greater flexibility."* MSD kindly requests that the committee take the rarity and heterogeneity of VHL into account when evaluating belzutifan in VHL disease, considering the challenges in generating high-quality evidence given the rarity of VHL disease and the urgent need for effective therapies necessitating a pragmatic decision.
Key issue 8: Data to inform effectiveness in the belzutifan arm (MK-6482-004 trial) are either immature or unavailable	No	The TA reports a data cut with a median duration of follow-up of 37.7 months, which should not be considered insufficient data in a novel technology. The company is concerned this issue is raised in the report for three reasons: Firstly, the EAG's characterisation of the MK-6482-004 trial data as immature indicates that the EAG is viewing VHL disease through the lens of a typical cancer; VHL is a chronic life-long genetic disorder presenting as spontaneous growth of tumours which can be benign and/or malignant. Hence, rather than referring to data

		immaturity, the focus should be around the number of surgery events that have occurred within the follow-up period compared to an equivalent timeframe in the patients' prior history of surgeries. Secondly, uncertainty around belzutifan long-term effectiveness can be resolved by future data collection. The best path to reimbursement for belzutifan would be funding via the CDF, given the CDF's mechanism to collect the important data to resolve uncertainty around effectiveness, treatment duration and treatment effect waning in the label population. MSD are preparing a proposal, a draft of which we have already asked for review by NHS England and the NICE Managed Access Team, to expedite a recommendation into CDF following ACM1. The managed access proposal describes the outcomes to be collected in the CDF. Additionally, the MK-6482-004 trial will continue to report data read outs annually until 2026, so providing further efficacy and safety data. A positive consequence of being the CDF will be that the company will have time to complete a retrospective review of patients in the UK who would be treated with belzutifan during the managed access period to further clarify the treatment effect of belzutifan in the decision problem population. While the CDF would not be a mechanism through which data on SOC can be collected, the period time belzutifan is in the CDF can be used to conduct the appropriate additional studies to better inform the SOC arm for the cost-effectiveness analyses in time for the CDF exit review. Lastly, the EAG explore the use of alternative parametric distributions for time to surgery, metastases, or death and present results for the RCC cohort in scenario analyses set 3 in Table 6.12. The use of Gompertz in the belzutifan arm is implausible as it leads to crossing curves between the belzutifan and SoC arms as mentioned in response to B2a) Table 6 in the clarification letter. In the absence of an EAG-preferred parametric function, the company base-case assumption should be accepted. Furthermore, the alternatives explored by the EAG produce results that show belzutifan to remain consistently cost-effective with the provided PAS.
Key issue 9: There is uncertainty in the derivation of the transition probabilities in the	No	To clarify, as stated in response to Key Issue 4, the pre-treatment period cannot be used to estimate transition probabilities for certain endpoints, namely metastases and death, as patients had to be alive and metastases-free to be eligible to participate in the trial. Hence, a full set of transition probabilities cannot be generated from this retrospective analysis and the VHL Natural History Study is therefore used. The EAG also questioned why the pre-treatment period was not used to estimate time to surgery (TTS) across all cohorts. The VHL Natural

standard of care arm		History Study is the best available source of evidence for the RCC cohort from which we can generate a full set of transition probabilities, giving the RCC cohort internal consistency. However, the company acknowledges the EAGs query regarding using the same source for all TTS estimates across the three cohorts. We present our revised base-case using a consistent method to estimate TTS across cohorts in Table 5. This results in an ICER [redacted] As part of their critique, the EAG explore doubling the risk of pre-surgery → surgery in the SoC arm for the VHL-CNS Hb & VHL-pNET cohorts in scenario analyses set 4, with results presented in Table 6.13. The stated rationale is that the surgery rates are approximately double in the VHL Natural History Study than with the pre-treatment period in the VHL-RCC cohort. It is assumed that this translates to the VHL-CNS Hb & VHL-pNET cohorts as well. MSD considers that this assumption does not take into account that an individual can have multiple tumours across two kidneys which is quite different for pancreatic and CNS surgeries. Running this scenario, under the revised company base-case, the ICER remains cost-effective as presented in scenario 10 of Table 5.
Key issue 10: There is uncertainty in the implementation of time on treatment and treatment effect waning	No	MSD accept the point that long-term time on treatment and treatment effect waning is an area of uncertainty and propose that future data collection via the CDF and future data read outs from the MK-6482-004 trial would help resolve this. **Time on treatment** The EAG explore the use of alternative parametric models to extrapolate time on treatment in each cohort in scenario analyses set 5 in Table 6.14. The log-normal distribution is implausible and the worst-fitting distribution by visual and statistical fit. Furthermore, the median time on treatment observed in the trial at the 01-APR-2022 data cut-off was [redacted] years whilst the modelled median time on treatment using a Gompertz distribution is [redacted] years and is therefore a conservative estimate. **Treatment effect waning**

		The EAG explored alternative arbitrary assumptions on the period of residual treatment benefit in Table 6.15 and when reproduced under the revised company base case (scenario 11 in Table 5) the ICERs remain cost-effective.
Key issue 11: There is uncertainty in the derivation and implementation of health-related quality of life in the model	Yes	**Mismatch between the decision problem population and evidence used to inform health-related quality of life in the model** We appreciate the EAG highlighting this concern and direct to the response to Key Issue 1, 2 and 6 around misalignment of the trial and decision problem populations. This response can be extended to the VHL real-world (RW) quality of life (QoL) disease burden study which was commissioned prior to the MHRA granting marketing authorisation which included the wording "for whom localised procedures are unsuitable or undesirable". Despite the study not being limited to such patients, it represents the best available evidence for quality of life in VHL patients as no other studies were identified in the SLR. This was also validated by clinical experts (see Appendix 2). The EAG raise questions regarding the face validity of the utility values and explore an adjustment in reducing the utility values by 20% in both arms in scenario analyses set 6. The company is not clear that there is a rationale for this adjustment beyond exploratory analysis, however we note the results from this scenario remain cost-effective (scenario 12 in Table 5). **Immediate health-related quality of life benefit for the belzutifan arm** It could be argued that there is an immediate quality of life benefit on taking belzutifan in reduced anxiety before a confirmed response is realised, as indicated by the efficacy of belzutifan at the first scan (Figure 1). However, we accept the critique on this issue and have implemented in the EAG scenario analyses set 6 Table 6.17 into the cost-effectiveness model. As explored by the EAG, we have implemented the median time-to-response into the QALY calculation and revised the company base case as demonstrated in Table 3.
Key issue 12: Cost-effectiveness analyses should be based on	No	**Subgroup specific input parameters** As mentioned in response to Key Issue 1, 2 and 6, the VHL cohorts included in the economic model are not mutually exclusive and the marketing authorisation population should be assessed as a whole. A weighted combined ICER has been provided for decision-making. Use of subgroup-specific parameters does not apply for this disease where patients are seldom suffering from only one type of tumour manifestation. The MHRA granted a marketing authorisation beyond VHL-associated RCC based on the data package provided in this

subgroup-specific parameters		appraisal; therefore, NICE are obligated to start any appraisal from the full licensed indication for whom we are seeking reimbursement. **Severity modifier** The EAG use a method of calculating the QALY severity weighting that is flawed in several ways: • Firstly, the EAG use an online tool that is based in the Netherlands and although this tool allows UK specific utility data, it uses a UK value set by derived by Heijink et al. (2011) to estimate general population QALYs. This is inconsistent with the NICE reference case which recommends the Hernández Alava et al. 2017 mapping function, using the 'EEPRU dataset' (Hernández Alava et al. 2020) which was done in the company submission. • Secondly, a UK based online tool has been developed by Schneider et al. (2022) (a collaboration between University of York, University of Sheffield and Lumanity) using the NICE reference case value set and can be accessed via this link https://shiny.york.ac.uk/shortfall/. This tool is in line with the NICE reference case and produces results consistent with those presented in the company submission. This tool has also been used in other EAG assessments conducted by KSR for other NICE appraisals (NICE TA911, ID4036). • Lastly, the approach of using a “likelihood of applicable QALY weight” based on PSA results is mentioned nowhere in the NICE reference case on QALY shortfall analysis and appears to simply be an artefact of the Netherlands-based tool used by the EAG. The section on Decision modifiers in the NICE Manual does not mention use of PSA results to produce such weighting of QALY weights. More specifically, the EAG calculated a QALY weight of 1.7 with 55.9% likelihood and 1.2 with 44.1% likelihood effectively equating to a QALY weight of 1.48. This would not be permitted under the NICE Manual as it is clear that based on an absolute shortfall and/or proportional shortfall calculation it is either the 1.0, 1.2 or 1.7 QALY weight that can apply. • It is inconceivable that a disease as debilitating as VHL, specifically the population in the marketing authorisation, should receive anything other than the 1.7 severity modifier. Any uncertainty in the data to determine this absolutely does not equate to evidence that the disease and specifically the indicated population have less severe disease. Application of anything other than the 1.7 modifier should be considered unreasonable and unfair in the light of the patients’ lived experiences.

		We note that this appraisal has been through the process to be routed via HST on multiple occasions. The final criteria checklist undoubtedly confirms the rarity and severity of this disease as defined by the TSOP. We ask that the committee view this technology through a rare disease lens. Although calculations may appear to produce a lower QALY weight, evidence around the severity of disease in the patient population clearly demonstrates that it would meet the highest severity modifier. Furthermore, the NICE methods guide states that modifiers that cannot be included in the estimated QALYs "*can be taken into account qualitatively through committee discussion or quantitatively through QALY weighting*" (paragraph 6.2.11). We ask that the committee take particular consideration to the rarity of this disease and the challenges in capturing its full impact on patients, noting this is also a highly innovative technology.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

There are no additional issues from the EAR that the company would like to respond to.

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 3 Changes to the company's cost-effectiveness estimate

<table>
<tr><th>Key issue(s) in the EAR that the change relates to</th><th>Company's base case before technical engagement</th><th>Change(s) made in response to technical engagement</th><th>Impact on the company's base-case incremental cost-effectiveness ratio (ICER)</th></tr>
<tr><td>N/A: Inclusion of PAS</td><td>No PAS was applied on submission.</td><td>PAS applied in line with that submitted to PASLU.</td><td>
<table>
<tr><th></th><th>RCC</th><th>CNS Hb</th><th>pNET</th><th>Weighted population</th></tr>
<tr><td>New base-case with PAS, ICER (unadjusted)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Δ from original base-case (unadjusted)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>New base-case severity-modifier (1.7 weight) adjusted ICER</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td></tr>
<tr><td>Key Issue 11: There is uncertainty in the derivation and implementation of HRQoL in the model</td><td>No inclusion of time to response of belzutifan in the QALY calculations</td><td>Incorporate median time to response in the QALY calculations</td><td>
Note includes PAS discount
<table>
<tr><th></th><th>RCC</th><th>CNS Hb</th><th>pNET</th><th>Weighted population</th></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td>
<table>
<tr><td>New base-case with PAS, ICER (unadjusted)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Δ from original base-case (unadjusted)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>New base-case with PAS, severity-modifier (1.7 weight) adjusted ICER</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
</tr>
<tr>
<td>Company’s base case following technical engagement (or revised base case)</td>
<td>Incremental QALYs:
<table>
<tr><td>RCC</td><td>[redacted]</td></tr>
<tr><td>CNS Hb</td><td>[redacted]</td></tr>
<tr><td>pNET</td><td>[redacted]</td></tr>
<tr><td>Weighted population</td><td>[redacted]</td></tr>
</table>
</td>
<td>Incremental costs:
<table>
<tr><td>RCC</td><td>[redacted]</td></tr>
<tr><td>CNS Hb</td><td>[redacted]</td></tr>
<tr><td>pNET</td><td>[redacted]</td></tr>
<tr><td>Weighted population</td><td>[redacted]</td></tr>
</table>
</td>
<td>Company revised base case ICER (with PAS)
<table>
<tr><td></td><td>RCC</td><td>CNS Hb</td><td>pNET</td><td>Weighted population</td></tr>
<tr><td>New base-case with PAS, ICER (unadjusted)</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
</tr>
</table>

			New base-case with PAS, severity-modifier (1.7 weight) adjusted ICER	[redacted]	[redacted]	[redacted]	[redacted]

Deterministic results with revised company base-case

Table 4 Deterministic results with revised company base-case (PAS price)

Technologies	Total costs (£)	Total QALYs	Total LYs	Incremental costs (£)	Incremental QALYs	Incremental LYs	ICER vs. comparator (£/QALY)	Severity-adjusted ICER*
VHL-RCC cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL-CNS Hb cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL-pNET cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL- GB marketing authorisation population (weighted cohort)								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Note: The x1.7 severity weight is used.

Technical engagement response form

Sensitivity analyses around revised base case

Table 5 Key scenario analyses around revised base case (with PAS)

#	Scenario	VHL-RCC	VHL-CNS Hb	VHL-pNET	GB MA VHL-weighted population	GB MA VHL-weighted population (severity-modifier adjusted)
-	Base case (revised)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	Proportion receiving immediate surgery in SoC arm reduced to 80% in RCC & pNET cohorts and to 40% (with additional 40% receiving equivalent sequelae) in CNS Hb cohort	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	Relative risk adjustment of non-metabolic surgical complications & perioperative mortality risk reduced to x1.5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Reduce metabolic consequences risk: ESRD/dialysis for RCC surgery to 60% (consequently CKD increased to 40%), stroke for CNS Hb surgery to 65%, diabetes and immunocompromisation for pNET surgery to 80%.	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Omit the adjustment to risk of surgery and metastasis based on	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

#	Scenario	VHL-RCC	VHL-CNS Hb	VHL-pNET	GB MA VHL-weighted population	GB MA VHL-weighted population (severity-modifier adjusted)
	the Optum Clinformatics Data Mart data.					
5	Pre-treatment period used to estimate pre-surgery → surgery in RCC cohort	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Distribution for pre-surgery → surgery in belzutifan arm for VHL-RCC cohort: Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	Distribution for ToT: Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	EAG scenario analysis set 1 – Table 6.2: Double transition probability from pre-surgery → surgery in belzutifan arm	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9	EAG scenario analysis set 1 – Table 6.9: Allowing only adjustment of risks for ESRD and/or dialysis in the VHL RCC cohort and secondary diabetes and immune-compromisation in the VHL pNET cohort, and cerebral vascular occlusion/stroke in the VHL pNET cohort	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	EAG scenario analysis set 4 – Table 6.13: Risk of pre-surgery →	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

#	Scenario	VHL-RCC	VHL-CNS Hb	VHL-pNET	GB MA VHL-weighted population	GB MA VHL-weighted population (severity-modifier adjusted)
	surgery in SoC arm of the VHL CNS Hb and VHL pNET cohorts increased by 100%					
11	EAG scenario analysis set 5 – Table 6.15: 1.5-year duration of residual treatment benefit	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	EAG scenario analysis set 6 – Table 6.16: 20% reduction of the utility values based on the VHL RW QoL Disease Burden Study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Scenario analyses results in a change in the severity modifier from x1.7 to x1.2.

Figure 2 DSA Tornado diagram for belzutifan vs. SoC with revised company base-case in the VHL-GB MA population (PAS price – ICERs unadjusted)

[redacted]

Probabilistic sensitivity analyses were run using a Monte-Carlo simulation with 1,000 iterations. Results indicate a [redacted] probability of cost-effectiveness of belzutifan vs. SoC with the revised company base-case at a willingness-to-pay threshold of £51,000 per QALY gained as per the x1.7 severity modifier adjustment.

Table 6 Probabilistic results with revised company base-case (PAS price)

Technologies	Total costs (£)	Total QALYs	Total LYs	Incremental costs (£)	Incremental QALYs	Incremental LYs	ICER vs. comparator (£/QALY)	Severity-adjusted ICER*
VHL-RCC cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL-CNS Hb cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL-pNET cohort								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VHL- GB marketing authorisation population (weighted cohort)								
Belzutifan	[redacted]	[redacted]	[redacted]	-	-	-	-	-
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Note: The x1.7 severity weight is used.

Technical engagement response form

Figure 3 PSA Cost-effectiveness plane with revised company base-case in the VHL-GB MA population (PAS price – ICERs unadjusted)

Figure 4 PSA Cost-effectiveness acceptability curve (CEAC) with revised company base-case in the VHL-GB MA population (PAS price – ICERs unadjusted)

References

1. MSD. Von Hippel-Lindau Natural History Study - Final Study Report v1.0. [MSD Data On File] [Internet]. 2021.
2. Wang L, Bensimon AG, Sundaram M, Xu R, Lai Y, Liu Y, et al. Burden of surgeries and surgical complications in patients with Von Hippel Lindau (VHL) disease before and after treatment with belzutifan. Journal of Clinical Oncology. 2023;41(6_suppl):733-.
3. Jonasch E, Song Y, Freimark J, Berman R, Nguyen H, Signorovitch J, et al. Epidemiology and Economic Burden of von Hippel-Lindau Disease-Associated Renal Cell Carcinoma in the United States. Clinical Genitourinary Cancer. 2022.

Appendix 1

The company considered it might be helpful to present its perspective on the consequence of any decision error for patients and the NHS due to uncertainty and how this might be mitigated. Considering the relevant elements of the Structured Decision Making (SDM) considerations for clinical effectiveness and value for money MSD believes the most relevant elements are ensuring the patients and clinical expert insights are appropriately captured.

The most obvious way to mitigate risk to either patients or the NHS is a rapid recommendation into the Cancer Drugs Fund, which will require ongoing data collection and a formal CDF exit, both of which will help address residual uncertainty.

From a patient point of view - acknowledging some uncertainty associated with a single arm study and a lack of data for patients with non-RCC tumours, but with CNS Hb or pNET tumours - the before and after design of the study demonstrates profound clinical effectiveness that offers patients enormous benefit. The technology appears to be well tolerated therefore any adverse event, tolerability risk to patients with belzutifan would be entirely in line with other treatments for similarly complex diseases. There is no evidence of differential objective response rate (ORR) treatment effect by patient type, presentation of disease, disease severity. Therefore again, risk appears to be low and potential benefit of a treatment option is extraordinarily high. We request the committee understands the patients' point of view with regards the potential value of this treatment or the impact on families, multiple generations of which have VHL, of this treatment not being available. Acknowledging these points are 'reserved' for HST technologies we consider them relevant for VHL: the QALY gains associated with this treatment are profound, this is because the disease can be horrific, and the drug has impact across multiple affected systems. QALY gains above 4.0 and even above 5.0 represent a paradigm shift in treatment options. The evidence base is small but the effectiveness estimates are substantial, a

recommendation to CDF will continue to strengthen the evidence base. Disease morbidity and clinical disability can be extremely high in this patient population. Again, we request the committee consults the patient experts to understand the breadth and depth of impact this disease has.

With regards SDM consideration on value for money, we consider the economic model produces plausible cost-effectiveness estimates, particularly with the PAS applied, even acknowledging the challenge with an ultra-rare evidence package and highly heterogeneous disease presentation. Given the very small population size the maximum financial risk [redacted]. Referencing NHS England policy documents, this is barely over the threshold at which the NHS did not review products or activities. [redacted]. As such the total financial decision risk is low. Any long-term opportunity cost associated with recommending a treatment that turned out to have a higher ICER, can be mitigated be a CDF recommendation and a CDF exit in the future.

Additional value for money considerations include uncaptured value in the economic model. Due to the nature of the data package and therefore the required simplicity of the economic model, there are critical elements that are not captured

- Because VHL ‘snowballs’ eg, multiple tumours, multiple interventions, multiple more tumours, more interventions, the impact of the disease if far greater than the sum of its parts. The economic model uses a very simple approach that does not fully capture this multisystem, multiyear impact.

- Belzutifan appears to have effectiveness across not only the three named tumours but also other VHL-associated tumours (we acknowledge very small patient numbers) AND the positive treatment effect occurs simultaneously across tumours in

individual patients. The submitted model has a very ‘one-at-a-time’ approach to tumour response, which does not recognise the value of the treatment.

- VHL affects multiple members of the same family, whilst we have included a family and carer disutility we consider this substantially underestimates the lived reality of some families with multiple people with severe manifestations. As belzutifan appears to have consistent ORR treatment effect within the indicated population, the positive impact on these families, again the benefit is more than the sum of its part, is not captured in the current economic model.

For the reasons described above we believe the committee should accept uncertainty associated with this technology appraisal.

Appendix 2

Expert Elicitation Exercise post-EAG report

Background

MSD has developed a model to determine the cost-effectiveness of belzutifan versus SoC in adults who require therapy for VHL-associated RCC, CNS hemangioblastomas or pNET and for whom localised procedures are unsuitable or undesirable. An advisory board and several 1:1 engagements were conducted with clinical experts as detailed in the company submission; however, the EAG report notes concerns to be further clarified with clinical experts. Hence, we have conducted an additional expert elicitation exercise following the EAG report. These were in the form of 1:1 discussions for two main reasons. Firstly, conducting a structured expert elicitation in relation to this rare disease setting brings feasibility challenges. Clinical experts describe this target population as only a handful of patients they would see in their career as it is a subpopulation of a rare disease; therefore, it is impractical to draw quantitative conclusions for. Secondly, due to time constraints in providing this documentation in time for technical engagement response, it was prioritised to have separate discussions with each expert due to difficulties in finding availability for all experts for a group discussion.

The purpose of this expert elicitation exercise was to gain insight into the key concerns raised by the EAG which they suggest could be clarified by clinical experts.

Selection & background of clinical experts

Three experts were selected based on their experience in managing and treating patients with VHL disease. Details on their role, expertise and conflicts of interest are detailed in Table 7 below.

Table 7 Details of clinical experts

Expert	Role	Expertise	Conflicts of Interest
#1	Consultant Endocrinologist	Experienced consultant endocrinologist with experience working within the largest VHL clinic with the UK. Currently coordinating the care of VHL patients in Sheffield. Respected endocrinologist, [redacted] Has also published extensively in the endocrine cancer (including VHL related tumours), surgery and dysfunction space.	Has provided paid consultancy services to MSD on VHL.
#2	Consultant Physician/Professor of Clinical Endocrinology	[redacted] Well respected clinical expert, and scientific leader. Extensive publication history in endocrine disease and heavily involved in VHL clinical research at key scientific centres including the Francis Crick Institute. Trustee and board member of clinical and patient groups including the VHL UK&I and the Addison's disease group. [redacted]	Has not received any payment from MSD.
#3	Consultant Oncologist in Renal cell carcinoma and Melanoma	Chief investigator in renal cancers at the leading Francis Crick Centre and medical oncologist at the Royal Marsden. In research, has a specialist interest in cancer evolution, translational studies and VHL disease. Is a widely recognized clinical expert, highly	Has provided paid consultancy services to MSD on VHL.

		regarded by peers, presenting at global congresses and [redacted]. Has been widely published and has worked with leading VHL experts in some of her publications on the disease. [redacted] Has a deep understanding of the disease at both a biological and patient level.	

Description of validation points

There were 6 key validation points identified in the EAG report for further consultation with clinical experts. These are described below.

Key Validation Point 1 – Interpretation of the label population

The EAG note the interpretation of the patient population described by the company as those who have exhausted all alternative treatments and for whom localised procedures would be a 'last-resort' option and likely result in loss of organ function with extremely poor outcomes. We wanted to provide further confirmation and documentation of this with this clinical expert elicitation exercise.

Key Validation Point 2 – Immediate surgery assumption in the SoC arm

The EAG raised concerns around defining those who require therapy as receiving immediate surgery in the comparator arm. The EAG stated whether in daily practice these patients would be able to wait almost one year (in median) without surgery until a response to treatment is observed. We believe this timeframe is based on the median time-to-response of belzutifan which in the 01-APR-2022 data-cut was 11.1 months. Response defined in the MK-6482-004 study is at least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters. We know that shrinkage in the tumour would begin before the median time-to-response is achieved. Hence, we adjusted the timeframe mentioned by the EAG and asked about a 6-month wait time before surgery.

Key Validation Point 3 – Outcomes for belzutifan non-responders

The EAG report noted the 36.1% of the trial population who did not respond to belzutifan treatment when discussing the immediate surgery assumption. The EAG noted that this concern should be clarified with clinical experts hence we asked clinical experts what would happen to belzutifan non-responders in practice.

Key Validation Point 4 – Perioperative mortality & surgical complications risks

The EAG report noted better justification of the perioperative mortality and surgical complication risk adjustments made to better align with the label population. We further consulted with clinical experts on the risks posed for the target population.

Key Validation Point 5 – Face validity of utility values

In the Clarification Letter provided by the EAG, the face validity of the utility values was questioned. We consulted with clinical experts again on the topic in order to provide further documentation.

Key Validation Point 6 – Use of belzutifan in VHL patient with metastatic disease

The EAG report noted that there is no contraindication in terms of metastatic or advanced disease in the marketing authorisation details. We further clarified with clinical experts the potential use of belzutifan in patients with VHL-associated metastatic disease.

Results from clinical expert elicitation & discussion

The feedback from clinical experts and further discussion in relation to assumptions made in this appraisal are detailed on each of the validation points below.

Key Validation Point 1 – Interpretation of the label population

Clinical experts were clear that the belzutifan-eligible population are those for whom surgical interventions would result in loss of organ function or functional neurological impairment. They define these patients as those who who, due to position or size of tumour, would inevitably result in significant consequences from surgery that would have a detrimental impact on their quality of life. One expert described these patients are ‘standing on the precipice of organ failure’ and that a surgical procedure would be ‘severely hazardous’. This interpretation from clinicians aligns with the position in the company submission and with the outcomes and complications modelled following surgery in both the belzutifan and SoC arm.

Key Validation Point 2 – Immediate surgery assumption in the SoC arm

Technical engagement response form

Clinical experts understood the target population as those “who require therapy” and in the absence of belzutifan, SoC would be surgery. When questioned on whether this surgery would take place immediately or if there could feasibly be a waiting time, all experts were strongly against the idea of suggesting a delay in surgery. They stated that the gold-standard is for a patient who is requiring surgery to have this immediately and any suggestion of waiting times would be not only arbitrary but need to take into consideration variations across the country. One expert suggested that for a patient whose tumour was nearing the threshold for intervention they may wait until the next scan before deciding to operate, but then clarified that this patient would not be considered as requiring therapy. We further clarified the critique raised by the EAG, specifically the concern that belzutifan could take time to work and therefore the suggestion that the two populations in each arm of the model may not be equal. Experts still stood firm on their response that if belzutifan were available then it would be used, but in its absence, surgery is the only option, and it should take place immediately for a patient who requires it. Two experts asked if this is an issue because the MK-6482-004 trial excludes patients who had immediate surgery to which we clarified that this was a point of critique. They then highlighted that clinical trials by nature need to be rigid on inclusion and exclusion criteria and this should not preclude the target population from belzutifan treatment. Following these discussions, the decision was made not to introduce a delay in surgery in the SoC arm. As stated above, all clinical experts were strongly against the idea and resonated with the initial assumption that patients would receive surgery immediately.

Key Validation Point 3 – Outcomes for belzutifan non-responders

When asked about what would happen to patients who did not respond to belzutifan, clinical experts stated that they would consider the classification of stable disease as a meaningful improvement in a previously growing tumour. One expert said ‘the aim is to prevent or at least delay surgery, so if a tumour is no longer growing and considered stable disease this is a positive’. Another

expert affirmed this notion. Experts noted that patients who progress on belzutifan would require a nuanced conversation on the risk vs. benefit of future treatment and could not draw definitive conclusions on what would happen to these patients. Hence, we have made no change to the modelling of belzutifan non-responders. As per clinical expert opinion, where stable disease is considered a positive, there is no requirement to have an assumption that all of these patients have surgery as implied by the EAG. Furthermore, the pre-surgery → surgery transition is derived from time-to-surgery from the MK-6482-004 trial. Therefore, non-response resulting in surgery is accounted for based on the 7 RCC surgery events and the 2 CNS Hb surgery events that occurred in the trial. Note that there were only 3 patients identified as progressive disease and only for CNS Hb, hence, the 7 RCC surgeries that occurred were in patients that had stable disease.

Key Validation Point 4 – Perioperative mortality & surgical complications risks

Clinical experts found it challenging to quantify the exact likelihood of complications a belzutifan-eligible patient would require. They were clear that they do have an elevated level of risk due to the severity of disease in the label population, when compared to those in the Optum study, but were unable to quantify how much more elevated this should be. When presented with the adjustments made in the company base-case, they commented that they appear to be reasonable. All experts were clear, however, on identifying the inevitable consequences of surgery resulting from loss of organ function or neurological impairment. We have held our initial assumption on the adjustments made to perioperative mortality & surgical complications risks and explored these in sensitivity analyses presented in Table 5.

Key Validation Point 5 – Face validity of utility values

As with Key Validation Point 4, clinical experts found it challenging to comment on the exact utility values presented. However, they were clear that in the absence of any alternative data sources, the VHL RW QoL Disease Burden Study would provide the best available evidence as it would be closest to the population of interest. When explaining where the remaining utility values were sourced, they agreed that it seemed plausible and relevant proxies in the absence of alternative data sources. However, they pointed to the heterogenous nature of disease and that no 2 VHL patients have the same experience. As no alternative utility values or sources were provided as part of this exercise, those reported in the company submission remain the base case.

Key Validation Point 6 – Use of belzutifan in VHL patient with metastatic disease

Clinical experts suggested that based on the mechanism of action of belzutifan, they do not see a reason why it would not be effective in metastatic disease. However, as it has not been studied in this population and is not explicitly stated in the label, they would not prescribe it for this purpose. One expert specifically stated 'belzutifan's aim in the pathway is to delay structural interventions as long as possible, if a patient presented with metastatic disease in VHL then we would have passed the point at which belzutifan could be used'. This is consistent with what is included in the company submission and economic modelling and affirms that metastatic disease therapies are not appropriate comparators but rather subsequent treatments.

Strengths & Limitations

There are several strengths of this expert elicitation exercise:

- Experts' judgements were elicited individually, minimising the risk of bias
- A limited number of questions were presented to minimise response fatigue

- A consistent approach was applied across experts, questions were not changed following each discussion
- Experts clearly understood the questions asked and were able to provide relevant and appropriate answers

Some limitations are:

- Due to time constraints, it was not possible to conduct a group expert elicitation exercise where a consensus easily be reached
- Questions that presented values were harder to validate due to the heterogeneity of disease and limited number of patients clinicians would see in practice
- All the concerns raised by the EAG for further discussion with clinical experts were unable to be explored

Slides presented to facilitate discussion

Confidential

Defining the label population

Belzutifan is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and ***for whom localised procedures are unsuitable or undesirable****.*

Based on the indication wording above, is the label population appropriately interpreted as patients who have exhausted all alternative options for the tumour(s) belzutifan intends to treat and are at the "end of the road" given that *localised procedures are unsuitable or undesirable*?

MSD 4

NICE National Institute for Health and Care Excellence

Confidential

Defining current standard of care

*Belzutifan is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease **who require therapy** for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable.*

As patients *require therapy*, in the absence of belzutifan we have taken this therapy to be surgery. Initially, we modelled surgery taking place immediately once the decision is made to treat.

Following EAG critique on this assumption, we are considering that patients could wait up to 6 months before receiving surgery due to NHS waiting times & scheduling challenges, is this a reasonable assumption?

MSD 5

NICE National Institute for Health and Care Excellence

Confidential

Belzutifan non-responders

In the latest data-cut from the clinical trial, there are 36.1% of patients without a response to belzutifan.

What do you expect to happen in practice to patients who do not respond to belzutifan?

MSD

Confidential

Perioperative mortality and surgical complication risk (1)

Considering the label population are unsuitable/undesirable for localised procedures, would it be reasonable to assume a risk adjustment (doubled) for perioperative mortality and surgical complication risk?

VHL-RCC surgery complication	Risk pre-adjustment *(Source: Optum database)*	Risk post-adjustment for surgery-unsuitable population
Perioperative mortality	1.96%	3.92%
End stage renal disease and/or dialysis	4.0%	80.0%*
Chronic kidney disease	24.0%	20.0%*
Hernia surgery	1.6%	3.2%
Chronic pain	8.8%	17.6%
Cerebral vasculature occlusion or stroke	3.2%	6.4%

*Metabolic consequence of last-resort surgery leading to loss of organ function

MSD 7

Confidential

Perioperative mortality and surgical complication risk (2)

Considering the label population are unsuitable/undesirable for localised procedures, would it be reasonable to assume a risk adjustment (doubled) for perioperative mortality and surgical complication risk?

VHL-CNS Hb surgery complication	Risk pre-adjustment *(Source: Optum database)*	Risk post-adjustment for surgery-unsuitable population
Perioperative mortality	1.82%	3.64%
Chronic pain (in CNS Hb population)	15.4%	30.8%
Cerebral vasculature occlusion or stroke	7.7%	85.0%*
Seizures	10.3%	20.5%
Neurological complications	43.6%	87.2%
VHL-pNET surgery complication	**Risk pre-adjustment** *(Source: Optum database)*	**Risk post-adjustment for surgery-unsuitable population**
Perioperative mortality	1.71%	3.42%
Chronic pain (in pNET population)	10.0%	20.0%
Secondary diabetes or exocrine pancreatic insufficiency	20.0%	100.0%*
Immunocompromisation	0.0%	100.0%*

*Metabolic consequence of last-resort surgery leading to loss of organ function

MSD 8

Confidential

Quality of life for the label population

Considering the interpretation of the label population, on a scale of 0 (death) to 1 (perfect health), do these utility values reflect what we would expect for the label population?

Health state/response status	Utility value	Source
Non-metastatic health states*		
Complete response	**0.868**	KEYNOTE-564 (patients who are disease-free post-nephrectomy in adjuvant RCC trial)
Partial response	**0.754**	VHL RW QoL Disease Burden Study (MSD)
Stable disease	**0.754**	VHL RW QoL Disease Burden Study (MSD)
Progressive disease (RCC & pNET)	**0.665**	VHL RW QoL Disease Burden Study (MSD)
Progressive disease (CNS Hb)	**0.550**	Kiebert et al. (2001) study on patients with motor neuron disease
Metastatic disease state		
Metastatic disease (pre-progression)	**0.525**	VHL RW QoL Disease Burden Study (MSD)
Metastatic disease (post-progression)	**0.412**	VHL RW QoL Disease Burden Study (MSD)

*Note disutilities associated with surgical complications are applied separately

MSD 9

NICE National Institute for Health and Care Excellence

Confidential

Belzutifan use in metastatic disease

Belzutifan is indicated for the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable.

Based on the indication of belzutifan, can we confirm that belzutifan would not be used for a VHL patient with metastatic disease?

MSD 10

Appendix 3

Additional data on the studies included prior to the final step of the clinical effectiveness evidence systematic literature review

Bibliographic details of the 26 records that were initially included in the clinical effectiveness systematic literature review:

Table 8 List of citations initially included at full-text screening phase (reproduced here from Appendix D Table 110 of the original company submission)

Citation
Asthagiri AR, Mehta GU, Zach L, et al. Prospective evaluation of radiosurgery for hemangioblastomas in von Hippel-Lindau disease. Neuro-Oncology. 2010;12(1):80-86.
Capitanio U, Rosiello G, Erdem S, et al. Clinical, surgical, pathological and follow-up features of kidney cancer patients with Von Hippel-Lindau syndrome: novel insights from a large consortium. World J Urol. 2021;39(8):2969-2975.
Chan VW, Lenton J, Smith J, et al. Multimodal image-guided ablation on management of renal cancer in Von-Hippel-Lindau syndrome patients from 2004 to 2021 at a specialist centre: A longitudinal observational study. Eur J Surg Oncol. 2022;48(3):672-679.
Chang SD, Meisel JA, Hancock SL, et al. Treatment of hemangioblastomas in von Hippel-Lindau disease with linear accelerator-based radiosurgery. Neurosurgery. 1998;43(1):28-34; discussion 34-25.
Cvek J, Knybel L, Reguli S, et al. Stereotactic radiotherapy for spinal hemangioblastoma - disease control and volume analysis in long-term follow up. Rep Pract Oncol Radiother. 2022;27(1):134-141.
Eggener SE, Rubenstein JN, Smith ND, et al. Renal tumors in young adults. J Urol. 2004;171(1):106-110.
Frydenberg M, Malek RS, Zincke H. Conservative renal surgery for renal cell carcinoma in von Hippel-Lindau's disease. J Urol. 1993;149(3):461-464.
Goldfarb DA, Neumann HP, Penn I, et al. Results of renal transplantation in patients with renal cell carcinoma and von Hippel-Lindau disease. Transplantation. 1997;64(12):1726-1729.
Jilg CA, Neumann HP, Glasker S, et al. Nephron sparing surgery in von Hippel-Lindau associated renal cell carcinoma; clinicopathological long-term follow-up. Fam Cancer. 2012;11(3):387-394.

Citation
Jonasch E, Donskov F, Iliopoulos O, et al. Belzutifan for Renal Cell Carcinoma in von Hippel–Lindau Disease. New England Journal of Medicine. 2021;385(22):2036-2046.
Jonasch E, McCutcheon IE, Waguespack SG, et al. Pilot trial of sunitinib therapy in patients with von Hippel-Lindau disease. Ann Oncol. 2011;22(12):2661-2666.
Jonasch E, McCutcheon, Gombos DS, et al. Pazopanib in patients with von Hippel-Lindau disease: a single-arm, single-centre, phase 2 trial. The Lancet Oncology. 2018;19(10):1351-1359.
Kano H, Niranjan A, Mongia S, et al. The role of stereotactic radiosurgery for intracranial hemangioblastomas. Neurosurgery. 2008;63(3):443-450; discussion 450-441.
Kano H, Shuto T, Iwai Y, et al. Stereotactic radiosurgery for intracranial hemangioblastomas: a retrospective international outcome study. J Neurosurg. 2015;122(6):1469-1478.
Koh ES, Nichol A, Millar BA, et al. Role of fractionated external beam radiotherapy in hemangioblastoma of the central nervous system. Int J Radiat Oncol Biol Phys. 2007;69(5):1521-1526.
Ma K, Hong B, Zhou J, et al. The Efficacy and Safety of Tyrosine Kinase Inhibitors for Von Hippel-Lindau Disease: A Retrospective Study of 32 Patients. Front Oncol. 2019;9:1122.
Morgan WR, Zincke H. Progression and survival after renal-conserving surgery for renal cell carcinoma: experience in 104 patients and extended followup. J Urol. 1990;144(4):852-857; discussion 857-858.
Persad RA, Probert JL, Sharma SD, et al. Surgical management of the renal manifestations of von Hippel-Lindau disease: a review of a United Kingdom case series. Br J Urol. 1997;80(3):392-396.
Ploussard G, Droupy S, Ferlicot S, et al. Local recurrence after nephron-sparing surgery in von Hippel-Lindau disease. Urology. 2007;70(3):435-439.
Roma A, Maruzzo M, Basso U, et al. First-Line sunitinib in patients with renal cell carcinoma (RCC) in von Hippel-Lindau (VHL) disease: clinical outcome and patterns of radiological response. Fam Cancer. 2015;14(2):309-316.
Roupret M, Hopirtean V, Mejean A, et al. Nephron sparing surgery for renal cell carcinoma and von Hippel-Lindau's disease: a single center experience. J Urol. 2003;170(5):1752-1755.
Simone CB, 2nd, Lonser RR, Ondos J, et al. Infratentorial craniospinal irradiation for von Hippel-Lindau: a retrospective study supporting a new treatment for patients with CNS hemangioblastomas. Neuro Oncol. 2011;13(9):1030-1036.
Steinbach F, Novick AC, Zinke H, et al. Treatment of renal cell carcinoma in von Hippel-Lindau disease: A multicenter study. J Urol. 1995;153(6):1812-1816.
Wessendorf J, König A, Heers H, et al. Repeat Percutaneous Radiofrequency Ablation of T1 Renal Cell Carcinomas is Safe in Patients with Von Hippel-Lindau Disease. Cardiovasc Intervent Radiol. 2021;44(12):2022-2025.

Citation
Yao M, Yoshida M, Kishida T, et al. VHL tumor suppressor gene alterations associated with good prognosis in sporadic clear-cell renal carcinoma. Journal of the National Cancer Institute. 2002;94(20):1569-1575.
Yousef A, Rutkowski MJ, Yalcin CE, et al. Sporadic and Von-Hippel Lindau disease-associated spinal hemangioblastomas: institutional experience on their similarities and differences. J Neurooncol. 2019;143(3):547-552.

Detailed information on the 26 records that were initially included in the clinical effectiveness systematic literature review:

Table 9 Study and treatment characteristics

Study	Tx	Tx schedule	N	F/U duration (median, months)	Location	Study Design	Date Initiated	Date Completed
RCC								
Non-randomized clinical trials								
Jonasch et al 2011	Sunitinib	50 mg	12	--	United States	Single-arm, Phase 2, Open-label	June 19, 2006	July 19, 2007
Jonasch et al 2018	Pazopanib	800 mg	32	12 (7-32)	United States	Single-arm, Phase 1, Open-label	January 17, 2012	April 11, 2023
Jonasch et al 2021	Belzutifan	120 mg	61	21.8 (20.2-30.1)	Multinational	Single-arm, Phase 2, Open-label	May 2, 2018	March 29, 2026
Observational studies								
Capitanio et al 2021	Surgery	--	96	96	Europe	Retrospective Study	1987	2001
Chan et al 2022	Image-guided ablation (Radiofrequency ablation, Cryoablation, Irreversible electroporation)	--	17	79 (51-134)	United Kingdom	Retrospective Study	2004	2021
Eggener et al 2004	Nephrectomy	--	12	--	USA	Retrospective Study	January 1998	October 2002
Frydenberg et al 1993	Nephrectomy	--	19	60*	USA	Prospective Study	1956	1991
Goldfarb et al 1997	Renal transplant	--	32	48 (35)*	North America, Europe	Retrospective Study	1974	1996

Jilg et al 2012	NSS, Nephrectomy, RFA	--	54	67 (2.6-205.7)	Germany	Retrospective Study	1991	February 2011
Ma et al 2019	Sunitinib 50 mg; Sorafenib 400 mg; Axitinib 5 mg; Pazopanib 800 mg	--	31	31.5	China	Retrospective Study	2009	2018
Morgan et al 1990	Nephrectomy and/or enucleation	--	6	40*	USA	Prospective Study	November 1965	December 1987
Persad et al 1997	Nephrectomy	--	11	26	UK	Prospective Study	1983	1993
Ploussard et al 2007	NSS	--	18	100 (7-223)	France	Retrospective Study	February 1987	August 2005
Roma et al 2015	Sunitinib	--	14	39.4	Italy	Retrospective Study	2007	2012
Roupret et all 2003	Nephron sparing surgery	--	56	55.9	France	Retrospective Study	1988	2001
Steinbach et al 1995	NSS	--	65	68*	USA	Retrospective Study	--	1993
Wessendorf et al 2021	Radiofrequency ablation	--	9	--	Germany	Retrospective Study	--	--
Yao et al 2002	Nephrectomy with adjuvant postoperative interferon and/or chemotherapy	--	78	26	Japan	Prospective Study	October 1986	December 1995
CNS Hemangioblastoma								
Non-randomized clinical trials								
Jonasch et al 2011	Sunitinib	50 mg	11	--	United States	Single-arm, Phase 2, Open-label	June 19, 2006	July 19, 2007
Jonasch et al 2018	Pazopanib	800 mg	23	12 (7-32)	United States	Single-arm, Phase 1, Open-label	January 17, 2012	April 11, 2023
Jonasch et al 2021	Belzutifan	120 mg	50	21.8 (20.2-30.1)	Multinational	Single-arm, Phase 2, Open-label	May 2, 2018	March 29, 2026
Observational studies								
Asthagiri et al 2010	Stereotactic radiosurgery	--	20	8.5 (3-17.6)*^	North America	Prospective study	--	--
Chang et al 1998	Linear Accelerator-based Radiosurgery	--	13	--	United States	Retrospective Study	1989	1996
Cvek et al 2022	Stereotactic body radiotherapy	--	5	5 (2-8)^	--	Retrospective Study	2010	2018
Kano et al 2008	Stereotactic radiosurgery	--	13	50.1 (6-165)	United States	Retrospective Study	June 1990	October 2006
Kano et al 2015	Stereotactic radiosurgery	--	80	5 (0.5-18)*^	North America and Japan	Retrospective Study	1990	2010

Koh et al 2007	Fractionated external beam radiotherapy	--	5	--	Canada	Retrospective Study	January 1, 1980	December 31, 2004
Ma et al 2019	Sunitinib 50 mg; Sorafenib 400 mg; Axitinib 5 mg; Pazopanib 800 mg	--	22	31.5	China	Retrospective Study	2009	2018
Simone et al 2011	Infratentorial craniospinal radiation therapy	--	7	73.8 (40.3-155.6)*	United States	Retrospective Study	1998	2008
Yousef et al 2019	Surgery	--	20	---	United states	Retrospective Study	1997	2016
pNET								
Non-randomized clinical trials								
Jonasch et al 2011	Sunitinib	50 mg	7	--	United States	Single-arm, Phase 2, Open-label	June 19, 2006	July 19, 2007
Jonasch et al 2018	Pazopanib	800 mg	9	12 (7-32)	United States	Single-arm, Phase 1, Open-label	January 17, 2012	April 11, 2023
Jonasch et al 2021	Belzutifan	120 mg	22	21.8 (20.2-30.1)	Multinational	Single-arm, Phase 2, Open-label	May 2, 2018	March 29, 2026

**Mean reported*
^years reported
Abbreviations: CNS, central nervous system; RCC, renal cell carcinoma

Table 10 Baseline demographics and performance scores

Study	Tx	N	Age (mean, range) y	Male, n (%)	ECOG 0, n (%)	Clear cell histology, n (%)	CNS hemangioblastoma, n (%)	Method of VHL status diagnosis	VHL Type
RCC									
Non-randomized clinical trials									
Jonasch et al 2011	Sunitinib	15	36 (22, 57)*	--	--	12 (80)	--	--	--
Jonasch et al 2018	Pazopanib	32	38 (32-42)*	14 (44)	--	--	--	--	--
Jonasch et al 2021	Belzutifan	61	41 (19-66)*	32 (52)	50 (82)	--	--	--	1 (n=51); 2A (n=2); 2B (n=6); missing (n=2)

Study	Tx	N	Age (mean, range) y	Male, n (%)	ECOG 0, n (%)	Clear cell histology, n (%)	CNS hemangioblastoma, n (%)	Method of VHL status diagnosis	VHL Type
Observational									
Capitanio et al 2021	Surgery	96	38 (32-47)*	51 (53.1)	--	96 (100)	--	--	--
Chan et al 2022	Image-guided ablation (Radiofrequency ablation, Cryoablation, Irreversible electroporation)	17	43.9 (13.6)	65 (6)	--	17 (100)	--	--	--
Eggener et al 2004	Nephrectomy	114	--	0 (0)	--	--	--	Strong family history of RCC or a personal history of other tumors distinctive for VHL and some were also symptomatic with flank pain or hematuria	--
Frydenberg et al 1993	Nephrectomy	19	40.3 (15, 65)	8 (73)	--	--	--	--	--
Goldfarb et al 1997	Renal transplant	32	36 (19-59)	23	--	--	--	--	--
Jilg et al 2012	NSS, Nephrectomy, RFA	54	38.5 (18, 73)^	24 (44.4)	--	54 (100)	--	--	--
Ma et al 2019	Sunitinib 50 mg; Sorafenib 400 mg; Axitinib 5 mg; Pazopanib 800 mg	32	41.5 (21-66)	18 (56)	--	--	--	--	--
Morgan et al 1990	Nephrectomy and/or enucleation	104	--	--	--	--	--	--	--
Persad et al 1997	Nephrectomy	11	42 (31, 62)	7 (64)	--	--	--	--	--
Ploussard et al 2007	NSS	21	38.5 (24, 69)	--	--	(98.8)	--	Family history, VHL germline mutation. Major manifestation (clinical) of VHL with a familial history	--

Study	Tx	N	Age (mean, range) y	Male, n (%)	ECOG 0, n (%)	Clear cell histology, n (%)	CNS hemangioblastoma, n (%)	Method of VHL status diagnosis	VHL Type
Roma et al 2015	Sunitinib 50 mg	14	48 (27-71)*	6 (43)	--	14 (100)	--	--	--
Roupret et all 2003	Nephron sparing surgery	56	37.2**	26 (46.4)	--	--	--	--	--
Steinbach et al 1995	NSS	65	39 (15, 67)**	39 (60)	--	(89)	--	By the presence of RCC and a family history of VHL or by the presence of RCC and other major manifestation of VHL	--
Wessendorf et al 2021	Radiofreqency ablation	9	--	--	--	--	--	--	--
Yao et al 2002	Nephrectomy with adjuvant postoperative interferon and/or chemotherapy	187	--	--	--	--	--	--	--
CNS Hemangioblastoma									
Non-randomized clinical trials									
Jonasch et al 2021	Belzutifan	61	41 (19-66)*	32 (52)	50 (82)	--	50(82)	--	1 (n=51); 2A (n=2); 2B (n=6); missing (n=2)
Jonasch et al 2018	Pazopanib	32	38 (32-42)*	14 (44)	--	--	23(72)	--	--
Jonasch et al 2011	Sunitinib	15	36 (22, 57)*	--	--	12 (80)	11(73)	--	--
Observational									
Asthagiri et al 2010	Stereotactic radiosurgery	20	37.5 (13-67)	10 (50)	--	--	20 (100)	Patients had VHL diagnosed by clinical and genetic criteria.	--
Chang et al 1998	Linear Accelerator-based Radiosurgery	13	40 (31-57)*	10 (76.9)	--	--	--	--	--
Cvek et al 2022	Stereotactic body radiotherapy	5	22 (18-60)	3 (60)	--	--	5 (100)	--	--

Study	Tx	N	Age (mean, range) y	Male, n (%)	ECOG 0, n (%)	Clear cell histology, n (%)	CNS hemangioblastoma, n (%)	Method of VHL status diagnosis	VHL Type
Kano et al 2008	Stereotactic radiosurgery	13	40.2	7 (53.8)	--	--	13 (100)	--	--
Kano et al 2015	Stereotactic radiosurgery	80	38	36 (45)	--	--	80 (100)	The tumors were diagnosed by MRI and angiography. Ten VHL patients had 2 or more hemangioblastomas at separate locations within the brain and a family history of VHL	--
Koh et al 2007	Fractionated external beam radiotherapy	18	31 (25-41)	--	--	--	5 (100)	Three VHL patients had a positive family history of the disease in first degree relatives. Major manifestation (clinical) of VHL with a familial history	--
Ma et al 2019	Sunitinib 50 mg; Sorafenib 400 mg; Axitinib 5 mg; Pazopanib 800 mg	32	41.5 (21-66)	18 (56)	--	--	22(69)	--	--
Simone et al 2011	Infratentorial craniospinal radiation therapy	7	40.8*	4 (57.1)	--	--	7 (100)	--	--
Yousef et al 2019	Surgery	20	---	15 (75)	--	--	20 (100)	--	--

**Median age was reported; ** age at diagnosis; ^Calculated (Mean age was 37 years in males (range 18–66 years) and 40 years in females (range 21–73 years)*
Abbreviations: CNS, central nervous system; ECOG, Eastern Cooperative Oncology Group; N, number of patients; NSS, nephron sparing surgery; RCC, renal cell carcinoma; RFA, radiofrequency ablation; Tx, treatment; VHL, Von Hippel-Landau; y, years

Table 11 Response outcomes for included studies

Study	Tx	N	Number of lesions	Response Criteria	ORR, n (%)	CR, n (%)	PR, n (%)	SD, n (%)	PD, n (%)	DOR, n	Median DOR, months
RCC											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	12	18	mRECIST	6(33)	--	6 (33)	10 (67)	2 (10)	--	--
Jonasch et al 2018	Pazopanib	31	--	RECIST v1.1	13 (42)	0	13 (42)	18 (58)	0	--	--
Jonasch et al 2021	Belzutifan	61	--	RECIST v1.1	30 (49)	0	30 (49)	30 (49)	0	61	NR (2.8-22.3)
Observational studies											
Ma et al 2019	Sunitinib 50 mg	12	15	RECIST v1.1	6 (40)	0	6 (40)	5 (33)	4 (27)	--	--
Ma et al 2019	Sorafenib 400 mg	11	12	RECIST v1.1	3 (25)	0	3 (25)	5 (42)	4 (33)	--	--
Ma et al 2019	Axitinib 5 mg	6	6	RECIST v1.1	2 (33)	0	2 (33)	4 (67)	0	--	--
Ma et al 2019	Pazopanib 800 mg	3	3	RECIST v1.1	0	0	0	3 (100)	0	--	--
Roma et al 2015	Sunitinib 50 mg	14	--	RECIST v1.1	9 (64.3)	0	9 (64.3)	5 (35.7)	6	--	--
CNS Hemangioblastoma											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	11	21	mRECIST	0	--	--	19(91)	2 (9)	--	--
Jonasch et al 2021	Belzutifan	50	--	RECIST v1.1	15(30)	3(6)	12 (24)	31 (62)	2(4)	--	NR (2.8-22.3)
Jonasch et al 2018	Pazopanib	--	49	RECIST v1.1	2(4)	0	2(4)	47(96)	0	--	--
Observational studies											
Ma et al 2019	Stereotactic radiosurgery	20	--	--	--	--	--	--	--	--	--

Study	Tx	N	Number of lesions	Response Criteria	ORR, n (%)	CR, n (%)	PR, n (%)	SD, n (%)	PD, n (%)	DOR, n	Median DOR, months
Ma et al 2019	Sunitinib 50 mg	12	2	RECIST v1.1	0	0	0	2(100)	0	--	--
Ma et al 2019	Sorafenib 400 mg	11	1	RECIST v1.1	0	0	0	1(100)	0	--	--
Ma et al 2019	Axitinib 5 mg	6	2	RECIST v1.1	1 (50)	0	1(50)	1(50)	0	--	--
Roma et al 2015	Sunitinib 50 mg	11	--	RECIST v1.1	0	0	0	11(100)	--	--	--
pNET											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	--	5	mRECIST	0	--	--	5(100)	0	--	--
Jonasch et al 2021	Belzutifan	22	--	RECIST v1.1	20(91)	3(14)	17(77)	2 (9)	0	--	NR (2.9-22.3)
Jonasch et al 2018	Pazopanib	--	17	RECIST v1.1	9(53)	0	9(53)	8(47)	0	--	--

**Median age was reported; ** age at diagnosis; ^Calculated (Mean age was 37 years in males (range 18–66 years) and 40 years in females (range 21–73 years)*
Abbreviations: CNS, central nervous system; ECOG, Eastern Cooperative Oncology Group; N, number of patients; NSS, nephron sparing surgery; RCC, renal cell carcinoma; RFA, radiofrequency ablation; Tx, treatment; VHL, Von Hippel-Landau; y, years

Study	Tx	N	Number of lesions	Response Criteria	ORR, n (%)	CR, n (%)	PR, n (%)	SD, n (%)	PD, n (%)	DOR, n	Median DOR, months
RCC											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	12	18	mRECIST	6(33)	--	6 (33)	10 (67)	2 (10)	--	--
Jonasch et al 2018	Pazopanib	31	--	RECIST v1.1	13 (42)	0	13 (42)	18 (58)	0	--	--
Jonasch et al 2021	Belzutifan	61	--	RECIST v1.1	30 (49)	0	30 (49)	30 (49)	0	61	NR (2.8-22.3)
Observational studies											

Study	Tx	N	Number of lesions	Response Criteria	ORR, n (%)	CR, n (%)	PR, n (%)	SD, n (%)	PD, n (%)	DOR, n	Median DOR, months
Ma et al 2019	Sunitinib 50 mg	12	15	RECIST v1.1	6 (40)	0	6 (40)	5 (33)	4 (27)	--	--
Ma et al 2019	Sorafenib 400 mg	11	12	RECIST v1.1	3 (25)	0	3 (25)	5 (42)	4 (33)	--	--
Ma et al 2019	Axitinib 5 mg	6	6	RECIST v1.1	2 (33)	0	2 (33)	4 (67)	0	--	--
Ma et al 2019	Pazopanib 800 mg	3	3	RECIST v1.1	0	0	0	3 (100)	0	--	--
Roma et al 2015	Sunitinib 50 mg	14	--	RECIST v1.1	9 (64.3)	0	9 (64.3)	5 (35.7)	6	--	--
CNS Hemangioblastoma											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	11	21	mRECIST	0	--	--	19(91)	2 (9)	--	--
Jonasch et al 2021	Belzutifan	50	--	RECIST v1.1	15(30)	3(6)	12 (24)	31 (62)	2(4)	--	NR (2.8-22.3)
Jonasch et al 2018	Pazopanib	--	49	RECIST v1.1	2(4)	0	2(4)	47(96)	0	--	--
Observational studies											
Ma et al 2019	Stereotactic radiosurgery	20	--	--	--	--	--	--	--	--	--
Ma et al 2019	Sunitinib 50 mg	12	2	RECIST v1.1	0	0	0	2(100)	0	--	--
Ma et al 2019	Sorafenib 400 mg	11	1	RECIST v1.1	0	0	0	1(100)	0	--	--
Ma et al 2019	Axitinib 5 mg	6	2	RECIST v1.1	1 (50)	0	1(50)	1(50)	0	--	--
Roma et al 2015	Sunitinib 50 mg	11	--	RECIST v1.1	0	0	0	11(100)	--	--	--
pNET											
Non-randomized trials											
Jonasch et al 2011	Sunitinib	--	5	mRECIST	0	--	--	5(100)	0	--	--
Jonasch et al 2021	Belzutifan	22	--	RECIST v1.1	20(91)	3(14)	17(77)	2 (9)	0	--	NR (2.9-22.3)

Study	Tx	N	Number of lesions	Response Criteria	ORR, n (%)	CR, n (%)	PR, n (%)	SD, n (%)	PD, n (%)	DOR, n	Median DOR, months
Jonasch et al 2018	Pazopanib	--	17	RECIST v1.1	9(53)	0	9(53)	8(47)	0	--	--

Abbreviations: CNS, central nervous system; CR, Complete response; DOR, Duration of response; mRECIST, modified RECIST; NSS, nephron sparing surgery; ORR, Overall response rate; PD, Progressive disease; pNET, pancreatic neuro-endocrine tumor; PR, Partial response; RECIST, Response Evaluation Criteria in Solid Tumors; RFA, radiofrequency ablation; SD, Stable disease; Tx: Treatment

Table 12 Adverse events

Study	Tx	N	AE	Safety N	Overall TRAE, n (%)	Grade 3 AE, n (%)	Grade 3-5 AE, n (%)	SAE, n (%)	Disc. due to AE/toxicity, n (%)	Disc. due to progression, n (%)
RCC										
Non-randomized clinical trials										
Jonasch et al 2011	Sunitinib	54	--	--	--	--	--	--	--	--
Jonasch et al 2021	Pazopanib 800 mg	32	--	--	--	--	--	2	--	--
Jonasch et al 2018	Belzutifan	61	--	61	(96.7)	NR (9.8)	--	--	2 (3)*	1 (2)
Observational Studies										
Capitanio et al 2021	Surgery	96	--	--	--	--	--	--	--	--
Chan et al 2022	Image-guided ablation (Radiofrequency ablation, Cryoablation, Irreversible electroporation)	17	--	10	58.8	1 (5.9)	--	--	--	--
Eggener et al 2004	Nephrectomy	12	--	--	--	--	--	--	--	--
Frydenberg et al 1993	Nephrectomy	19	--	--	--	--	--	--	--	--
Goldfarb et al 1997	Renal transplant	32	--	--	--	--	--	--	--	--
Jilg et al 2012	NSS, Nephrectomy, RFA	54	--	--	--	--	--	--	--	--
Ma et al 2019	Sunitinib 50 mg; Sorafenib 400 mg;	32	--	--	--	--	--	--	--	--

Study	Tx	N	AE	Safety N	Overall TRAE, n (%)	Grade 3 AE, n (%)	Grade 3-5 AE, n (%)	SAE, n (%)	Disc. due to AE/toxicity, n (%)	Disc. due to progression, n (%)
	Axitinib 5 mg; Pazopanib 800 mg									
Morgan et al 1990	Nephrectomy and/or enucleation	6	--	--	--	--	--	--	--	--
Persad et al 1997	Nephrectomy	11	--	--	--	--	--	--	--	--
Ploussard et al 2007	NSS	18	--	--	--	--	--	--	--	--
Roma et al 2015	Sunitinib 50 mg	14	--	--	--	--	--	--	--	--
Roupret et all 2003	Nephron sparing surgery	56	--	--	--	--	--	--	--	--
Steinbach et al 1995	NSS	65	--	--	--	--	--	--	--	--
Wessendorf et al 2021	Radiofreqency ablation	9	--	9	55.6	--	--	--	--	--
Yao et al 2002	Nephrectomy with adjuvant postoperative interferon and/or chemotherapy	78	--	78	--	--	--	--	--	--
CNS Hemangioblastoma										
Asthagiri et al 2010	Stereotactic radiosurgery	20	--	20	4 (20)	--	--	--	--	--
Chang et al 1998	Linear Accelerator-based Radiosurgery	13	--	--	--	--	--	--	--	--
Cvek et al 2022	Stereotactic body radiotherapy	5	--	--	--	--	--	--	--	--
Kano et al 2008	Stereotactic radiosurgery	13	--	--	--	--	--	--	--	--
Kano et al 2015	Stereotactic radiosurgery	80	--	--	--	--	--	--	--	--
Koh et al 2007	Fractionated external beam radiotherapy	5	--	--	--	--	--	--	--	--
Simone et al 2011	Infratentorial craniospinal radiation therapy	7	--	3	42.9	--	--	--	--	--
Yousef et al 2019	Surgery	20	--	--	--	--	--	--	--	--

**Calculated, Abbreviations: AE: Adverse event; Disc.: Discontinuation; NSS, nephron sparing surgery; RFA, radiofrequency ablation; SAE: Serious adverse event, TRAE, treatment-related adverse event*

Appendix 4

Abstracts of the two non-English language records excluded from the clinical effectiveness systematic literature review

Sankaredja J, Brac B, Thines L, Baroncini M, Zairi F, Cardot-Bauters C, Lejeune JP. Épidémiologie, traitement et suivi des hémangioblastomes du système nerveux central dans le cadre de la maladie de von Hippel-Lindau. Rev Neurol (Paris). 2014 Apr;170(4):288-96. doi: 10.1016/j.neurol.2013.12.005.

Introduction:

Central nervous system (CNS) hemangioblastomas (HGB) are rare vascular tumors. The goal of this study was to analyze their epidemiology, treatment and prognosis in association with von Hippel-Lindau (VHL) disease.

Methods:

We retrospectively reviewed a series of patients treated in our department for a CNS HGB with VHL disease between 1996 and 2008. We analyzed pre- and postoperative clinical and radiological characteristics, number of visceral lesions (fundoscopy, abdomino- pelvian CT, metanephrines), clinical course (modified Rankin Scale and McCormick scale) and late prognosis (Kaplan-Meier survival curves).

Results:

We studied 19 cases (sex-ratio 0.9, mean age 36). The mean time to diagnosis was 61 days. The main symptom was intracranial hypertension for cerebellar lesions (7/15) and a sensitive-motor deficit for medulla oblongata (2/5) or spinal lesions (5/11). Preferred loca- tions were cerebellum (15/31), often nodulo-cystic appearance, followed by spinal cord (11/ 31), frequently coming with adjacent syringomyelia. Multiple locations and visceral lesions were found in two-third of the cases. Surgical removal was complete in more than three- quarter of the cases. Mean follow-up duration was 9 years. Postoperative mortality rate was 16%. In cerebellar and medulla oblongata locations together, final mRS was 1 in 17 of the 20 cases. In spinal cord locations, final McCormick score was 2 in all the cases. After delayed follow-up, about two-third of patients experienced recurrence or new progressive CNS lesions.

Conclusion:

HGB are rare CNS tumors. VHL disease should be considered when an HGB is diagnosed before 30, is located at the spinal cord, comes with multiple other CNS lesions or with typical peripheral lesions. Microsurgical removal is the gold standard treatment and can offer good functional results.

Wang EM, Wang BJ, Zhang N, Pan L, Dong YF, Zhou LF, Dai JZ, Cai PW, Chen H. Analysis of the results of gamma knife radiosurgery for hemangioblastomas of the brain and the factors related to the tumor recurrence. Zhonghua Yi Xue Za Zhi. 2004 May 17;84(10):813-7.

Objective:

To assess the 5-year-result of leksell gamma knife (LGK) in controlling hemangioblastomas of the brain (HB) and to analyze the factors related to tumor recurrence or development of new tumors.

Methods:

From November 1993 to September 2001, 35 patients, 28 males and 7 females, aged 36 (16 approximately 61), 18 with multiple tumors and 17 with solitary tumor, the number of tumors being 93 in total, were treated by LGK. Twenty-one patients with HBs were associated with von Hippel-Lindau disease (VHLD). The tumor size ranged 5 approximately 55 mm with a mean size of 13 mm. The mean maximum irradiation dose was 35.6 Gy (20.0 approximately 50.0 Gy) at tumor center and the mean minimum dose was 17.2 Gy (12.0 approximately 24.0 Gy) at tumor periphery. Fisher exact test, independent T test and Wilcoxon rank sum W test were used to analyze the results of LGK on solitary and multiple HBs, the recurrent time of the HBs, and the relation between minimum irradiation dose and tumor control.

Results:

35 patients had been followed for 24 - 114 months with a mean value of 66 months. 29 patients were alive and 6 died. Of the 29 patients 21 achieved satisfying tumor control, and 8 patients underwent open surgery because of tumor-associated cysts enlarging or development of new tumors after LGK. 21 patients had improvement or remained stable in neurological status. Of the 8

reopened patients, 2 had deteriorated symptoms and the other 6 remained neurologically stable. Of the 35 patients, 7 developed new tumor during the follow-up period, and 5 had second LGK. Tumor control: Of the 29 cases, solitary or multiple tumors in 23 patients decreased in volume or remained the same, although two developed new tumors. The result of LGK in controlling HBs showed no significant difference between the solitary and multiple HBs ($P > 0.05$), but the dose of long-term tumor control was significantly higher than that of uncontrolled tumors ($W = 98$, $P < 0.01$). The tumor control rate was 94% 1 year after; 85% 2 years after; 82% 3 years after; 79% 4 years after; and 71% 5 years after. For the patients with solitary tumor, the mean time of development of new tumor was 63 months, but for the patients with multiple HBs, the time was 25 months. There was a significant difference between the two groups ($t = 3.987$, $P < 0.001$). With margin dose of 18 Gy, histopathology showed that no tumor cell was found and there were coagulation necrosis, hyaline degeneration and fibrosis tissues in the tumor nodule 48 months after LGK.

Conclusion:

LGK is a good choice for small- or medium-sized, solid HB in long term, especially when tumor margin dose is 16 - 20 Gy. Although LGK can treat multiple tumors in one single treatment session, for HB in patients associated with VHLD, LGK faces the problem of tumor recurrence or development of new tumor.

Appendix 5

Compilation of data available on transition rates from the MK-6482-004 study, its pre-treatment period, the VHL Natural history Study and the Optum Clinformatics Data Mart Claims Study

Please note that the table below shows a selection of the data that were *available* and is not limited to only what were, or what should be, *used* as inputs into the cost-effectiveness analyses. Data from these that were selected for use as inputs into the cost-effectiveness analyses are shown in green, additional information on what inputs were used in the cost-effectiveness analyses and why they were selected are provided in section B.3.3 of the company submission.

Table 13 Compilation of data available on transition rates from the MK-6482-004 study, its pre-treatment period, the VHL Natural history Study and the Optum Clinformatics Data Mart Claims Study

Event for rate/hazard calculated	Events/person-week (standard error) [effective sample size]			
	MK-6482-004 study results	MK-6482-004 pre-treatment period data	VHL Natural History Study (post-matching, exponential parametric multistate modeling transition probability)	Optum Clinformatics Data Mart claims study (unadjusted data)
RCC				
Pre-surgery → 1st surgery	0.00071 (0.00027) [61]	NR	0.00487 (0.00034) [92.2]	[redacted]
Pre-surgery → metastatic disease	NR	NR	0.00004 (0.00002) [92.2]	[redacted]
Pre-surgery → death	NR	NR	0.00012 (0.00003) [92.2]	[redacted]

Event for rate/hazard calculated	Events/person-week (standard error) [effective sample size]			
	MK-6482-004 study results	MK-6482-004 pre-treatment period data	VHL Natural History Study (post-matching, exponential parametric multistate modeling transition probability)	Optum Clinformatics Data Mart claims study (unadjusted data)
Event-free after 1st surgery → next surgery	NR	NR	0.00166 (0.00015) [75.7]	[redacted]
Event-free after 1st surgery → metastatic disease	NR	NR	0.00006 (0.00002) [75.7]	[redacted]
Event-free after 1st surgery → death	NR	NR	0.00023 (0.00005) [75.7]	[redacted]
Metastatic disease → death	NR	NR	0.00199 (0.00074) [7.2]	[redacted]
pNET				
Pre-surgery → 1st surgery	NR	0.00017 (0.00010) [22]*	0.00027 (0.00008) [60.4]	[redacted]
Pre-surgery → metastatic disease	NR	NR	0.00013 (0.00005) [60.4]	[redacted]
Pre-surgery → death	NR	NR	0.00021 (0.00007) [60.4]	[redacted]
Event-free after 1st surgery → next surgery	NR	NR	Not evaluable, model did not converge	[redacted]
Event-free after 1st surgery → metastatic disease	NR	NR	0.00038 (0.00023) [14.1]	[redacted]
Event-free after 1st surgery → death	NR	NR	0.00036 (0.00021) [14.1]	[redacted]
Metastatic disease → death	NR	NR	0.00158 (0.00097) [7.6]	[redacted]
CNS-Hb				

Event for rate/hazard calculated	Events/person-week (standard error) [effective sample size]			
	MK-6482-004 study results	MK-6482-004 pre-treatment period data	VHL Natural History Study (post-matching, exponential parametric multistate modeling transition probability)	Optum Clinformatics Data Mart claims study (unadjusted data)
Pre-surgery → 1st surgery	0.00010 (0.00010) [50]	0.00202 (0.00032) [50]*	0.00072 (0.00009) [37.9]	[redacted]
Pre-surgery → metastatic disease	NR	NR	0.00006 (0.00002) [37.9]	[redacted]
Pre-surgery → death	NR	NR	0.00028 (0.00005) [37.9]	[redacted]
Event-free after 1st surgery → next surgery	NR	NR	0.00193 (0.00032) [37.9]	[redacted]
Event-free after 1st surgery → metastatic disease	NR	NR	Not evaluable, model did not converge	[redacted]
Event-free after 1st surgery → death	NR	NR	0.00024 (0.00008) [37.9]	[redacted]
Metastatic disease → death	NR	NR	0.00118 (0.00079) [37.9]	[redacted]

* For active surveillance in the CNS-Hb and pNET populations, an important limitation of the VHL Natural History Study data was the inability to identify patients who had CNS-Hb and pNET tumours at the patient-level index date. Consequently, for these populations, the best available data source for pre-surgery → first surgery was the surgery event data collected for MK-6482-004 trial participants during the pre-treatment period of the MK-6482-004 study, as patients' CNS-Hb and pNET tumour status was identifiable at the baseline visit of the MK-6482-04 study.
† These values were used to derive adjustment factors used in the cost-effectiveness analyses, but these values themselves are not used in it.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

Technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR, in sections 1.1 and 1.3 to 1.5. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Clinical expert statement

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **28 September 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Clinical expert statement

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating Von Hippel-Lindau disease (VHL) in adults with and current treatment option

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	[redacted]
2. Name of organisation	UK Kidney Association
3. Job title or position	Regius Professor and Head of School of Clinical Medicine, Cambridge
4. Are you (please tick all that apply)	☒ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with VHL? (although not treating such patients currently) ☐ A specialist in the clinical evidence base for VHL or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☒ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

8. What is the main aim of treatment for VHL? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	To avoid death from metastatic cancer; to avoid loss of kidney function; to avoid loss of pancreas function; to avoid loss of vision; to avoid disability from central nervous system haemangioblastomas
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	Avoiding / reducing the need for repetitive surgeries in the kidney and pancreas. Any significant reduction in size and symptomatic effect of CNS haemangioblastomas
10. In your view, is there an unmet need for patients and healthcare professionals in VHL?	Yes
11. How is VHL currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	Annual surveillance imaging of the kidneys from age 16. Once a lesion is identified, it is monitored and when a lesion approached 3cm it is removed surgically by partial nephrectomy or by ablation and at the same time other lesions in the kidney are removed. There are also guidelines concerning identification and monitoring of pancreatic tumors and of CNS haemangioblastomas, which are covered by the MHRA label and the company's submission Guidelines from Maher ER et al European J Human Genetics 2011 19:617-623 The treatment is expected to reduce the rate of surgical interventions, reduce the risk of requiring renal replacement therapy due to removal of renal tissue, reduce the loss of pancreatic function, and reduce the risk of disability from CNS haemangioblastoma. Precisely who should be treated with it, and when it should be initiated, will be challenging. It is also possible that current criteria for when an operation should be performed (eg size of renal tumor) will need changing for patients who are on treatment with bezultifan since the effect on tumour biology is likely to be profound.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	The technology will be quite different as it will alter growth of tumors. Currently the approach to managing patients is screening to identify tumors, imaging to monitor known tumors and surgical removal when they are causing problems or to prevent metastasis.

• How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	The technology should be used in secondary / tertiary care in multidisciplinary VHL clinics.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	I think it is possible that it will extend life in some individuals. I think it is very likely to improve health related quality of life compared to current care in some VHL patients.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	I think the largest benefit will be for patients with symptomatic CNS haemangioblastoma
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	Working out who should be treated with the technology and when they should be treated, monitoring treatment appropriately and dealing with side effects will add complexity to looking after VHL patients.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	The rules based on the label concerning need for intervention and being unsuitable for surgery are sensible.

17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Yes, the technology is clearly a step change.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The main side effect is a reduction in red blood cell production leading to anaemia, which is straightforward to monitor and manage.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The main clinical trial focussed on individuals having surgery for renal cell carcinoma. The approach to monitoring and surgical intervention in the UK is essentially the same as that in the trial. As VHL patients (even within one family with a single gene mutation) have theior own range of tumours in the different target organs, and the individual tumours grow at different rates, it is quite difficult to extrapolate results from one patient to another. Surgical procedures on the kidney do lead to loss of renal function, so decreasing the number of surgical procedures will preserve renal function and reduce the risk of needing dialysis and transplantation. Similarly decreasing the

	number of procedures on the pancreas will decrease the likelihood of pancreatic insufficiency
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	My understanding is that the response in the trial reads across well into the real world, with a response in many VHL related tumors and a decrease in the rate of needing surgical interventions.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people.	VHL affects people of both sexes, and all races. Many VHL patients end up with significant disability making it hard for them to attend further hospital appointments, and reducing their ability to be effective advocates for better care.

Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. **Find more general information about the Equality Act and equalities issues here.**	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Implication of differences between intervention and comparator populations given interpretation of the marketing authorisation that standard of care for most patients is immediate surgery	This is very difficult. Most surgery in VHL patients is “urgent”, rather than “immediate”. It is carefully planned and needs to happen in a timescale of weeks. It will almost always have undesirable consequences. A real problem is that to have more certainty about the probabilities affecting the issues below would require very large studies given the variability of the condition. But as VHL is rare, such studies are unlikely to be possible; especially given that bezultifan is clearly effective so that achieving equipoise in any randomised trial would now be very difficult.
Key issue 2: Misalignment between the decision problem and MK-6482-004 study populations; and between the latter and the UK target population	The decision problem is, in my view, somewhat artificial given the variability in clinical course between individual VHL patients, and even between different tumours in the same patient.
Key issue 3: Potential risk of study selection bias resulting in possible omission of relevant comparator studies	I think this is unlikely as the awareness of studies is high

Key issue 4: Lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator	I am not able to comment on this. I find the cost effectiveness hard to understand and evaluate
Key issue 5: Limitations in the indirect treatment comparison hinder the assessment of the effectiveness of belzutifan compared to standard of care	This is difficult as bezultifan is so different from current approaches, and precisely how it will alter outcomes over time is somehat uncertain
Key issue 6: There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses	I find the economic analyses hard to understand and evaluate
Key issue 7: The comparator data might not be representative for the UK	The approach to managing VHL is very similar in the UK to the approach taken in the trial
Key issue 8: Data to inform effectiveness in the belzutifan arm (MK-6482-004 trial) are either immature or unavailable	I think the effectiveness of bezultifan on tumor growth and the need for surgery is sufficiently clear in the study. I also think it is clear that responses are sustained and it is well tolerated. It is important to recognise that the scientific rationale for thia as a treatment is very strong indeed
Key issue 9: There is uncertainty in the derivation of the transition probabilities in the standard of care arm	The uncertainty here is unsurprising since the clinical course is so variable
Key issue 10: There is uncertainty in the implementation of time on treatment and treatment effect waning	The uncertainty is unsurprising. Scientifically I consider it much less likely that there will be a waning of treatment effect compared to treatments in most cancer conditions.
Key issue 11: There is uncertainty in the derivation and implementation of health-related quality of life in the model	The uncertainty is unsurprising, given the unpredictable nature of VHL disease.
Key issue 12: Cost-effectiveness analyses should be based on subgroup-specific parameters	I am not able to comment on this, but think that in a condition that is as rare as VHL disease, and as variable in its clinical manifestations and their behaviour the subgroups will be small, and will still be heterogenous
Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Bezultifan is a completely different approach to current management of VHL disease which does not include any measure that alters tumour development and growth
Bezultifan has been shown to reduce the requirement for surgical procedures in VHL disease, and this is based on an exceptionally strong scientific rationale
Bezultifan is well tolerated and there does not appear to be waning of treatment effect
The comparisons required for the decision are challenging because of (a) the variable course between VHL patients, and between individual tumors in the same patient (b) the rarity of VHL disease (c) the magnitude of the step change copmpared to the current approach
Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with von Hippel-Lindau disease or caring for a patient with von Hippel-Lindau disease. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR, in sections 1.1 and 1.3 to 1.5.

A patient perspective could help either:

- resolve any uncertainty that has been identified OR

- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **28 September 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with von Hippel-Lindau disease (VHL)

Table 1 About you, VHL current treatments and equality

1. Your name	[redacted]
2. Are you (please tick all that apply)	☒ A patient with VHL? ☐ A patient with experience of the treatment being evaluated? ☒ A carer of a patient with VHL? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	VHL UK/Ireland
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☒ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☒ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☒ I am drawing from personal experience ☒ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: Support patient and carer members of VHL UK/Ireland

Patient expert statement

	☒ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with VHL? **If you are a carer (for someone with VHL) please share your experience of caring for them**	Father with VHL: (current age 78) Loss of vision in one eye as a teenager, multiple laser coagulation treatments in the other eye and now registered blind, 1 brain surgery, 1 partial nephrectomy. Early medical retirement. Me: (Current age 42) - diagnosed under age 10 through eye screening and then confirmed with genetic testing when it became possible several years later. I have had: Eyes: Blindness and then removal of one eye, countless laser coagulation treatments in the other eye Brain: 2 surgeries, 1 shunt, multiple lesions remain 'on watch' with one causing particular concern as it is thought to be a 'regrowth' presenting greater complications if/when surgery is required Spine: Several lesions with over a decade of resulting numbness in one leg. Kidney: 1 Cryoablation for RCC, multiple cysts remain Pancreas: Whipple's procedure and now on PERT for life Family planning: PGD, carrying a pregnancy and surrogacy Although this is long, I hope you will read it as I have tried to illustrate exactly how VHL has affected every aspect of my life: • Multiple interruptions to my primary, secondary and university **education** for screening and/or treatments. Day to day lives are regularly interrupted with surveillance and follows ups across multiple disciplines. An average year for

	me (without surgery) is 14 appointments, most of which involve a whole day away from work and family duties. • For many years, I was frightened to tell my employers that I had VHL for fear of being a burden and missing **career** opportunities (from my support work I know this is common in young people with VHL). Now I am self-employed I struggle with committing to clients and am fearful of being unable to fulfil contractual obligations. • One of my biggest vulnerabilities is that I have been unable to **secure life insurance** as I am considered a 'write off' and I fear for the financial security of my family should I be unable to contribute to our mortgage to the end of the term. I do not want them to lose their home, we have worked so hard for. • **Travel** insurance is expensive and planning for travel burdensome. I feel **uncomfortable planning** any significant **social or leisure activity** too far in advance for fear of being unwell or unable and letting anyone else down in the process (family, friends etc) • My first **brain surgery** was at 23. I was doing well until I developed a CSF leak 2 weeks post op. This led to months of ambulances, hospital admissions, lumber punctures, steroids, crippling painful spasms, meningitis, Christmas day in hospital, missing one of my close friends weddings and some **dark thoughts** that I didn't want to carry on if there was no way to improve things, thankfully after a shunt things improved. • All my **relationships** have been affected and it was around the time of this first surgery that I resisted commitment to my now husband because I wasn't sure we could/should have a family in future. By then was certain I did not want to risk passing on VHL (50/50 chance). • However, when we discovered **PGD** was a possibility, I was able to carry my eldest daughter. However, the pregnancy badly accelerated the health of my bad eye and that is when it went totally **blind** and as my family left for a short holiday without me, I headed to the hospital, some 26 weeks pregnant to have a cyclodiode procedure. The surgeons warned me it would be a very

	painful procedure and were concerned for me. They were shocked when I reported it was nothing compared to the pain that had led to it. Some 10 years later the worst pain I have ever had to endure came from the blind eye that was now calcifying and a good eye that was experiencing '**sympathy pain'**. I was unable to care for my (by then, 2 children) and laid for weeks on end, in a dark room with both eyes shut. For the only time in my life, I **broke down** in front of one of my much-respected medical team and begged for the eye to be removed. I had never before and never since had to lobby for myself on the basis of my **quality of life,** which had become quite simply, unbearable. • Advised not to carry again but with 2 VHL free embryos, a desire to be a Mum again and for my daughter to have a sibling (something important to me from a VHL perspective, so that they may support each other in hard times) we sought out the help of a **surrogate**. Whilst this resulted in our youngest daughter, an overall wonderful experience and friends for life, it significantly impacted us **financially** and **emotionally**. • There were then a good number of years where VHL wasn't too troublesome. Laser treatments were effective for my good eye and I was lucky enough to qualify for cryoablation when my **kidney RRC** reached a size it could no longer stay. It was a 'manageable' experience with a small post op bleed and a fever that lasted around 2 weeks. The surgeon explained that had the ablation not been possible, I would have lost my kidney due to the location of the tumour. • However, in 2018, after I questioned a spot showing up on scans of my **pancreas** and seeking out the best endocrinology team I could find, it was confirmed I had a PNET, 'the one no VHLer wants to develop'. 4 years later, I had no choice but to undergo the Whipple's procedure in 2022. • 2022 was the worst year of my life. I have felt **guilt** about VHL all my life. I have no choice to endure what VHL throws at me but my husband did and although I am extremely proud that we made sure my children do not have

	VHL, I always worry about the impact on them of having a parent with it. That became a reality last year. • Every VHLer dreads the **Whipple's**. And rightly so. I don't think I am able to put into words what **a propound impact it has had on me physically and mentally** and over a year out now, I am not sure I will ever be the same person again. To add to the eye the **missing parts of my body** now include, the head of my pancreas, some stomach, my gallbladder and a chunk of my duodenum. I now take medicine with everything I eat to hep me digest it (and must remember to take the enzymes everywhere I go). I spent 2 days in ICU, 2 weeks in hospital barely able to speak or move and 2 months battling fevers, severe weight loss, nausea, pain and fatigue. I wouldn't say I was medically fit for 6 months and deemed 'fully recovered until a year post op. My husband and kids made **so many sacrifices** over that time and so often I could do nothing for them. • 5 months post Whipple, I had no choice but to go for **brain surgery**. It wasn't a surprise, a 'quiet' lesion had been on closer watch for about a year and when I saw the latest scan, I simply said to my surgeon, 'ok so it's when not if'. • I thought I was recovered enough from the Whipple but with hindsight I wasn't and although the surgery went well and I returned home after 2 nights, I blacked out and hit my eyelid missing my only seeing eye by millimetres. I became so **weak and unwell**, losing more weight and developed another suspected CSF leak, threatening more surgery. • If was at this time that my husband suffered a **mental breakdown**. The months of **worry**, supporting us financially and playing 'Mum and Dad', had become too much. I tried to support him, pick up the slack as much as I was able but often physically couldn't. And then my biggest fear came true, as I discovered my eldest daughter was in the midst of her own **severe mental health crisis** seeing both her parents waning. My world was imploding and I have never felt so **desperate and out of control**, in all my life. Thankfully, I

found some strength from somewhere to pull us through and 2023 has been much kinder to us all.

- I must also mention the challenges I have faced being the child of a VHL patient (and at times, co carer). Despite all I have endured, I speak with experience when I say I believe it is worse to watch than to endure. The worry and helplessness is profound. I collected my a-level results the day my Dad went in for emergency brain surgery. Although rare, sometimes the onset of symptoms can be sudden and severe and the previous few days I'd seen my strong and powerful Dad losing his balance, falling and being sick every time his head moved. Some years later his partial nephrectomy was horrific to witness, the surgeon told us it was one the most difficult operations he had ever performed. His post op pain and suffering was intense. Then the deterioration of his one seeing eye was devastating. His one pleasure in life was to read books and that was no longer possible. These days his general health means he would be unlikely to manage any more surgeries and so we hope that conversation never comes. Where possible I try to group our appointments and we often marvel at ourselves as the 'blind leading the blind'. I feel deeply sorry for my Mum, who had has to watch and support both her daughter and her husband endure so much with VHL.
- I always struggle to make **resolutions for the future.** Whilst I try to remain positive and certainly don't dwell on the effects I have listed above, I often can't help but feel I live with multiple ticking time bombs and at some point one of them might well take me out or make my life unbearable.

Despite all that I have described, I genuinely believe **I am one of the lucky ones**. I speak with so many VHL patients and carers who are enduring so much worse. These are just some of them:

- A patient who woke from her brain surgery to discover her daughter had been admitted via A&E and was having brain surgery the following day herself. Their support system for each other in complete turmoil.

	• A number of patients who are wheelchair bound after spinal surgery • A young girl who is paralysed and requires support to breath • A dear charity member who lost her life to VHL this year when her Whipple's procedure was abandoned mid surgery due to metastasis and subsequent and futile attempts to treat with available medicines all failed. • A number of patients who have endured more than 10 brain surgeries • A patient who has had over 40 operations and whose husband has given up work to be her full time carer. • At least 3 young people who have lost sight in one eye and are running out of options for vision in their remaining eye. Blindness is a real possibility for them • A lady who had brain surgery shortly after giving birth to her first child and having spent much of her maternity leave recovering, is now facing RCC.
7a. What do you think of the current treatments and care available for VHL on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	7a. The only effective current treatment is surgery. Whilst I am grateful for the skills, knowledge and kindness of the surgeons who have helped myself my father and many VHL patients across the UK, until now we have had no choice but to face the huge risks and consequences associated with them, or we will die. 7b. Thanks to modern day wonders such as social media, patients are well connected across the world and therefore aware of patients already using belzutifan either via the trials or since approval in their countries. Patients are regularly reporting their good results of regressing or disappearing tumours, **something no VHL patient has ever seen before** without high risk surgeries.
8. If there are disadvantages for patients of current NHS treatments for VHL (for example, how they are given or taken, side effects of treatment, and any others) please describe these	The disadvantage are: • The often-high risks associated with surgery but having no other choice. (Paralysis, stroke, death for brain and spine, loss of part/whole organs and

	metastases with RCC and PNET, diabetes, malabsorption and need or lifelong medication such as PERT or dialysis. • The physical effect on the body of often multiple procedures either in the same organ or multiple organs throughout patients' lives can be devastating • The effect on every aspect of theirs and their lives as fully detailed above in Q6
9a. If there are advantages of belzutifan with over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does belzutifan help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	**9a.** As per the VHL UK Patient Group submission, our knowledge of patients and carers and the surveys run, clearly demonstrate the ability of belzutifan to reduce the number of surgeries which will have a profound impact on every aspect of a patient and carers lives. It also has the ability to affect more than just the target tumour at the same time. Something surgically removing an induvial tumour cannot do. **Please refer to the surveys in the submission** **9b.** They all go hand in hand. Less surgery, less physical burden on the body, better wellbeing and quality of life in all aspects and some hope for the future for both patient and carer. **9c.** Yes. All of them. As per answer to 9b.
10. If there are disadvantages of belzutifan current treatments on the NHS please describe these. For example, are there any risks with belzutifan? If you are concerned about any potential side effects you have heard about, please describe them and explain why	As per the VHL UK Patient Group submission (please refer to this), although side effects are common, they are considered 'on target' and tend to improve over time as the body adjusts. In our survey it is clear most patients will gladly tolerate them than face the risks of the alternatives, this cannot be underestimated.
11. Are there any groups of patients who might benefit more from belzutifan or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	VHL is such a complex and varied disease and no 2 patients are the same. Personally, I feel that the patients who would benefit the most are those who face the greatest risks. For example: • Complete blindness • Loss of part or all of an organ which will have life changing effects (dialysis, diabetes, PERT etc)

	• Those who have metastasis as all known current therapies are widely accepted as 'useless' AND have devastating side effects • Those who face the most complex brain or spine surgeries with the highest risk of a poor outcome or have already endured so many that scar tissue etc pose a greater risk • Those for whom surgery is not possible, either because of a tumours position or the patients general health Those that would not currently benefit are those who have stable disease and/or where there are clear, low risk options available to them (such as laser for the eyes or ablations for RCC).
12. Are there any potential equality issues that should be taken into account when considering VHL and belzutifan? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	I agree with VHL/UK Ireland Patient Group submission.

13. Are there any other issues that you would like the committee to consider?	Patients have or are considering moving continent to access belzutifan. The charity is urging others who are considering delaying surgery in anticipation of being able to able to access it, to speak with their medical teams.

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement – I plan to add some lived examples once we are happy with the VHL UK/Ireland submission

Key issue 1: Implication of differences between intervention and comparator populations given interpretation of the MA that standard of care for most patients is immediate surgery	
Key issue 2: Misalignment between the decision problem and MK-6482-004 study populations; and between the latter and the UK target population	
Key issue 3: Potential risk of study selection bias resulting in possible omission of relevant comparator studies	

Patient expert statement

Key issue 4: Lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator	
Key issue 5: Limitations in the indirect-treatment comparison hinder the assessment of the effectiveness of belzutifan compared to standard of care	
Key issue 6: There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses	
Key issue 7: The comparator data might not be representative for the UK	
Key issue 8: Data to inform effectiveness in the belzutifan arm (MK-6482-004 trial) are either immature or unavailable	
Key issue 9: There is uncertainty in the derivation of the transition probabilities in the standard of care arm	
Key issue 10: There is uncertainty in the implementation of time on treatment and treatment effect waning	
Key issue 11: There is uncertainty in the derivation and implementation of health related quality of life in the model	
Key issue 12: Cost-effectiveness analyses should be based on subgroup-specific parameters	

Patient expert statement

Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- VHL is a devastating lifelong disease that effects every aspect of patients lives and wellbeing.
- VHL patients have NEVER before seen their tumours regress of disappear without the use of often multiple, life changing or life threating surgeries. The best they have ever been able to hope for is stability.
- VHL is not the same in any 2 patients, even within the same family where the gene mutation is the same. Which manifestations, how often, when they manifest, how long they take to need intervention and the resulting outcomes and impact on quality of life varies enormously, presumably that results in real difficulty estimating cost efficiency.
- Belzutifan offers patients the real chance to avoid high risk surgeries and hope for theirs and their carers futures. It would be devastating for it not to be approved for use in the UK, especially for those most in need.
- Patients have or are considering moving continent to access belzutifan. The charity is urging others who are considering delaying surgery in anticipation of being able to able to access it, to speak with their medical teams.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with von Hippel-Lindau disease or caring for a patient with von Hippel-Lindau disease. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR, in sections 1.1 and 1.3 to 1.5.

A patient perspective could help either:

- resolve any uncertainty that has been identified OR

- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **28 September 2023.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with von Hippel-Lindau disease (VHL)

Table 1 About you, VHL current treatments and equality

1. Your name	[redacted]
2. Are you (please tick all that apply)	☐ A patient with VHL? ☐ A patient with experience of the treatment being evaluated? ☒ A carer of a patient with VHL? ☐ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	VHL/UK Ireland
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☒ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☒ I am drawing from personal experience ☐ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: ☐ I have completed part 2 of the statement **after attending** the expert

	engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with VHL? **If you are a carer (for someone with VHL) please share your experience of caring for them**	In 2015 I was engaged to a VHL 'warrior' and we were married in 2016. My husband had his first manifestation of a VHL tumour in the brain at the age of 11 in the year 2000 and he has had multiple manifestations and surgeries since. It was spontaneous in him, neither his parents or siblings tested positive for VHL. He had growths in brain, eyes, spine, kidney and pancreas. Treatments have been surgeries (10 alone to the brain of which 1 was to insert a VP shunt), gamma knife radiation twice to brain stem tumours, plus an attempt at embolisation, spinal surgery, laser treatment to eye, partial nephrectomy to kidney and pancreatic cyst aspiration. Married life was a roller-coaster due to his medical condition. Each year there were constant scans and tests, numerous medical visits, plus I had to deal with all the emotions and anxiety that went along with it. I was dealing with my husband's emotions from past trauma and anxiety of declining health, other members of the family who were anxious and concerned and my own emotions too. I lived on a knife edge with this fear of "where is this going to crop up in his body next" and "will he survive this next surgery – if he does how will he be impaired afterwards". There was no pattern to where the next manifestation would be, but the worst area of growth for my husband was in the brain (cerebellum). There wasn't only the scans and tests, but then appointments that followed to find out the results, which for me as a carer was the worst part. Report of stable growths was news to be celebrated, but our experience mostly was new growths and existing ones that had increased in size. There was no let-up in the appointments, they only ever increased in number. Gastric issues were profound and every day would start with retching or vomiting and bowel challenges. We had a constant battle to find medications that would help his symptoms of nausea, dizziness and pain. My husband had days, weeks and months off work due to ill health. He was physically very weak as a result of

	everything that had taken place in his life. Not being well enough to work or able for much practical occupation, he got very mentally low and depressed too. We then had the psychotherapy journey, alongside everything else, to try and find something that could help. We couldn't socialise much, as my husband was not well enough to get out or travel far. Occasionally I would ask a friend if they could please come to see him for 15 minutes to give him a boost. He had physio and speech therapy following brain surgery for months on end, learning to walk and talk again and striving to get back to somewhere near where he was before. The brain surgeon who operated on him the last time (in 2018) said to me when I conveyed my husband's downcast statement that he didn't think he would walk again "I am afraid it is the number of times they've been in". The toll it was taking on his body was indescribable and immensely hard to watch someone go through this level of suffering in their body and mind. I felt helpless most of the time and my husband felt hopeless and useless. There was no way we could consider starting a family - my husband wasn't well enough, nor did we want to risk passing on this awful disease onto a child for which there was no treatments with many successful outcomes. As I have described above, his quality of life went downhill from when I first met him until he got onto the Belzutifan trial for VHL patients in 2018 and hope was held out to him for the first time since being diagnosed with VHL in 2000. With the help of a friend, we searched for top advice for VHL and were put in touch with the MD Anderson Cancer Centre in Houston and visited there at the end of 2017. Dr Eric Jonasch started my husband on the trial as soon as it opened in May 2018. The manifestations and surgeries stopped when he started on the Belzutifan drug …. by August 2018 he was having the first round of scans which showed something we had never experienced in our lives before – SHRINKAGE by 33%. This is something VHL patients rarely hear – our life changed that year! My husband learnt to walk and talk again after his last surgery in January 2018 and now works a full 5-day week because of the effect of Belzutifan in shrinking his growths and keeping them stable! We can travel again (even abroad!), get out and about with friends, help out in the community and enjoy food and company like never before! Even his

	pancreas function has recovered to a normal level again, so he no longer has to take enzymes with every snack or meal.
7a. What do you think of the current treatments and care available for VHL on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	7a. From my experience I only know of limited treatments available that may momentarily halt some growth, but rarely does anything have a lasting effect. Surgery removes problematic tumours and cysts can be drained, but these either form again or there are new manifestations. This obviously has a huge negative impact on quality of life. The medical staff do the best they can with what is currently available to treat and care for the patient, for which we are very grateful. 7b. I only know of other VHL patients having similar experiences to ourselves, some patients have loss of sight from the disease and long-term damage from surgeries, sometimes having parts of vital organs removed.
8. If there are disadvantages for patients of current NHS treatments for VHL (for example, how they are given or taken, side effects of treatment, and any others) please describe these	Current treatments for VHL can cause lasting damage to the body due to impairment from surgery. Multiple interventions over time weakens the body and each one takes longer to get over. Loss of sight, damage to nerves causing speech, swallowing, numbness, autonomic, continence and mobility issues. Medications given can cause side effects of nausea, vomiting, dizziness, constipation, etc and sometimes severe reactions. My husband had anaphylaxis as a result of one anti-sickness drug the Drs tried him on.
9a. If there are advantages of belzutifan with over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does belzutifan help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	9a. Belzutifan is highly successful in shrinking growths that no treatment has ever done before for VHL related growths. A huge advantage is that Belzutifan has comparatively low side effects compared to other current treatments; it is well tolerated by patients. Belzutifan saves the patient and NHS vast amounts of time and resource by: • Reduced visits to the doctors/surgeon/pharmacy/therapy • Less time off work sick/recovering from treatments/surgeries

	• Reduced number of medications for the many symptoms that come with increased VHL growth and new manifestations, or side effects as a result of surgeries and treatments. All the above combined makes for an improved quality of life where the patient can care for themselves and other family and community members more, as well as work and contribute to the economy and society a lot more. Another enthusing benefit is that my husband has a work colleague in the same office who has kidney cancer which was getting worse and spreading, until he started on a trial of Belzutifan and experienced shrinkage in his tumours too. We are so happy for them and any other patients who have got on this trial. It brings such positivity and hope into the lives of many who are benefitting from it. 9b. Living with a much better quality of life that is more predictable and the huge positive impact on the oncoming generations to have a highly successful treatment which has comparatively low side effects. 9c. Yes, Belzutifan addresses all issues, as it prevents lasting damage to the body by shrinking the growths &/or holding them stable and therefore preventing medications, surgeries and the recovery from multiple surgeries, which takes its toll on the patient as time goes on.
10. If there are disadvantages of belzutifan current treatments on the NHS please describe these. For example, are there any risks with belzutifan? If you are concerned about any potential side effects you have heard about, please describe them and explain why	Not that I am aware of. I have no concerns.
11. Are there any groups of patients who might benefit more from belzutifan or any who may benefit less? If so, please describe them and explain why	Not to my knowledge

Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	
12. Are there any potential equality issues that should be taken into account when considering VHL and belzutifan? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	Not to my knowledge
13. Are there any other issues that you would like the committee to consider?	

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Implication of differences between intervention and comparator populations given interpretation of the MA that standard of care for most patients is immediate surgery	
Key issue 2: Misalignment between the decision problem and MK-6482-004 study populations; and between the latter and the UK target population	
Key issue 3: Potential risk of study selection bias resulting in possible omission of relevant comparator studies	

Key issue 4: Lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator	
Key issue 5: Limitations in the indirect-treatment comparison hinder the assessment of the effectiveness of belzutifan compared to standard of care	
Key issue 6: There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses	
Key issue 7: The comparator data might not be representative for the UK	
Key issue 8: Data to inform effectiveness in the belzutifan arm (MK-6482-004 trial) are either immature or unavailable	
Key issue 9: There is uncertainty in the derivation of the transition probabilities in the standard of care arm	
Key issue 10: There is uncertainty in the implementation of time on treatment and treatment effect waning	In my experience as a carer, my husband is still responding positively to the drug over 5 years later - he started on the trial in May 2018.
Key issue 11: There is uncertainty in the derivation and implementation of health related quality of life in the model	
Key issue 12: Cost-effectiveness analyses should be based on subgroup-specific parameters	

Patient expert statement

Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Belzutifan is a life-changing drug which improves quality of life for VHL patients, enabling them to live a longer life doing usual activities, like others, with reduced pain and discomfort and many other unpleasant symptoms.
- There has been no other treatment before like Belzutifan for VHL patients; the first of its kind with a high success rate on growths in multiple areas of the body and is proving successful in kidney cancer patients too.
- Belzutifan prolongs the life of VHL patients, enables them to care for themselves and contribute more to society, local communities, family life, the economy, the workplace – the list is endless!
- Having a treatment for VHL patients prevents lasting damage to the body by shrinking the growths &/or holding them stable and therefore preventing surgeries and the recovery from multiple surgeries, which takes its toll on the patient as time goes on.
- Belzutifan hugely improves mental health as this first of its kind treatment gives VHL patients hope, reassurance and more predictability, knowing there is a highly successful treatment available to them, which has never been proved before.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Belzutifan for treating tumours associated with Von Hippel-Lindau disease [ID3932]

EAG response to TE submission including cost effectiveness sections

Produced by Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Susan O'Meara, Reviews Manager, KSR Ltd, United Kingdom (UK)
Venetia Qendri, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands (NL)
Eline Krijkamp, Health Economist, ESHPM, EUR, NL
Jiongyu Chen, Heath Economist/Systematic Reviewer, KSR Ltd, UK
Xiaoyu Tian, Health Economist/Systematic Reviewer, KSR Ltd, UK
Mubarak Patel, Systematic Reviewer, KSR Ltd, UK
Rachel Croft, Information Specialist, KSR Ltd, UK
Lisa Stirk, Senior Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to Susan O'Meara, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 25/10/2023

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/56/85.

Declared competing interests of the authors None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.
Confidential comparator prices are highlighted in green throughout the report.
Any de-personalised data are highlighted in pink throughout the report.

Table of Contents

EAG's comments on company's Stakeholder Engagement Response From

Key Issue 1, 2 and 6: Implication of differences between intervention and comparator populations given interpretation of the MA that standard of care for most patients is immediate surgery, Misalignment between the decision problem and MK-6482-004 study populations; and between the latter and the UK target population, and There is mismatch between the population in the economic analyses and the population included in the sources of evidence used to inform such analyses

The EAG note that the company recognise the difference between the MA population, which they maintain should be the same as the decision problem population, and the population of the MK-6482-004 study. They also, in the history of obtaining MA, show that a population description that would have been consistent with MK-6482-004 is "not requiring immediate surgery". As stated in the EAG report, this shows the large discrepancy, as opposed to "slightly different": Not requiring immediate surgery implies a much less severe stage of disease than surgery being "unsuitable or undesirable". As also stated in the EAG report, this also highlights the discrepancy between the populations implied by immediate surgery only occurring with standard care i.e., more like the DP or the MA, as opposed to no immediate surgery with belzutifan i.e., more like the MK-6482-004 study.

The company go on to argue that no surgery is required if belzutifan is given because belzutifan is "an effective therapy" and that the "onset of efficacy is rapid". This might make sense if no substantial harm might befall patients who have to wait to see if belzutifan is effective. However, as the EAG report stated, the company stated that surgery was "the only treatment option available to keep patients alive…". As stated in the EAG report, it cannot be true that such immediate surgery is required without belzutifan, but not with it. Indeed, since the population in the MK-6482-004 study is those not requiring immediate surgery, then active surveillance as comparator seems to be, in contrast to the company's assertion, entirely plausible. This would imply a change to the DP, which might not be consistent with the MA: if this is the case then, as stated in the EAG report, perhaps there needs to be some differentiation between patients who need lifesaving surgery and those who need it for symptom relief or progression prevention.

The company also refer to an expert elicitation exercise, a summary of which is presented in Appendix 2 and in response to which the EAG have already presented an addendum. Although the experts' responses have not been provided, this summary reveals nothing that would change the EAG critique. Therefore, key issues 1, 2 and 6 remain relevant.

Finally, the company discuss potential belzutifan benefits that are not captured in the current economic model. While these could be relevant, the lack of evidence remains a main concern. In this respect, the EAG agrees with the company that additional data collection might resolve or reduce some of the uncertainties associated to the remaining key issues.

Key Issue 3: Potential risk of study selection bias resulting in possible omission of relevant comparator studies

It is unclear to the EAG how "…a specific treatment that alone would not be representative of overall standard of care in UK clinical practice…" would not be informative, particularly given the lack of evidence for the clinical effectiveness for what the company are regarding as

standard of care i.e., immediate ablative procedures. Indeed, the only clinical effectiveness evidence that was submitted was the VHL Natural History study, where no immediate surgery or its sequalae were observed. This therefore remains a key issue.

Key Issue 4: Lack of clear link between clinical and cost-effectiveness sections in relation to sources of data for comparator

The company continue to misunderstand the purpose of the clinical effectiveness evidence and any ITC within that i.e., its importance regardless of its use in the economic analysis. Also, although the additional clarification regarding the various data sources is helpful, it is still unclear why the rates of events from the pre-surgery state could not be estimated from the Optum study. The company states: "…the outcomes calculated from this study (i.e., rates of events after 1st surgery) are not the same as the outcomes collected from the MK-6482-004 study (i.e. rates of events from pre-surgery)…", but it does not provide any further explanation. Therefore, this remains a key issue.

Key Issue 5: Limitations in the indirect treatment comparison hinder the assessment of the effectiveness of belzutifan compared to standard of care

The lack of pooling of all IPD remains a problem, particularly in the context of a method of adjusting for confounding that has been identified by the NICE DSU in TSD 18 as very likely to be unsatisfactory.

The company continue to not provide any objective evidence as to the most important variables to adjust for in the MAIC. They also make the spurious argument that there is unlikely to be confounding because the treatment effect estimated is too large: a very large treatment effect might suggest that no treatment effect is unlikely, but can say nothing about the size or direction of any confounding. Indeed, it is concerning that this very large treatment effect is estimated based on natural history data that does not include the sequalae of immediate surgery, which the company assume is standard of care, and which is at least party intended to reduce the risk of disease progression.

Limitation in the ITC therefore remain a key issue.

Key Issue 7: The comparator data might not be representative for the UK

Additional data collection from UK clinical practice might reduce the uncertainties associated to the comparator data, but, until then, this remains a key issue.

Key Issue 8: Data to inform effectiveness in the belzutifan arm (MK-6482-004 trial) are either immature or unavailable

The EAG would like to refer to the corresponding FAC response: Key Issue 8 refers to immaturity in the sense that there is uncertainty in the long-term extrapolations of treatment effectiveness in general. Again, this key issue might be resolved with additional data collection from UK clinical practice.

In addition, we would like to emphasize that statements such as "the alternatives [parametric distributions] explored by the EAG produce results that show belzutifan to remain consistently

cost-effective with the provided PAS" made by the company should be considered with extreme caution. Based on the data shown in Section 5.3.2.2 of the EAG report for example, the EAG would conclude that none of the parametric distributions should be considered reliable enough to support such statements.

Key Issue 9: There is uncertainty in the derivation of the transition probabilities in the standard of care arm

The EAG would like to thank the company for the additional clarification and scenarios. As above, there remains the issue that the data were not collected on the decision problem population.

Key Issue 10: There is uncertainty in the implementation of time on treatment and treatment effect waning

The EAG would like to thank the company for the additional clarification and refer to the corresponding FAC response. This remains a key issue.

Key Issue 11: There is uncertainty in the derivation and implementation of health-related quality of life in the model

The EAG would like to thank the company for the additional clarification and refer to the corresponding FAC response. The mismatch between the decision problem population and evidence used to inform health-related quality of life in the model remains a key issue.

Regarding the immediate health-related quality of life benefit for the belzutifan arm, the EAG would also like to refer to the corresponding FAC response: The EAG considers that this issue might have been resolved by including time to treatment response in the model and by linking the objective response level to time to response to calculate utility values in the pre-surgery, surgery, and event-free after surgery states. The EAG explored a scenario in which a fixed cut-off at the median time to treatment response was included to the QALY calculation (please refer to Scenario analyses set 6 in Section 6.1.2.6). The company also explored the impact of using fixed proportions at each response level in response to clarification question B21c.4 However, these scenarios were only exploratory, and their results should be interpreted with caution. The latter should be emphasised since the company seem to imply that implementing the median time-to-response into the QALY calculation would resolve this issue, whereas the EAG considers that this is not the case. In fact, we believe that in the current version of the model which includes the median time-to-response into the QALY calculation, the immediate effect of belzutifan is still present in the economic model. This can be seen for example by running the model for a time horizon of 0.9 years (approximately 11 months). This would be before the median time to response observed in the RCC cohort (11.11 months). If we remove immediate surgery from the SoC arm, it would be expected that before that time, no differences would be predicted between belzutifan and SoC. However, the model predicts [redacted] incremental QALYs in the RCC cohort in favour of belzutifan when the model is run under these settings.

Key Issue 12: Cost-effectiveness analyses should be based on subgroup-specific parameters.

The EAG would like to highlight that our point of view regarding key issue 12, as explained to the company in our responses to the FAC comments and during the TE call, has not changed. For convenience, these are summarized below.

Subgroup specific parameters

The EAG understands how the "subgroups" were defined. It is unclear however that the company is stating that the "marketing authorisation population should be assessed as a whole" when the whole submission is based on three different subgroups. Given that the evidence presented suggests that the clinical effectiveness or the disease severity may be different per subgroup, the EAG considers it to be more appropriate to use subgroup-specific parameters in this submission.

Severity modifier

We would like to express, again, that our method for calculating the QALY severity weighting is not methodologically flawed. However, if the company think otherwise, we would like to invite the company to formally challenge the methods by Versteegh et al. 2019 published in *PharmacoEconomics*.[1] We acknowledge that peer reviewed publications are not exempt from being flawed, but we also consider that in case it is, it should be formally proven.

We would also like to stress that we consider unacceptable the argument of using a tool based in the Netherlands, as opposed to an UK-based tool, to demerit our approach. As acknowledged by our UK colleagues in their recent publication, they also made available "*an R-Shiny online tool (https://shiny.york.ac.uk/shortfall)* ***inspired*** *by the iDBC platform of Versteegh et al. (https://imta.shinyapps.io/iDBC/)*".[2] Both tools therefore should be considered as appropriate to estimate QALY weighting. One of the main reasons why the QALY Shortfall Calculator tool has also been previously used in other EAG assessments conducted by KSR was that the iDBC tool was being updated to include more recent life tables and discounting, the latter to conform with the latest NICE methods.

In response to FAC comments we already mentioned that one of the differences with respect to the iDBC tool, but not the only one, is the use of Hernández Alava instead of Heijink for the UK value set. While the EAG agrees with the company that using Hernández Alava would be in line with the NICE reference-case, we would like to invite the company to conduct the severity analyses with both tools and check whether results are substantially different. As an example, we compared the company's PSA results obtained with the iDBC and the QALY Shortfall Calculator tools. Note that the only difference would be in the number of QALYs without the disease. These would be 18.02 with the iDBC tool and 18.15 with the QALY Shortfall Calculator tool using Hernandez-Alava (as reported in the company's model). The results for the RCC cohort are as follows:

- RCC cohort with iDBC tool: 42.7% for weight 1.2 and 57.3% for weight 1.7. Weighted ICER £[redacted].

- RCC cohort with QALY Shortfall Calculator tool: 42.2% for weight 1.2 and 57.8% for weight 1.7. Weighted ICER £[redacted].

Therefore, as it can be seen, the impact of using Hernández Alava instead of Heijink for the UK value set is minimal.

The EAG considers that the discussion on the severity of the condition should not be focused on whether one specific tool or another should have been used. We consider this irrelevant. The main EAG issue with the CS relates to the fact that the same severity weights were used for all three subgroups. The EAG believes this is incorrect.

In addition, the EAG considers that, if it is accepted that the estimated QALYs under standard of care are uncertain, then the estimated proportional and absolute shortfall should be considered uncertain as well. Regardless of which tool is used, the EAG considers that a fairer assessment of proportional and absolute shortfall is to account for this uncertainty, as there may be submissions that happen to have a deterministic QALY loss resulting in a QALY multiplier group that is not fitting to the entire sample. While it is understandable that this might cause some resistance with submissions where the deterministic QALY happens to correspond with the upper QALY multiplier, there will equally be cases conceivable where the opposite occurs. Both can be dealt with equally by accounting for the uncertainty in the QALY loss predicted by the model.

Furthermore, the fact that the updated NICE methods manual does not mention a particular methodology, in this case the severity-adjusted probability of being cost effective, should not prevent the EAG from using it. We would also like to clarify that, in the example provided by the company, interpreting the severity-adjusted probability of being cost effective as effectively equating to a QALY weight of 1.48 is indeed incorrect, since as the company correctly indicate, there are only three possible weights. The severity-adjusted probability of being cost effective should be interpreted in relation to the cost-effectiveness thresholds.

EAG's comments on company's updated cost-effectiveness analyses included in the Stakeholder Engagement Response From

New model version

A new version of the model was submitted after the Technical Engagement meeting. The company indicated that the following changes were made to the model:

- Specifications sheet:
  - Row 96: option to delay utility benefit from response achievement until median TTR – this was also incorporated in the updated company base-case.
  - Row 121: the PAS submitted to PASLU, described below, was added.
- DSA & PSA results sheets:
  - Results have been rerun with the updated base case assumptions & included discount. Both also included the weighted GB MA cohort results.
- Utility sheet:
  - Rows 15-18: the option to delay utility benefit from response achievement until median TTR (as per specifications sheet). The median TTR reported from MK-6482-004 for each cohort is used.

Based on these changes the ICERs were different, as shown below.

PAS price

The company submitted a patient access scheme (PAS) in the form of [redacted]
The cost effectiveness results presented by the company after the Technical Engagement meeting are based on the new version of the model, with the changes mentioned above, including thus the PAS for belzutifan.

The EAG would like to emphasise that the inclusion of a PAS price for belzutifan would not resolve the uncertainties associated to the key issues since these mostly relate to the lack of appropriate clinical data.

Updated results

Given the time constraints associated with this project, the EAG could not reproduce all tables with the updated cost effectiveness results in this document. Comments regarding the updated results are provided below:

- In general, statements regarding the cost effectiveness of belzutifan should be considered with extreme caution given the remaining uncertainties highlighted in the key issues above.
- We still disagree with the company in the way severity weighting was implemented: we consider that applying a weight of 1.7 for all cohorts is incorrect, regardless the

approach to estimating the proportional and absolute QALY shortfall (deterministic or probabilistic). For the pNET cohort a (deterministic) weight of 1.2 should be applied.

- With the most recent version of the model that the EAG has (NICE ID3932 STA Submission CEA v5.0 (CIC).xlsm) it was possible to replicate the ICERs presented by the company in response to Technical Engagement comments. Base-case ICERs (including belzutifan PAS and QALY weight) in the model: £[redacted], £[redacted], £[redacted] for the RCC, CNS Hb and pNET cohorts respectively.
- The updated base-case ICER for the pNET cohort, including the new PAS for belzutifan, and a deterministic QALY weight of 1.2 is £[redacted].
- The updated base-case ICERs for the RCC, CNS Hb and pNET cohorts, including the new PAS for belzutifan and the EAG probabilistic approach to the proportional and absolute QALY shortfall are £[redacted], £[redacted] and £[redacted], respectively.
- The impact of using the EAG's approach to severity weighting on the ICERs presented by the company in the scenario analyses should be similar to the one observed in the base-case analysis.

EAG's preliminary comments on Appendix 2 of the company's Stakeholder Engagement Response From

General comment

The EAG acknowledges that the company has presented the "Results from clinical expert elicitation & discussion". However, it appears that the experts' responses to the expert elicitation exercise have not been provided alongside this. This means that the company has just presented a summary of the discussion and the EAG cannot check if this summary indeed corresponds to actual statements made by the experts.

Key validation point 1

The experts' answers seem to confirm that the population in the MK trial is not the same as the label population. However, there is no actual mention of the MK trial – this confirms that surgical interventions would be harmful for those who would receive belzutifan. What it doesn't do is state that such surgical interventions should only be given in a world without belzutifan. In fact, it is confirmed under key validation point 2 that SoC without belzutifan would be surgery and that there would be no wait. This undermines the company's assertion that patients can wait until the outcome of belzutifan has been determined.

Key validation point 2

The company appears to have misunderstood the earlier EAG critique. The EAG did not suggest delaying immediate surgery in SoC, the issue being that belzutifan patients are also in need of immediate surgery, but they do not get it, since they wait until they respond to belzutifan. The question was more what happens to them until they respond to treatment, (assuming that they do respond). This undermines the company's assertion that one can wait until the outcome of belzutifan has been determined before undergoing surgery.

Key validation point 3

Some parts of this point are not clear. The experts indicated that patients would be classified as having stable disease but if patients do not respond, presumably it cannot be assumed that the tumour has stopped growing. The EAG agrees with the company when they said that non-response is accounted for in the transition probability from pre-surgery to surgery, however this is based on the MK trial, where we have a different population. It appears that there may have been a misunderstanding during the expert elicitation process as it makes no sense that patients can wait to see if belzutifan works if without it they need immediate harmful surgery and it makes even less sense that if they don't respond to belzutifan they somehow have been transformed into not needing surgery at all.

Key validation point 4

The EAG does not have anything new to add here but wishes to reiterate their previous point that more severe patients (as in the label population) should also have more surgeries also in the belzutifan arm. This continues to highlight the misapprehension that 'belzutifan eligible' patients are not the same as those who get SoC without belzutifan – we need to know what SoC is for the belzutifan eligible patients and if those patients wouldn't get immediate harmful surgery with belzutifan then they wouldn't with SoC.

Key validation point 5

The EAG does not have anything new to add here however, would point out again that the QoL study was conducted in a different population.

Key validation point 6

The EAG is not clear why the experts chose to refrain from making explicit comments relating to treatment of the metastatic population. Why is the same logic not applicable to the label population? It has only been studied in RCC patients following the MK trial, but conclusions have been generalised to other types of patients. This continues to highlight that belzutifan cannot be considered as a substitute for immediate surgery – it is to delay progression. The only way to reconcile no surgery with belzutifan is to assume that the belzutifan eligible population, which should be that of the DP, is not so severe that immediate surgery is required.

Final point

The EAG does not consider that the submitted material has any implications for changing assumptions in the CEA model.

References

[1] Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-adjusted probability of being cost effective. Pharmacoeconomics 2019; 37(9):1155-1163

[2] McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-adjusted life expectancy norms for the English population. Value Health 2023; 26(2):163-169