NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Contents:

The following documents are made available to consultees and commentators:

The **final scope** and **final stakeholder list** are available on the NICE website.

1. **Company submission** from Amgen UK

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. Roy Castle Lung Cancer Foundation
   b. British Thoracic Oncology Group

4. **Evidence Review Group report** prepared by Kleijnen Systematic Reviews (KSR) Ltd

5. **Evidence Review Group report – factual accuracy check**

6. **Technical engagement response** from Amgen UK

7. **Technical engagement response from consultees and commentators:**
   a. British Thoracic Oncology Group
   b. Boehringer Ingelheim

8. **Evidence Review Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews (KSR) Ltd

9. **Evidence Review Group addendum** prepared by Kleijnen Systematic Reviews (KSR) 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

Sotorasib for previously treated *KRAS p.G12C* mutated, locally advanced or metastatic non-small-cell lung cancer

[ID3780]

Document B

Company evidence submission

30th June 2021

File name	Version	Contains confidential information	Date
Sotorasib_NICE sub_Doc B	**v1.0**	**YES** **No**	**30th June 2021**

Contents

Tables and figures

Abbreviations

AIC	Akaike information criterion
ALK	Anaplastic lymphoma kinase
ASCO	American Society of Clinical oncology
ATT	average treatment effect of the treated
BIC	Bayesian information criterion
BICR	blinded independent central review
BNF	British National Formulary
BRAF	B-Raf Proto-oncogene
CI	confidence interval
CR	complete response
CRC	colorectal cancer
CSR	clinical study report
CT	computerised tomography
DCR	disease control rate
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal Growth Factor Receptor
eMIT	electronic market information tool
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Quality-of-life Questionnaire Core 30
EQ-5D-3L	EuroQol-3 Dimension
EQ-5D-5L	EuroQol-5 Dimension
ERK	extracellular signal-regulated kinase
ESMO	European Society of Medical oncology
ESS	effective sample size
FACT-G	Functional Assessment of Cancer Therapy Tool General form
FDA	Food and Drugs Administration
GDP	guanosine diphosphate
GTP	guanosine triphosphate
HER2	human epidermal growth factor receptor 2
HR	hazard ratio
HRQoL	health-related quality-of-life
IV	intravenous
KM	Kaplan-Meier
KRAS	Kirsten rat sarcoma viral oncogene homolog (protein)
KRAS	Kirsten rat sarcoma viral oncogene homolog (DNA)
KRASG12C	KRAS protein with a G12C amino acid substitution
KRAS p.G12C	KRAS DNA with a mutation
MAIC	matching-adjusted indirect comparison
MET	met proto-oncogene (hepatocyte growth factor receptor)
MHRA	Medicines and Healthcare products Regulator Agency
MIMS	Monthly Index of Medical Specialities
MRI	magnetic resonance imaging
NCCN	N**ational Comprehensive Cancer Network**
NE	not estimable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NR	not reported
NSCLC	non-small cell lung cancer
NTRK	neurotrophic tyrosine kinase
ORR	objective response rate

OS	overall survival
PAS	patient access scheme
PGIC	patient global impression of change
PGIS	patient global impression of severity
PD	progressive disease
PD-1	programmed cell death-1
PD-L1	programmed death-ligand 1
PD-L2	programmed death-ligand 2
PFS	progression free survival
PSSRU	Personal Social Services Research Unit
PSWA	propensity score weighting analysis
PR	partial response
PRO	patient-reported outcomes
PRO-CTCAE	patient-reported outcomes version of the common terminology criteria for adverse events
PS	performance status
QALY	quality-adjusted life year
QLQ-LC13	Quality-of-Life Questionnaire Lung Cancer Module
QOL	quality-of-life
QQ	quantile-quantile
QTc	corrected QT (interval)
RAF	RAF proto oncogene serine/threonine protein kinase
RAS	rat sarcoma viral oncogene homolog
RCT	randomised controlled trial
RECIST	response evaluation criteria in solid tumors
ROBINS-I	Risk Of Bias In Non-randomised Studies - of Interventions
ROS	proto-oncogene tyrosine-protein kinase
RTOR	Real-Time Oncology Review
SAQ	symptom assessment questionnaire
SD	standard deviation / stable disease
SLR	systematic literature review
TEAE	treatment-emergent adverse events
TK	tyrosine kinase
TRAE	treatment-related adverse events
UK	United Kingdom
US	United States
WHO	World health organisation

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

Disease overview

- Lung cancer is the third most common cancer and the leading cause of cancer deaths in the UK.
- Non-small cell lung cancer (NSCLC) is the predominant type, accounting for around 80-85% of cases. Most cases (around 75%) are diagnosed at an advanced stage, when the cancer has already spread locally (stage III) or to other parts of the body (metastatic, stage IV).
- A range of molecular and genetic tumour characteristics have been identified as key drivers of tumour proliferation and growth in NSCLC. This has transformed treatment options for patients with those specific oncogenic mutations for which targeted therapies have been developed (e.g., *ALK*, *BRAF, EGFR, NTRK* inhibitors).
- The *KRAS p.G12C* mutation is one of the most common oncogenic mutations, occurring in 13% of all NSCLC cases. Despite being recognised around 40 years ago, no targeted therapy for this mutation has been previously developed.
- Clinical outcomes and quality of life in patients with *KRAS p.G12C*-mutated NSCLC with currently available, non-targeted therapy are very poor, particularly for those who have progressed following first-line therapy.
- Current standard of care following first-line therapy is limited to cytotoxic chemotherapy with IV docetaxel monotherapy, or IV docetaxel in combination with oral nintedanib in some patients with adenocarcinoma.
    - In patients with KRAS-mutated NSCLC treated with docetaxel monotherapy median PFS is 2.9 months and median overall survival is 7.9 months.
    - There are no data for docetaxel in combination with nintedanib specifically in KRAS-mutated NSCLC patients.
    - Cytotoxic chemotherapy is associated with a range of off-target adverse events including neutropenia and febrile neutropenia, dyspnoea, fatigue, infection, anaemia, diarrhoea, stomatitis and nausea and vomiting that can lead to hospitalisation and severely compromise health-related quality of life.
- There is therefore an urgent need for more effective and tolerable targeted therapy for patients with *KRAS p.G12C*-mutated NSCLC.

Sotorasib and its position in the current clinical pathway

- Sotorasib is a highly innovative, targeted, oral monotherapy for adult patients with *KRAS p.G12C*-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated.
    - It is the first *KRAS p.G12C* inhibitor to be filed for regulatory approval.
    - It was granted an Innovation Passport under the recently introduced Innovative Licensing and Access Pathway.
    - It has been designated as a Promising Innovative Medicine under the Early Access to Medicines Scheme.
    - [redacted].
- As screening of *KRAS* mutations is now routinely included in the NHS England national service for cancer genomic testing, use of sotorasib does not require additional testing beyond the usual diagnostic work-up of NSCLC patients.
- Based on the NICE NSCLC clinical pathway and feedback from UK clinical experts, sotorasib would be used as an alternative to docetaxel monotherapy, or possibly nintedanib plus docetaxel in patients with adenocarcinoma, as a second or subsequent line therapy following prior immunotherapy-based treatment.
- As outcomes are worse with each successive line of therapy, sotorasib should be used in line with its licensed indication as early as possible in the treatment pathway

B.1.1. Decision problem

Lung cancer is the third most common cancer and the most common cause of cancer death in the UK, accounting for 13% of all new cancer cases and 21% of all cancer deaths [1]. This equates to around 48,000 new lung cancer cases and 35,000 deaths from lung cancer in the UK every year. NSCLC is by far the most common type of lung cancer, accounting for 80%–85% of lung cancers in the UK, with non-squamous cell carcinoma being the predominant sub-type [2].

Specific molecular and genetic tumour characteristics have been identified as key drivers of tumour proliferation and growth in NSCLC [3]. This has transformed treatment options for patients with those specific oncogenic mutations for which targeted therapies have been developed. However, there remain other oncogenic mutations that have been recognised as key drivers of tumour proliferation and growth for many years but for which no targeted therapy has previously been successfully developed. These include mutations in the Kristen rate sarcoma viral oncogene (*KRAS*), of which the most frequently occurring in NSCLC is the *KRAS p.G12C* mutation [4]. This mutation occurs in around 13% of NSCLC cases [5] and, until the advent of sotorasib, was so elusive to effective therapy it was considered to be "undruggable" [6]. Patients harbouring this mutation currently have a very poor prognosis, especially those who progress on first-line therapies for whom the only treatment options are non-targeted, cytotoxic regimens that are associated with off-target toxicities [7, 8]. There is therefore a significant unmet need for a highly targeted, effective, tolerable, and convenient treatment that improves clinical outcomes and quality of life for patients with *KRAS p.G12C*-mutated NSCLC.

Sotorasib is an oral, once daily therapy targeted specifically to inhibit the *KRAS p.G12C*-mutated protein in NSCLC. It is the first $KRAS^{G12C}$ inhibitor to be submitted for a marketing authorisation, and in February 2021 was granted an Innovation Passport under the recently introduced Innovative Licensing and Access Pathway [9]. Sotorasib has also been designated as a Promising Innovative Medicine under the Early Access to Medicines Scheme [10] and [redacted]. Sotorasib is therefore recognised as a highly innovative therapy that has the potential to meet the significant unmet needs of patients with *KRAS p.G12C*-mutated NSCLC who currently have no targeted therapy options available.

An application for UK marketing authorisation for sotorasib was submitted to MHRA in January 2021 with a proposed indication for use *as monotherapy for treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated*. A conditional licensing approval via the Project Orbis regulatory route in the UK is anticipated [redacted] based on data from a phase 2 single-arm trial (CodeBreaK100), with confirmatory results from a comparative phase 3 trial (CodeBreaK200) [11] anticipated within the next 2 years.

This submission covers the full proposed marketing authorisation for sotorasib. Given the urgent need for early access to sotorasib and the evidence that will be available initially and

following licensing, sotorasib may be considered a candidate for the Cancer Drugs Fund. The decision problem addressed in this submission is summarised in Table 1.

Table 1. 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 with previously treated *KRAS p.G12C* mutated, locally advanced or metastatic NSCLC	As per the anticipated licensed indication: *as monotherapy for the treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated* [12].	n/a
Intervention	Sotorasib	As per scope	n/a
Comparator(s)	**Non-squamous NSCLC:** • pemetrexed with carboplatin with or without pemetrexed maintenance • other platinum doublet chemotherapy with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • docetaxel monotherapy • atezolizumab • nivolumab (subject to ongoing CDF review) • pembrolizumab (PD-L1-expressing tumours) • best supportive care **Squamous NSCLC:** • gemcitabine with carboplatin or cisplatin • vinorelbine with cisplatin or carboplatin • docetaxel monotherapy • pembrolizumab (PD-L1-expressing tumours) • atezolizumab • nivolumab • best supportive care	**Primary comparator:** Docetaxel monotherapy **Secondary comparator:** Nintedanib + docetaxel	The NICE lung cancer pathway and international clinical guidelines (e.g. ESMO) recognise the increasing role of combination immunotherapy and chemotherapy in the first-line setting for NSCLC [3, 13]. Clinical expert opinion obtained from a UK advisory board held in February 2021 confirmed this is the case in UK clinical practice. The vast majority of patients will receive immunotherapy as a first line option and combination chemotherapy as part of first or second-line therapy. As clinical experts have confirmed that re-challenge with immunotherapy is not routine clinical practice [14], immunotherapy and combination chemotherapy are not relevant comparators for sotorasib. As *KRAS p.G12C* mutations rarely occur with other driver mutations (eg ALK, EGFR, ROS1 and BRAF co-occur with *KRAS p.G12C* at a rate of <1% [15]), therapies that target other driver mutations are not relevant comparators for sotorasib. Docetaxel monotherapy is recognised in the NICE lung cancer guideline (NG122) and pathway as a key second- and

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	People with *KRAS p.G12C* mutation and another driver mutation (including EGFR-TK, ALK or ROS1): Established clinical management without sotorasib, including: • atezolizumab combination (after EGFR-TK or ALK-targeted therapies) • lorlatinib (after ALK-targeted therapies) • brigatinib (after ALK-targeted therapies) • ceritinib (after ALK-targeted therapies) • osimertinib (EGFR T790M mutation-positive after EGFR-TK targeted therapies) • pemetrexed with carboplatin • platinum doublet chemotherapy • with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • nivolumab (subject to ongoing CDF review)		subsequent-line option in NSCLC across non-squamous and squamous disease and across PD-L1 expression and tumour proportion scores [13, 16]. The phase 3 CodeBreaK 200 study will compare sotorasib against docetaxel monotherapy, which has been agreed in NICE/EUnetHTA Scientific Advice sought by Amgen, and in the UK advisory board held in February 2021, to be a current standard of care treatment option for most patients with previously treated *KRAS p.G12C* mutated NSCLC with no other mutations for which a targeted product is available in the UK [17]. Docetaxel monotherapy is therefore the appropriate primary comparator for this appraisal. Patients with adenocarcinoma who are eligible for docetaxel may also be eligible for docetaxel in combination with nintedanib, in line with NICE TA 347. Clinical expert opinion obtained at the UK advisory board February 2021, indicates that use of docetaxel in combination with nintedanib is variable across different regions in the UK. On the basis of its variable use and only in adenocarcinoma, docetaxel in combination with nintedanib may be considered a secondary comparator.
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • time to treatment discontinuation • adverse effects of treatment	As per scope, with addition of duration of response.	n/a

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	• health-related quality of life.		
Special considerations including issues related to equity or equality	n/a	• In contrast to NSCLC patients with other oncogenic mutations, patients with advanced or metastatic *KRAS p.G12C*-mutated NSCLC who have failed prior therapy currently have no targeted therapy options, and very few other effective therapy options. Their prognosis is very poor, with OS significantly less than 2 years. • Sotorasib is a highly innovative, first in class therapy for *KRAS p.G12C*-mutated NSCLC. It provides an effective and tolerable targeted treatment option where previously there was none. It has been designated as a Promising Innovative Medicine via the UK Early Access to Medicines Scheme, and was granted an Innovation Passport under the Innovative Licensing and Access Pathway. UK orphan designation is pending. • Subject to approval, sotorasib is anticipated to be granted conditional marketing authorisation by the MHRA via the Project Orbis regulatory route on the basis of the results of the phase 2 CodeBreaK100 single arm trial. • As sotorasib is the first $KRAS^{G12C}$inhibitor to progress to licensing by any regulatory authority there is a lack of data specifically in patients with *KRAS p.G12C* mutated NSCLC for the relevant comparators, or any other agents. • Indirect comparative data using the most robust methods possible indicate that sotorasib is highly effective in achieving clinically meaningfully improvements in PFS and OS by >3 months compared with relevant comparators. • Based on these data, sotorasib provides a step change in therapy for patients with *KRAS p.G12C* mutated NSCLC and is highly likely to be cost effective under the NICE end of life policy. • Phase 3 data from the CodeBreaK200 RCT are anticipated within the next 2 years. • Sotorasib may therefore be a candidate for the Cancer Drugs Fund.	

B.1.2. Description of the technology being appraised

Sotorasib (LUMYKRAS™) is the first KRAS^{G12C} inhibitor to be submitted for marketing authorisation. It is a once daily oral therapy that, subject to approval, is anticipated to be licensed in the UK by MHRA in [redacted] *as monotherapy for the treatment of adult patients with KRAS p.G12C*-mutated *locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated* [12].

Sotorasib is an inhibitor of KRAS^{G12C}, a tumour-restricted, mutant-oncogenic form of the RAS GTPase, KRAS. It forms an irreversible, covalent bond with the unique cysteine of KRAS^{G12C}, locking the protein in an inactive state that prevents downstream signalling, inhibits cell growth, and promotes apoptosis only in *KRAS p.G12C* tumour cell lines with minimal detectable off-target activity [18]; no other wildtype or mutant protein or receptor has been identified to which sotorasib binds, nor has any effect been observed in cells without the *KRAS p.G12C* mutation. As the *KRAS p.G12C* mutation has been found only in tumour tissues, and not in normal tissue [19, 20], sotorasib has the potential to be highly tolerable compared with standard of care chemotherapy.

Sotorasib has been granted an Innovation Passport under the Innovative Licensing and Access Pathway [9], is designated as a Promising Innovative Medicine under the Early Access to Medicines Scheme in the UK [10], and [redacted]. It received accelerated approval by the US FDA 28th May 2021 under its Real-Time Oncology Review (RTOR), a pilot program that aims to explore a more efficient review process that ensures safe and effective treatments are made available to patients as early as possible [21]. With clinical evidence indicating meaningful improvements in PFS and OS, sotorasib is highly innovative and provides a step change in therapy for patients with *KRAS p.G12C*-mutated NSCLC who currently have no targeted therapy options available. A summary of sotorasib is provided in Table 2, and the draft Summary of Product Characteristics (SmPC) is provided in the reference pack (see Appendix C).

Table 2. Summary description of Sotorasib (LUMYKRAS™)

UK approved name and brand name	Sotorasib (LUMYKRAS™)
Mechanism of action	Sotorasib is a first-in-class irreversible small molecule inhibitor of $KRAS^{G12C}$. It binds specifically to the G12C mutant form of the KRAS protein, locking it in an inactive, guanosine diphosphate (GDP)-bound conformation. This inactivation of the KRAS protein prevents it from signalling to downstream effectors (including extracellular signal-regulated kinase, ERK) that control proliferation and mechanisms of cell survival [18].
Marketing authorisation / CE mark status	Sotorasib is currently being reviewed by the MHRA via the Project Orbis regulatory route. Conditional marketing authorisation is anticipated [redacted]. [redacted]
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Subject to approval, the licensed indication is anticipated to be for use as monotherapy for the treatment of adult patients with *KRAS p.G12C*-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated. See draft SmPC in Appendix C for full details of warnings and contraindications.
Method of administration and dosage	Sotorasib is administered orally at a dose of 960mg (given as 8 x 120mg tablets) once daily until disease progression or unacceptable toxicity.
Additional tests or investigations	As a highly targeted therapy, the presence of *KRAS p.G12C* mutation should be confirmed using a validated test prior to initiation of sotorasib. *KRAS p.G12C* is now included routinely in the national service for cancer genomic testing. Therefore, no additional tests beyond those used in the routine diagnostic work up and management of patients with NSCLC are required.
List price and average cost of a course of treatment	The anticipated list price for sotorasib is [redacted] per 30-day supply (pack of 240 tabs of 120mg).
Patient access scheme (if applicable)	[redacted] Based on the modelled drug utilisation and duration of therapy utilised in the economic evaluation, the anticipated average per patient cost of treatment of sotorasib is [redacted] (undiscounted).

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

B.1.3.1 Disease overview

B.1.3.1.1 Lung cancer

Lung cancer is the third most common cancer, accounting for over 40,000 new cancer cases (13% of all new cancer cases) in the UK. Around 75% of lung cancer cases are diagnosed at an advanced stage, when the cancer has already spread to lymph nodes and other organs in the chest (locally advanced disease; stage III) or to other parts of the body (metastatic disease; stage IV), which contributes to a disproportionately poor prognosis; despite being the third most common cancer, lung cancer is the most common cause of cancer death in the UK, accounting for 35,000 deaths (21% of all cancer deaths) each year [1].

Histologically, lung cancer can be classified into different types and subtypes. These include small cell lung cancer (SCLC); and non-small cell lung cancer (NSCLC), which includes squamous cell carcinoma, adenocarcinoma and large-cell carcinoma. The latter two, collectively, are referred to as 'non-squamous' lung cancer. NSCLC is by far the most common type, accounting for 80%–85% of lung cancers in the UK{Cancer Research UK, 2021 #2}, with non-squamous cell carcinoma the predominant sub-type of NSCLC (adenocarcinoma 66%; large cell 2%; squamous 23%; others 8%) [22].

In addition to histological characteristics, specific molecular and genetic tumour characteristics have been identified as key drivers of tumour proliferation and growth in NSCLC [3]. This has transformed treatment options for patients with those specific oncogenic mutations for which targeted therapies have been developed; NSCLC patients with anaplastic lymphoma kinase *(ALK)*, proto-oncogene tyrosine-protein kinase (*ROS),* epidermal growth factor receptor (*EGFR)*, B-raf (*BRAF)*, and neurotrophic tyrosine kinase *(NTRK)* mutations have several highly specific therapy options available (see section B.1.3.2.2). However, there remain other oncogenic mutations that have been recognised as key drivers of tumour proliferation and growth for many years but for which no targeted therapy has previously been successfully developed. These include mutations in the Kristen rate sarcoma viral oncogene (*KRAS*), of which the most frequently occurring in NSCLC is the *KRAS p.G12C* mutation [4].

B.1.3.1.2 KRAS p.G12C mutation in NSCLC

KRAS genes express proteins called guanosine triphosphate (GTP)ases, which regulate cellular proliferation, apoptosis and survival [23]. Mutations in *KRAS* genes can therefore disrupt the processes involved in the proliferation and survival of tumour cells. Of the *KRAS* mutations, an estimated 80% occur at codon 12. The *KRAS* gene with a mutation resulting in a G12C amino acid substitution (*KRAS p.G12C* mutation) in codon 12 is a single guanine to thymine substitution that results in a glycine to cysteine substitution at amino acid position 12. This structural change in the protein results in a defect in the association of GTPase-

activating proteins, which reduces the normal hydrolysis of GTP by the KRAS protein. The resulting accumulation of active, GTP-bound KRAS disrupts the process of apoptosis and promotes tumour proliferation and survival [24]. The *KRAS p.G12C* mutation is therefore an oncogenic driver and *KRAS* mutations are generally acknowledged as negative prognostic factors for treatment response and, hence, survival outcomes in patients with NSCLC and other solid tumours [4, 24, 25].

The *KRAS p.G12C* mutation occurs in around 13% of NSCLC cases [3, 26]. This would equate to an upper estimate of around 5,000 *KRAS p.G12C* mutated NSCLC cases in the UK each year, which is well below the prevalence threshold of 5 cases per 10,000 population used by the MHRA to define a disease as an orphan disease [27]. Despite NSCLC being a leading cause of cancer death, and *KRAS p.G12C* being one of the most common types of mutation in NSCLC, *KRAS p.G12C* mutated NSCLC is still a rare disease.

Several sources have demonstrated that the *KRAS p.G12C* mutation rarely occurs concomitantly with other oncogenic mutations for which targeted therapies have been developed [28-30]. This has been confirmed in a large Amgen study (Study 20200097) of *KRAS p.G12C* mutated NSCLC patients, conducted in the US Flatiron Health-Foundation Medicine Clinico-Genomic Database, who were diagnosed January 2011 to March 2019 [15]. Currently approved targeted therapies are therefore rarely an option for patients with the *KRAS p.G12C* mutation.

Current treatment of KRAS p.G12C-mutated NSCLC are non-targeted and associated with very poor outcomes

Current therapy options for patients with *KRAS p.G12C* mutated NSCLC are limited to non-targeted therapy (see section B.1.3.2.2), and outcomes for patients with advanced NSCLC who are not candidates for currently approved targeted therapy are poor, particularly in the second- or subsequent-line setting. As sotorasib is the first $KRAS^{G12C}$ inhibitor to have progressed in development to the point of regulatory submission, there is a general lack of randomised controlled trial (RCT) evidence on outcomes specifically in NSCLC patients harbouring the *KRAS p.G12C* mutation. However, RCTs of second-line non-targeted treatments (e.g. docetaxel, nintedanib plus docetaxel) in advanced NSCLC patients who received first-line platinum-based chemotherapy have demonstrated low objective response rates (ORRs; objective response = complete response + partial response) of 5% to 14%, poor progression-free survival (PFS) of 2.8 to 4.2 months and poor overall survival (OS) of 7.9 to 12.6 months) [31, 32].

Literature reviews conducted by Amgen to identify all available trials in NSCLC patients identified with *KRAS* mutations confirmed the lack of outcomes data specifically in patients with *KRAS p.G12C* mutated NSCLC (See Appendix D). Only one RCT (SELECT-1) published since 2014 that included sufficient information on patients with the *KRAS p.G12C* mutation [31] was identified. This RCT found that addition of selumetinib (a mitogen-activated protein kinase [MEK] inhibitor, not currently licensed in the UK for use in NSCLC) to docetaxel (taxane chemotherapy) as second-line therapy in patients with advanced *KRAS*-mutated NSCLC (including *G12C* and non-*G12C* mutations) did not significantly improve outcomes compared with docetaxel alone. In the docetaxel group, ORR was 13.7%,

median duration of response was 4.5 months, median PFS was 2.9 months and median OS was 7.9 months. There were no differences in PFS by *KRAS* mutation subtypes [31].

Given the lack of outcomes data from RCTs specifically in patients with *KRAS p.G12C* mutated NSCLC, observational, real-world evidence studies provide a valuable source of data on the natural history and prognosis for these patients. A recently published meta-analysis of observational studies found a statistically significant negative prognostic effect for mutant *KRAS* on OS (random effects model; HR 1.71; 95% CI 1.07, 2.84) and PFS (fixed effect model; HR1.18; 95% CI1.02, 1.36) [25]. Although this did not provide data specifically in *KRAS p.G12C*-mutated NSCLC, several other observational studies in Western (European, Australian and US) populations show that OS in patients with *KRAS p.G12C*-mutated NSCLC is similarly poor as that in patients with other *KRAS* mutations and in wild type disease who are not eligible for existing targeted therapies [33-35]. One study observed median OS in 352 patients with *KRAS p.G12C*-mutated NSCLC diagnosed between 2014 and 2018 in the US was 13 months, and in those receiving first or second-line immunotherapy the ORR was 24% [33]. An analysis of a prospective, multicentre, German (CRISP) registry for patients with stage IV NSCLC also found that 171 patients with *KRAS p.G12C*-mutated non-squamous tumours had similar survival outcomes to those with other KRAS-mutated tumours; following second-line treatment, median PFS was 4.4 months versus 4.8 months and median OS was 10.1 months versus 9.4 months for those with *KRAS p.G12C*-mutated and other *KRAS*-mutated tumours, respectively [35].

To provide further outcomes data specifically in *KRAS p.G12C*-mutated NSCLC, Amgen has conducted two large real-world evidence studies in US advanced NSCLC patients harbouring the *KRAS p.G12C* mutation (Study 20180277, n=416 and Study 20200097, n=743), and a further study in patients with advanced NSCLC patients with known KRAS status, including KRAS wild type, KRAS *p.G12C* mutated patients from Study 20200097, and other KRAS mutated patients (Study 20200132, n=7069) [5, 36, 37]. These data, which are highly consistent with those observed in Europe [35], confirm that survival is similarly very poor for patients with *KRAS p.G12C*-mutated NSCLC, other *KRAS*-mutated NSCLC, and other patients who are not candidates for existing targeted therapies, irrespective of KRAS status. Of note, PFS and OS are particularly poor as patients progression through lines of therapy, such that by third- or fourth-line therapy, median PFS is 3 months or less and median OS is less than 7 months (Table 3).

Table 3. Real-world Outcomes of Patients With Advanced NSCLC and *KRAS p.G12C*-mutated Advanced NSCLC by Line of Therapy

	***KRAS p.G12C*-mutated NSCLC**		**KRAS-mutated (non-*p.G12C*)**	**All NSCLC**
	Study 20180277 AACR project GENIE	Study 20200097 Flatiron Health Foundation	Study 20200132 Flatiron Health Foundation	
Median (95% CI) OS (months)				
First line	14.9 (12.2, 24.3)	12.0 (9.6, 15.3)	12.2 (10.5, 14.4)	12.9 (11.9, 14.2)
Second line	10.1 (7.1, 16.9)	9.5 (8.1, 13.1)	9.6 (7.7, 12.4)	10.2 (9.5, 11.3)
Third line	6.5 (5.0, NE)	6.7 (5.9, 10.7)	6.6 (5.0, 9.0)	7.9 (6.8, 8.8)
Fourth line	3.0 (2.2, NE)	5.9 (4.3, 12.9)	5.5 (3.9, 8.6)	7.4 (6.4, 8.6)
Median (95% CI) rwPFS (months)				
First line	6.1 (4.4, 9.3)	5.0 (4.4, 5.8)	5.6 (5.4, 6.0)	5.6 (5.3, 5.8)

	***KRAS p.G12C*-mutated NSCLC**		**KRAS-mutated (non-*p.G12C*)**	**All NSCLC**
	Study 20180277 AACR project GENIE	Study 20200097 Flatiron Health Foundation	Study 20200132 Flatiron Health Foundation	
Second line	3.2 (2.1, 5.3)	4.0 (2.8, 5.3)	4.1 (3.7, 4.4)	4.0 (3.7, 4.4)
Third line	2.3 (1.4, 4.1)	3.1 (2.4, 4.3)	3.5 (3.2, 4.0)	3.5 (3.1, 3.9)
Fourth line	1.8 (1.4, 15.0)	2.6 (2.1, 4.7)	3.1 (2.7, 3.5)	3.0 (2.7, 3.4)

KRAS p.G12C = *KRAS* gene with a mutation resulting in a G12C amino acid substitution; NE = not evaluable; NSCLC = non-small cell lung cancer; OS = overall survival; rwPFS = real-world progression-free survival
Retrospective Studies 20200097 and 20200132 were conducted using the United States Flatiron Health - Foundation Medicine Clinico-Genomic Database in 743 patients with *KRAS p.G12C*-mutated advanced NSCLC and 7069 patients with advanced NSCLC (ie, regardless of *KRAS p.G12C* mutation), respectively. Retrospective Study 20180277 was conducted using the American Association for Cancer Research Project Genomics Evidence Neoplasia Information Exchange database in 416 patients with *KRAS p.G12C*-mutated advanced NSCLC.[5, 36, 37]

Current non-targeted treatment for KRAS p.G12C-mutated NSCLC is associated with frequent toxicities and impaired quality of life

As NSCLC is diagnosed at an advanced stage in most patients, it is particularly important to consider the impact of treatment on symptom burden and health-related quality of life (HRQoL) [38, 39]. Patients whose disease progresses following first-line treatment are often already fatigued, anorexic and in pain. It is therefore important that further treatment does not worsen these symptoms or add to the symptom burden [40].

As there are no targeted therapies currently available for patients with advanced *KRAS p.G12C*-mutated NSCLC that has progressed on first-line therapy, treatment often consists of intravenous myelosuppressive, cytotoxic chemotherapy such as docetaxel, which as a non-targeted therapy is associated with a number of off-target toxicities, including neutropenia and febrile neutropenia, dyspnoea, fatigue, infection, anaemia, diarrhoea, stomatitis and nausea and vomiting that can lead to hospitalisation and severely compromise HRQoL [7, 8]. Approximately 50% of patients with previously treated *KRAS*-mutated advanced NSCLC experience symptom progression one month after initiating treatment with docetaxel [31].

Febrile neutropenia (fever in the presence of neutropenia) is a particularly concerning complication of myelosuppressive chemotherapy such as docetaxel [41]. Patients who develop febrile neutropenia are more susceptible to potentially fatal infections, such as severe sepsis or septic shock, that may necessitate chemotherapy dose modification, delays, or discontinuation, jeopardizing treatment response and clinical outcomes and compromising quality of life [42, 43].

B.1.3.1.3 Summary: there is an urgent need for targeted therapy with sotorasib in KRAS p.G12C-mutated NSCLC

In contrast to other driver mutations, for which there are several targeted treatment options available, there are no targeted treatment options available for patients with NSCLC harbouring the *KRAS p.G12C* mutation. With the earlier use of immunotherapy in the NSCLC pathway, treatment options for patients with *KRAS p.G12C*-mutated NSCLC who

progress to second and subsequent line therapy are increasingly limited to cytotoxic chemotherapy that is minimally effective and associated with a number of severe, off-target toxicities. There is, therefore, a particularly high unmet need for effective and tolerable therapies for *KRAS p.G12C*-mutated NSCLC, in particular targeted treatments, that can improve outcomes without compromising HRQoL for patients who have progressed to second or subsequent line therapy.

With the advent of sotorasib as a specifically targeted therapy for the *KRAS p.G12C* mutation, *KRAS p.G12C* is now an important predictive biomarker for one of the most common types of NSCLC.

As a specifically targeted oral therapy for the *KRAS p.G12C* mutation, sotorasib offers the potential for improved clinical outcomes, tolerability, convenience and HRQoL compared with existing highly toxic, non-targeted, standard of care therapy. Its early use in line with its anticipated licensed indication as a second or subsequent line therapy can delay the need to use less effective, intravenous cytotoxic chemotherapy.

There is therefore an urgent need for the earliest possible access to sotorasib in this patient population.

B.1.3.2 NSCLC clinical care pathway

B.1.3.2.1 Diagnosis of NSCLC and KRAS p.G12C-mutated NSCLC

The importance of molecular biomarker testing has increased with the development of a range of highly targeted therapies for NSCLC. As the vast majority of oncogene-driven lung cancers are adenocarcinomas, the latest ESMO and NCCN guidelines call for all patients with advanced, possible, probable or definite, adenocarcinoma to be tested for oncogenic drivers [3, 44].

NICE Guideline 122 – *Lung cancer: diagnosis and management* and the NICE Lung cancer pathway [13, 16] outline the approach to diagnosis and staging of lung cancer, emphasising the need for rapid assessment of suspected cases, and appropriate use of sputum cytology, bronchoscopy, CT and MRI scans and X-rays. Regarding the use of biomarkers the guidance is limited, noting only the need to ensure that biopsy samples are adequate (without unacceptable risk to the person) to permit pathological diagnosis, including tumour subtyping and assessment of predictive markers [16]

The *National Optimal Lung Cancer Pathway* produced by NHS England [45] provides the diagnostic standards of care for lung cancer, detailing the sequence and timing of investigations. This specifies that, where a pathological diagnosis could influence treatment, biopsy results providing the subtype of cancer should be available within 3 working days, and results of testing for molecular biomarkers should be available within 10 working days from the point of referral [45].

The latest National Lung Cancer Audit *Spotlight report on molecular testing in advanced lung cancer*, published January 2020, states the current standard of care is for all patients with

advanced lung adenocarcinoma (and other non-squamous subtypes) to have molecular testing at diagnosis for at least *EGFR* mutations, *ALK* and *ROS-1* rearrangements, and expression of PD-L1 on neoplastic cells. Patients with squamous cell lung cancer should have testing for PD-L1 expression [22]. However, with the advent of other targeted therapies, such as entrectinib targeting NTRK gene fusions [46] and sotorasib targeting $KRAS^{G12C}$, as well as others in development to target *MET* mutations and *HER2* alterations [3], there is a need to expand routine molecular testing in NSCLC.

The UK Government's ambitions are to create the most advanced genomic healthcare system in the world, incorporating the latest genomics advances into routine healthcare to improve the diagnosis, stratification and treatment of illness [47]. The NHS Long Term Plan, published in 2019, outlined the UK Government's strategy to expand the molecular diagnostics and to offer genomic testing routinely to all people with cancer for whom it would be of clinical benefit; from 2020/21 the NHS will offer more extensive genomic testing [47, 48]. To achieve this aim, NHS England has established a national service for cancer genomic testing to replace all local testing. This involves the setting up 7 genomic laboratory hubs across England to do testing by next generation sequencing (NGS) when locally advanced or metastatic solid tumours are first diagnosed [46]. KRAS mutation testing is confirmed and listed on the National Test Directory [49], meaning that no additional tests beyond those that will be used in the routine diagnostic work up and management of patients with NSCLC are anticipated to be required to determine eligibility for sotorasib.

B.1.3.2.2 Current treatment guidelines

As sotorasib is the first $KRAS^{G12C}$ inhibitor to have been submitted for regulatory approval, there is currently no defined clinical pathway for people with *KRAS p.G12C*-mutated NSCLC. Current treatment therefore follows the NSCLC clinical pathway for people without actionable mutations.

NICE Guideline 122 and the associated NSCLC pathway [13, 16], and ESMO [3], NCCN [44], ASCO [50] guidelines, all updated in 2019/20, emphasise the role of immunotherapy with checkpoint inhibitors, either alone or in combination with platinum-containing chemotherapy, as first line treatment for patients with advanced NSCLC without actionable mutations. Although the NICE lung cancer pathway still includes the option of first-line cytotoxic chemotherapy in these patients, the increasing evidence base in support of front-line immunotherapy has brought about a shift away from front-line treatment with cytotoxic chemotherapy alone [50]. NICE guidance therefore includes pembrolizumab in combination with platinum-based chemotherapy as a first-line treatment option, as well as pembrolizumab monotherapy, in patients with either non-squamous or squamous disease [13]. In patients whose disease progresses following first line pembrolizumab in combination with platinum-based chemotherapy, second line treatment options include docetaxel monotherapy, or (for those with adenocarcinoma) docetaxel in combination with nintedanib. In patients whose disease progresses following first line pembrolizumab alone, second-line therapy is with platinum-based chemotherapy, followed by docetaxel monotherapy, or (for some patients with adenocarcinoma) nintedanib in combination with [13].

B.1.3.3 Proposed positioning of sotorasib within the NSCLC clinical pathway

In line with the current treatment guidelines, clinical experts at a recent advisory board held by Amgen February 2021 confirmed that an increasing majority of patients in the UK with NSCLC without current actionable mutations now receive anti-PD-1/PD-L1 immunotherapy alone or in combination with platinum-based chemotherapy as first-line therapy; few patients receive immunotherapy as second-line therapy and re-challenge with immunotherapy is not routine practice [14]. They confirmed that sotorasib would be used as an alternative to docetaxel monotherapy or (for some patients with adenocarcinoma) nintedanib in combination with docetaxel, following initial therapy with immunotherapy and/or platinum-based chemotherapy.

The proposed positioning of sotorasib, in line with its anticipated licensed indication, is therefore as a second or subsequent line therapy following prior treatment with platinum-based chemotherapy and/or anti-PD-1/PD-L1 immunotherapy.

This positioning is aligned with the pivotal clinical trial data supporting its anticipated conditional regulatory approval (see section B.2.3), and confirms that the only relevant comparators for sotorasib are:

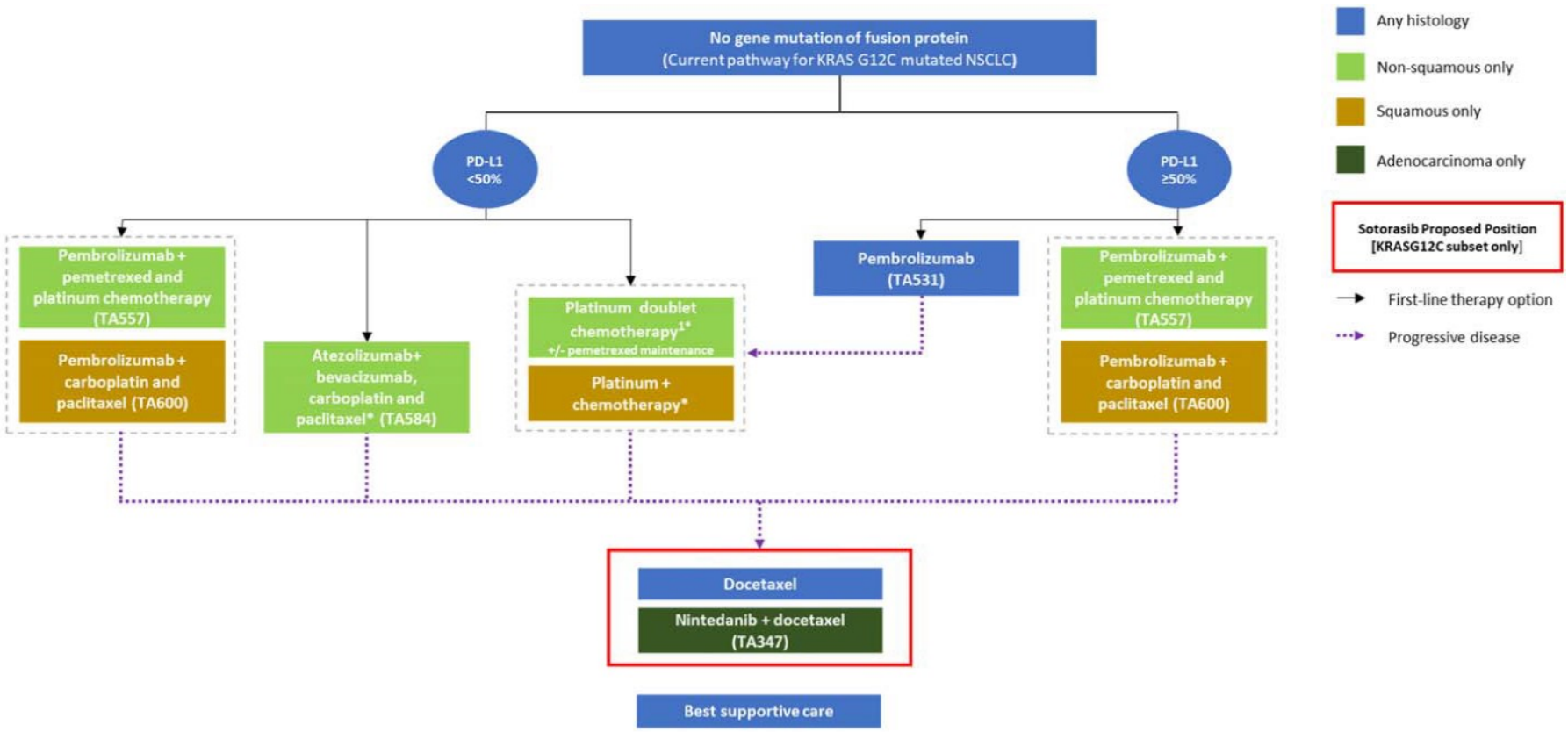
- docetaxel monotherapy, or
- nintedanib in combination with docetaxel (for some patients with adenocarcinoma) (see Figure 1).

As docetaxel monotherapy is recognised in the NICE lung cancer guideline (NG122) and pathway as a key second- and subsequent-line option in NSCLC across non-squamous and squamous disease and across PD-L1 expression and tumour proportion scores [13, 16], this is the appropriate primary comparator for this appraisal. As nintedanib in combination with docetaxel is only recommended for use in patients with adenocarcinoma [51], and clinical expert opinion obtained at the UK advisory board, February 2021, indicates its use is highly variable across different regions in the UK [14], nintedanib in combination with docetaxel is considered a secondary comparator for this appraisal.

B.1.4. Equality considerations

No specific equality considerations are anticipated.

Figure 1. Current treatment pathway for *KRAS p.G12C*-mutated NSCLC, and proposed positioning of sotorasib



Adapted from NICE Lung cancer pathway, available at: https://pathways.nice.org.uk/pathways/lung-cancer (accessed 21 Feb 2021)
1. When using Pemetrexed+Cisplatin as first-line treatment see TA181)
*This combination/some of these combinations of drugs do not have a UK marketing authorisation for 1 or more indications

B.2. Clinical effectiveness

Summary of Clinical Effectiveness

- The CodeBreaK100 phase 2, single-arm trial provides the relevant efficacy and safety data for sotorasib in patients with KRAS G12C-mutated NSCLC.
    - CodeBreaK100 is accepted as sufficient for conditional regulatory approval of sotorasib, and is reflective of and generalisable to patients in UK clinical practice
    - Evidence of anti-tumour activity from single-arm trials has previously been sufficient for NICE to make positive recommendations for several other targeted therapies for NSCLC, within the Cancer Drugs Fund and in routine commissioning.
- Evidence from CodeBreaK100 indicates that sotorasib is highly effective when used in line with its full anticipated licensed indication as a second- or subsequent line therapy.
    - Objective response rate was 37.1%, which UK clinical experts considered to be much better than they would expect to see with existing standard of care IV docetaxel monotherapy or IV docetaxel plus oral nintedanib.
    - Response to sotorasib was rapid (median 1.35 months to objective response) and durable (median duration of response 10 months).
    - Median PFS was 6.8 months, and median OS was 12.5 months, which are somewhat greater than the 2.9 months PFS and 7.9 months OS observed with the primary comparator docetaxel monotherapy.
- Safety data from CodeBreaK100 indicates that sotorasib is generally well tolerated, with a very manageable adverse event profile
    - The majority of treatment-related adverse events (TRAEs) with sotorasib were mild to moderate; 9 subjects (7%) discontinued sotorasib due to TRAEs.
    - UK clinical experts agreed the safety and tolerability of sotorasib appears to be superior to that seen with docetaxel or nintedanib plus docetaxel.
- There is a lack of trial data for the relevant comparators specifically in KRAS G12C-mutated NSCLC; however, indirect treatment comparisons using the most robust data sources and methods possible provide plausible early evidence of clinically meaningful improvements in survival outcomes with sotorasib compared with current standard of care, non-targeted therapy.
    - A primary MAIC analysis indicates that sotorasib provides a median gain in PFS of [redacted] months [redacted] and gain in OS of at least [redacted] months [redacted], hazard ratio [redacted]) compared with docetaxel monotherapy.
    - Based on an extrapolated analysis implemented in the economic model, sotorasib provides a plausible gain in mean PFS of [redacted] and gain in mean OS of [redacted] compared with the secondary comparator nintedanib plus docetaxel.
- Sotorasib is highly innovative and available evidence indicates clearly that it can address the significant unmet need for a targeted, more effective, tolerable, and convenient treatment that improves clinical outcomes and preserves quality of life for patients with *KRAS G12C*-mutated NSCLC. Sotorasib should therefore be made available as early as possible.
- Based on the available data, sotorasib fulfils the criteria for consideration under the NICE End of Life policy.
- With confirmatory phase 3 trial data anticipated to be available in the next 2 years, sotorasib may be a candidate for use via the Cancer Drugs Fund.

B.2.1. Identification and selection of relevant studies

See appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

B.2.2. List of relevant clinical effectiveness evidence

Sotorasib is anticipated to receive conditional marketing authorisation for the treatment of advanced *KRAS p.G12C*-mutated NSCLC primarily based on results from the phase 2 portion of the single-arm CodeBreak100 trial [52]. This trial provides the primary clinical trial evidence for sotorasib in this submission (Table 4).

The CodeBreaK100 trial is an ongoing phase 1/2 study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of sotorasib in subjects with *KRAS p.G12C*-mutated NSCLC, colorectal cancer, and other solid tumour types [52, 53]. Only data relating to the phase 2 portion of the trial in the treatment of NSCLC are presented. A primary analysis of efficacy, safety and patient-reported outcomes (PROs) data was conducted in September 2020. An updated analysis of efficacy and safety data for regulatory purposes was conducted 01 December 2020[54, 55], with a further updated analysis conducted 15 March 2021 and published 04 June 2021 [56]. As the results of the 01 December 2020 and 15 March 2021 analyses are highly consistent (see comparisons provided in Table 8 and in Appendix E) but patient-level data required for the indirect treatment comparisons are currently available only from the 01 December 2020 data cut, this submission focuses on the 01 December 2020 efficacy and safety analyses [54, 55], with the latest available PROs data from September 2020 [57].

A phase 3 randomised controlled trial (RCT) comparing sotorasib against standard of care docetaxel in patients with NSCLC is ongoing [11], with first results anticipated in 2022 (see section B.2.11).

Table 4 Clinical effectiveness evidence: CodeBreaK100

Study	CodeBreaK100 (NCT03600883)				
Study design	Single-arm, phase 2 trial conducted in 47 centres in the United States, Australia, Austria, Belgium, Canada, France, Germany, Japan, South Korea, and Switzerland.				
Population	Adults with confirmed *KRAS p.G12C*-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per Response Evaluation Criteria In Solid Tumors (RECIST) 1.1 criteria, and had ECOG performance status of 0 or 1.				
Intervention(s)	Sotorasib 960mg administered orally once per day without interruption (i.e., no planned off-treatment days) until disease progression, intolerance, withdrawal of consent or death.				
Comparator(s)	None.				
Indicate if trial supports application for marketing authorisation	Yes	✓	Indicate if trial used in the economic model	Yes	✓
	No			No	
Rationale for use/non-use in the model	CodeBreaK100 provides the primary clinical efficacy and safety data for sotorasib at the time of this submission.				

Reported outcomes specified in the decision problem (bold indicates outcome used in economic model)	Primary endpoint: Objective response rate assessed by blinded independent central review Secondary endpoints: Duration of response PFS OS Incidence and severity of adverse events Exploratory endpoints: Health-related quality of life (EORTC QLQ-C30, QLQ LC13, EQ-5D-5L)
All other reported outcomes	Disease control Time to response 6- and 12-month PFS 12-month OS Patient-reported outcomes (NSCLC SAQ, FACT-G, PRO-CTCAE) PK parameters and biomarkers (not further discussed in this submission)
EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality-of-life Questionnaire Core 30; FACT-G, Functional Assessment Of Cancer Therapy - General; NSCLC, non-small cell lung cancer; PRO-CTCAE, patient-reported outcomes version of the common terminology criteria for adverse events; QLQ LC13, Quality-Of-Life Questionnaire Lung Cancer Module; SAQ, symptom assessment questionnaire References: [52, 57, 58]	

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

B.2.3.1 Trial design

CodeBreaK100 is an ongoing phase 1/2 study of sotorasib. The phase 2 portion is a multicentre, nonrandomized, open-label study designed to evaluate efficacy and safety/tolerability of sotorasib as monotherapy in subjects with *KRAS p.G12C*-mutated advanced solid tumours (NSCLC, colorectal cancer, and other tumours) [52]. This single-arm design was considered appropriate to balance the need for sufficiently robust evidence generation against the need to ensure timely availability of sotorasib in patients with profound unmet clinical needs (see section B.1.3.1.3). Evidence of anti-tumour activity from single-arm trials has been sufficient for regulatory approval of several other targeted therapies for NSCLC, and for NICE to make positive recommendations for their use within the Cancer Drugs Fund [46, 59, 60] and in routine commissioning [61, 62].

The dose (and schedule) administered in phase 2 was the recommended dose of 960 mg once daily, determined in the phase 1 portion of the study as providing the highest ORR and the deepest average observed responses, with adverse events consistent with those observed at lower doses [52, 53].

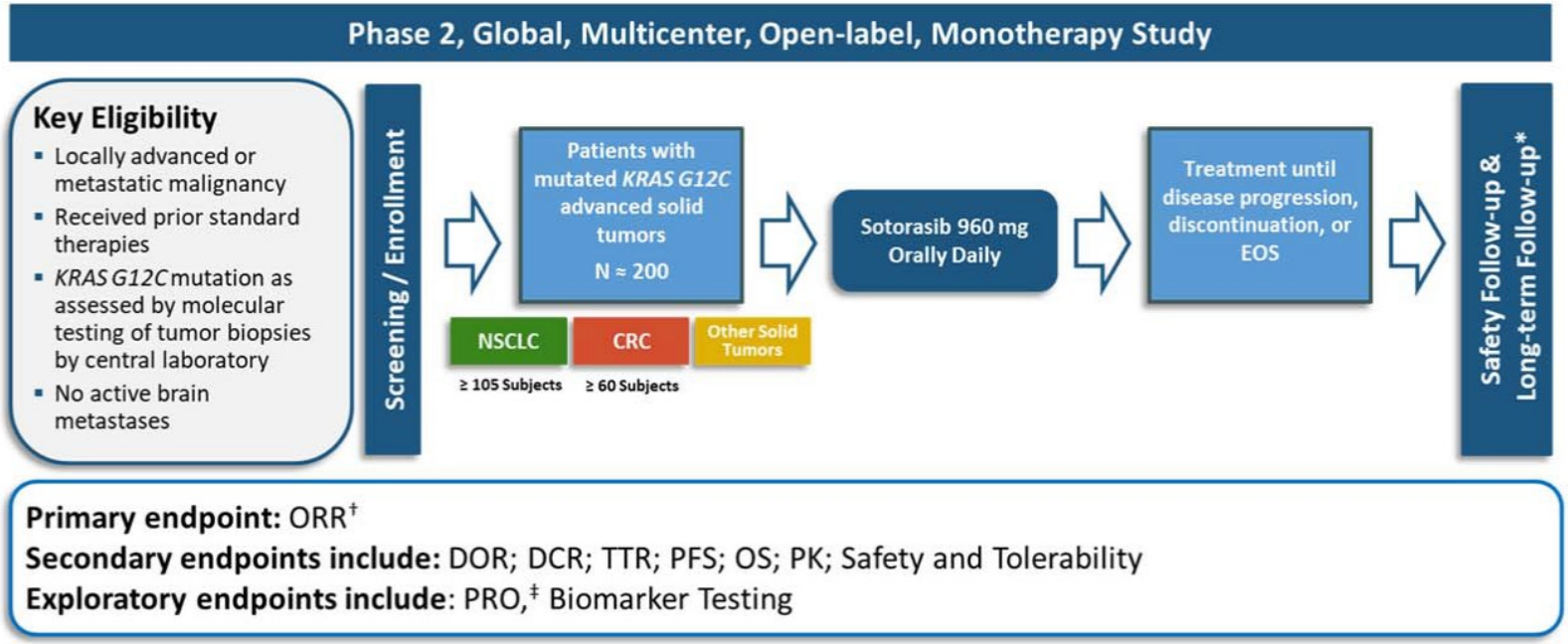
Approximately 250 subjects (with at least 105 with NSCLC) were planned to be enrolled. Tumour response was measured by contrast enhanced CT/MRI and assessed per RECIST 1.1 criteria by an independent radiological central laboratory. To demonstrate durability of ORR, the phase 2 primary analysis was to occur approximately 8.5 months after 105

evaluable subjects with NSCLC or 60 with CRC had enrolled in the phase-2 portion of the study.

Interim safety reviews were conducted after 30, 50, 70, and 100 subjects had been enrolled and treated with sotorasib for at least 21 days. Daily treatment with sotorasib in phase 2 was to continue without interruption (i.e., no planned off-treatment days) until disease progression, treatment intolerance, withdrawal of consent, or death. Subjects were to have a safety follow-up visit 30 days after the last dose of sotorasib, before any new anticancer treatment was started. After the safety follow-up visit, subjects were to be followed long-term for health condition, disease status, and subsequent anticancer treatment every 12 weeks for up to 3 years after the last subject was enrolled or until withdrawal of consent, loss to follow-up, or subject death, whichever occurred first. Study duration is therefore approximately 4 years (28-day screening, 6 to 12 months of treatment, and 3 years of long-term follow-up) for each subject. A schematic of the study is provided in Figure 2.

.

Figure 2. CodeBreaK100 study schematic



*30 (+7) days after end of treatment for safety follow-up; every 12 weeks for 3 years for long-term follow-up of OS.
†per RECIST 1.1; confirmatory assessment 4 weeks after first detection of response.
‡on disease-related symptoms and HRQOL.

CRC, colorectal cancer; **DCR,** disease control rate; **DOR,** duration of response; **EOS,** end of study; **HRQOL,** health-related quality of life; **KRAS,** Kirsten rat sarcoma viral oncogene homolog; **NSCLC,** non-small cell lung cancer; **ORR,** objective response rate; **OS,** overall survival; **PFS,** progression-free survival; **PK,** pharmacokinetic; **PRO,** patient-reported outcome; **RECIST,** response evaluation criteria in solid tumors; **TTR,** time to response.

Hong DS, et al. *N Engl J Med*. 2020;383:1207-1217. Supplementary material: Protocol.

B.2.3.2 Eligibility criteria

Adult subjects (≥ 18 years of age) with advanced solid tumours were eligible for the study. Enrolment was restricted to subjects with *KRAS p.G12C*-mutated solid tumours as assessed by molecular testing and confirmed by central testing (therascreen® KRAS RGQ PCR from Qiagen) prior to enrolment. Subjects must have had ≥ 1 prior line of systemic anticancer therapy, progressed on prior line of therapy, measurable disease per RECIST 1.1 criteria; Eastern Cooperative Oncology Group (ECOG) performance status of ≤ 1; no active brain metastases (but patients with treated and stable brain metastases could enrol); QTc ≤ 470 msec (based on average of screening triplicates); ability to take oral medications; and adequate haematological, renal, hepatic, and coagulation laboratory assessments. A full list of inclusion and exclusion criteria is provided in Appendix D.

B.2.3.3 Endpoints

The primary and secondary efficacy endpoints used in CodeBreaK100 are summarised in Table 5. The primary endpoint to support marketing approval is the objective response rate (ORR; ORR = complete response + partial response). As spontaneous regression of NSCLC tumours in the absence of treatment is a rare phenomenon, ORR is a compelling measure of antitumor activity showing the proportion of subjects with a response [63]. Analyses exploring the association between ORR and survival have demonstrated patient-level and study-level associations between ORR, PFS, and OS [64, 65]. ORR is therefore reasonably likely to predict clinical benefit, and it is of note that response rate was the primary endpoint in studies for several other targeted therapies that received positive recommendations by NICE for use in the Cancer Drugs Fund [46, 59, 60] and in routine commissioning [61, 62].

To reduce the risk of bias in this single-arm study, tumour response assessment was conducted in a blinded fashion by an independent, external radiologic central laboratory using RECIST 1.1 criteria [18, 66]. A pre-specified benchmark of efficacy for sotorasib was established based on the ORR observed for ramucirumab plus docetaxel, one of the most effective existing licensed therapies that could theoretically be used in the patient population proposed to be treated with sotorasib (but which is not recommended for use by NICE [67]. The ORR observed with ramucirumab plus docetaxel was 23% (95% CI: 20, 26) in 1253 patients with NSCLC with disease progression on or after one platinum-based therapy for locally advanced or metastatic disease [68]. In CodeBreaK100, if the lower bound of the 95% CI for ORR excluded this prespecified benchmark of 23%, this was taken to confirm the clinical significance of the ORR with sotorasib.

Since the clinical significance of ORR is also assessed by its duration, duration of response was evaluated as a key secondary endpoint. Other supportive secondary efficacy endpoints include time to response, disease control rate (percentage of patients with complete response, partial response or stable disease, which is another strong predictor of clinical benefit since not all patients with advanced NSCLC have tumour shrinkage after cancer therapies [69, 70]), PFS, and OS. Health-related quality of life and other patient-reported outcomes were assessed as exploratory analyses [52, 57].

Table 5. Efficacy endpoints in CodeBreaK100

Efficacy Endpoint	Definition	Statistical Test
Primary		
ORR	Proportion of subjects with best overall response of complete response or partial response as assessed by RECIST 1.1 Response was assessed by BICR. Complete response and partial response required confirmatory CT or MRI repeat assessment at least 4 weeks after the first detection of response	Percentage of subjects with an objective response summarized with Clopper-Pearson exact 95% CI. Clinical relevance of ORR was determined by the lower bound of the 95% CI excluding a prespecified benchmark of 23%
Secondary		
Duration of response	Time from first evidence of partial response or complete response to disease progression or death due to any cause for subjects who achieved a confirmed best overall response of partial response or complete response	Kaplan-Meier quartiles and rates for response select durations (responders only)
Disease control rate	Proportion of subjects with best overall response of complete response, partial response, or stable disease $\geq$ 5 weeks	Summarised as for ORR
Time to response	Time from first dose of sotorasib until the first evidence of partial response or complete response for subjects who achieved a confirmed best overall response of partial response or complete response	Summarised descriptively (responders only)
PFS	Time from first dose of sotorasib until disease progression or death from any cause	Summarised with Kaplan-Meier curves, quartiles, and rates at 6 and 12 months
OS	Time from first dose of sotorasib until death from any cause	Summarised as for PFS
Exploratory		
PROs and HRQoL	EORTC QLQ-C30; EORT QLQ-LC13; PGIS; PGIC; FACT-G (GP5); NSCLC-SAQ; PRO-CTCAE; EQ-5D-5L	Summarised descriptively. Changes from baseline over time tested using mixed effects model for repeated measures (MMRM) model

95% CI, 95% confidence interval; BICR, blinded independent central review; CT, computerised tomography; EORTC, European Organization for Research and Treatment of Cancer; EORTC QLQ-C30, EORTC Quality of life questionnaire 30-item core module; EORTC QLQ-LC13, EORTC Quality of life questionnaire lung cancer module; EQ-5D-5L, EuroQol-5 dimension-5 level; FACT-G, Functional Assessment of Cancer Therapy Tool General form; GP5, the single item "I am bothered by side effects of treatment," rated on a 5-point Likert scale and part of the FACT-G; HRQoL, Health Related Quality of Life; MRI, magnetic resonance imaging; NSCLC, Non-small cell lung cancer; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PGIC, patient global impression of change; PGIS, patient global impression of severity; PROs, patient-reported outcomes; PRO-CTCAE, patient-reported outcome version of the Common Terminology Criteria for Adverse events; SAQ, symptom assessment questionnaire. References: [52, 57, 58]

B.2.3.4 Baseline demographics and disease characteristics

Of the 126 enrolled patients, 81.7% were white, 15.1% were Asian, and 1.6% were black. There was an equal balance of males and females, and the median (range) age was 63.5 (37, 80) years. Nearly all subjects had metastatic disease (96.0%), and as per the protocol, no patients had active brain metastases at baseline. Most had an ECOG performance status of 1 (69.8%). Subjects received a median of 2 prior lines of anticancer therapy. Most subjects were previously treated with platinum-based chemotherapy (113 subjects [89.7%), anti-PD-1/PD-L1 immunotherapy (115 subjects [91.3%]), or both anti PD-1/PD-L1 immunotherapy and platinum-based chemotherapy, either in combination or across lines (102 subjects [81.0%]) [52, 54].

Table 6. Baseline characteristics of subjects in CodeBreak100, phase 2

	Sotorasib 960mg (N = 126)
Sex - n (%)	
Male	63 (50.0)
Female	63 (50.0)
Race - n (%)	
Asian	19 (15.1)
Black or African American	2 (1.6)
White	103 (81.7)
Other	2 (1.6)
Age (years)	
Mean	62.9
SD	9.3
Median	63.5
Min, Max	37, 80
Smoking history - n (%)[a]	
Never	6 (4.8)
Current or former	117 (92.9)
NSCLC stage – n (%)	
III	5 (4.0)
IV	121 (96.0)
Metastases – n (%)	
Brain (non-active)	26 (20.6)
Liver	26 (20.6)
NSCLC histology – n (%)	
Non-squamous	125 (99.2)
adenocarcinoma	120 (95.2)
Squamous	1 (0.8)
ECOG performance status – n (%)	
0	38 (30.2)

1	88 (69.8)

Prior lines of systemic anticancer therapy – n (%)

	Sotorasib 960mg (N = 126)
1	54 (42.9)
2	44 (34.9)
3	28 (22.2)
Types of prior systemic anticancer therapy[b] – n (%)	
Platinum-based chemotherapy	113 (89.7)
PD-1 or PD-L1 inhibitors	115 (91.3)
Platinum-based chemotherapy and PD1/L1 inhibitors	102 (81.0)

ECOG, Eastern Cooperative Oncology Group; N, Number of subjects in the analysis set; n, Number of subjects in the corresponding category; PD-1, programmed cell death protein 1; PD-L1, programmed death-ligand 1. SD = Standard Deviation.
[a] smoking status missing for 3 patients
[b] prior systemic anticancer therapy also included targeted biologics (23.8%), targeted small molecules (7.1%), and other (0.8%)
Reference: [52, 54]

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

B.2.4.1 Definition of study populations

The analysis sets used in CodeBreaK100 are defined in Table 7. Of a total of 126 subjects with NSCLC, 124 subjects were included in the full analysis set, and 2 subjects were excluded as they did not have at least 1 measurable lesion according to BICR. All 126 subjects were included in the safety analysis set [52, 54].

Table 7. Analysis sets used in CodeBreaK100

Analysis Set	Definition
Phase 2 full analysis set	All subjects in phase 2 who received at least 1 dose of sotorasib and have 1 or more measurable lesions at baseline as assessed by BICR using RECIST 1.1. This analysis set was to be used to evaluate response-related endpoints in the primary and final analyses.
Phase 2 safety analysis set	All subjects that enrolled in phase 2 and received at least 1 dose of sotorasib. This analysis set was to be used to evaluate safety and overall survival in the primary and final analyses.

BICR, blinded independent central review
Reference: [52, 54]

B.2.4.2 Primary analysis and updated analyses

The primary analysis was performed with a data cut-off date of 01 September 2020, approximately 8.5 months after at least 105 subjects with NSCLC were enrolled in the phase 2 portion of the study [52]. The updated analysis conducted with a data cut-off of 01 December 2020 provides 90 days of additional efficacy data and includes 6 months of

follow-up following onset of response for all responders included in the primary analysis, and updated ORR and DOR based on new responders as of 01 December 2020, as well as updated progression-free survival (PFS) and overall survival (OS) data [54]. Efficacy results from a further data cut of 15 March 2021, which provide an additional 3 months of follow up, are available [56] and are practically identical to the results from the 01 December 2020 data cut (see Table 8). As patient-level data required for the indirect treatment comparisons (see section B.2.9) are not yet available from the 15 March 2021 data cut, this submission focuses on the results of the 01 December 2020 data cut, which are used in the indirect comparisons and economic model.

B.2.4.3 Statistical analyses

The approach to statistical analysis of the efficacy endpoints is summarised in Table 5. A sample size calculation estimated that 105 NSCLC subjects would provide approximately a 90% probability that the lower limit of the ORR 95% CI would exceed the prespecified benchmark of 23% used to define a clinically significant ORR (see section B.2.3.3). The study therefore aimed to enrol at least 105 subjects with NSCLC [52]. As the study enrolled 126 subjects with NSCLC, the study was well powered for the primary endpoint.

B.2.5. Quality assessment of the relevant clinical effectiveness evidence

Quality assessment of the single-arm, non-randomised CodeBreaK100 study has been conducted using the Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I), tool [71].

The full quality assessment is presented in Appendix D. Notwithstanding the single-arm, non-randomised design, the CodeBreaK100 study is considered to be at a low risk of bias, with good external validity. There is no reason to doubt that the results of CodeBreaK100 are generalisable to the anticipated use of sotorasib in *KRAS p.G12C*-mutated NSCLC patients in clinical practice in the UK.

B.2.6. Clinical effectiveness results of the relevant trials

Summary of clinical effectiveness for sotorasib

- Current efficacy data in support of sotorasib in its anticipated licensed indication and in this appraisal are available from the well-conducted, ongoing, phase 2, single-arm CodeBreaK100 trial.
- Patients enrolled in the CodeBreaK100 had progressed after prior therapy and had received 1-3 prior lines of therapy. The vast majority had received prior immunotherapy, in line with current clinical practice in the UK.
- Primary and secondary efficacy endpoint data from the latest data cut (01 December 2020) consistently indicate that sotorasib is a highly effective targeted therapy in patients with *KRAS G12C*-mutated NSCLC:
    - Sotorasib achieved a primary endpoint of objective response rate (ORR, complete response + partial response) of 37.1% (95% CI: 28.6 to 46.2)
        - The lower bound of the 95% CI for ORR excluded the prespecified benchmark ORR of 23%, signifying the clinical significance of the response with sotorasib.
        - For further reference this compares well with an ORR of 13.7% observed with docetaxel monotherapy in KRAS-mutated NSCLC.
        - UK clinical experts at a recent advisory board held by Amgen, February 2021, agreed the ORR with sotorasib was far better than they would expect to see in patients treated with current standard of care docetaxel or nintedanib plus docetaxel.
    - With a median time to response of 1.35 months and median duration of response of 10.0 months, these data indicate that the clinically significant response to sotorasib is rapid and durable.
    - Median progression-free survival (PFS) was 6.8 months (95% CI: 5.1 to 8.2) and median overall survival (OS) was 12.5 months (95% CI: 10.0 to not estimatable)
        - Although CodeBreaK100 was not specifically powered for survival outcomes these data support the clinical benefit of sotorasib predicted by the ORR
        - As a point of reference, with docetaxel monotherapy in KRAS-mutated NSCLC median PFS was 2.9 months and median OS was 7.9 months.
    - Pre-specified subgroup analyses indicate that results are generally consistent across patient characteristics but are based on small sample sizes that warrant caution in their interpretation. None of the analyses are sufficient to support the preferential use of sotorasib in any subgroup, and therefore sotorasib should be used within its full anticipated licensed indication.
- These early clinical data indicate strongly that targeted treatment with sotorasib has potential to provide a step change in the management of *KRAS p.G12C*-mutated NSCLC in the UK.

B.2.6.1 Disposition of subjects

As of the 01 December data cut, of the 126 subjects with NSCLC, 56 subjects (44.4%) were continuing the study and 70 subjects (55.6%) had discontinued from the study. Of the 70 subjects who had discontinued from the study, 58 subjects (46.0%) had died, and 10 subjects (7.9%) withdrew consent. Two subjects were lost to follow-up. A total of 95

subjects (75.4%) had discontinued sotorasib treatment. The primary reasons for treatment discontinuation were disease progression (75 subjects, 59.5%) and adverse event (11 subjects, 8.7%) [54]. Full details of subject disposition are provided in Appendix D.

B.2.6.2 CodeBreaK100 trial results

A summary of the results for the primary and secondary efficacy endpoints by central review based on the primary analysis (data cut of 01 September 2020) and the updated analyses (data cuts 01 December 2020, and 15 March 2021) are presented in Table 8. Results of the updated analyses are highly consistent, supporting the focus of the submission on the 01 December 2020 updated analysis for which patient-level data are available for use in the indirect treatment comparisons and economic model. Results for the objective response and progression-free survival for 01 December data cut by investigator assessment were consistent with central review and are provided in Appendix E.

Table 8. Summary of results for primary and secondary efficacy endpoints in CodeBreaK100

	Phase 2 NSCLC 960 mg daily		
	Primary analysis (Data cut 1st September 2020) (n=123*) [52]	**Updated analysis (Data cut 1st December 2020)¶ (n=124*) [54]**	Updated analysis (Data cut 15th March 2021)¶ (n=124*) [56]
Best overall response - n (%)			
Complete response (CR)	2 (1.6)	3 (2.4)	4 (3.2)
Partial response (PR)	44 (35.8)	43 (34.7)	42 (33.9)
Stable disease (SD)	53 (43.1)	54 (43.5)	54 (43.5)
Progressive disease (PD)	20 (16.3)	20 (16.1)	20 (16.1)
Not evaluable	2 (1.6)	2 (1.6)	2 (1.6)
Not done	2 (1.6)	2 (1.6)	2 (1.6)
Primary endpoint:			
Objective response rate (ORR)			
Number of overall responders - N1 (%)	46 (37.4)	46 (37.1)	46 (37.1)
95% CI[a]	(28.84, 46.58)	(28.60, 46.23)	(28.60, 46.23)
Secondary endpoints:			
Duration of response (KM) (months)[b]			
Median (95% CI)	8.4 (6.9, 8.4)	10.0 (6.9, 11.1)	11.1 (6.9, NE)
Min, Max (+ for censored)	1.3+, 8.4	1.2+, 11.1	NR
Disease control rate (DCR) - n (%)	99 (80.5)	100 (80.6)	100 (80.6)
95% CI[a]	(72.37, 87.08)	(72.58, 87.19)	(72.58, 87.19)

	Phase 2 NSCLC 960 mg daily		
	Primary analysis (Data cut 1st September 2020) (n=123*) [52]	**Updated analysis (Data cut 1st December 2020)¶ (n=124*) [54]**	Updated analysis (Data cut 15th March 2021)¶ (n=124*) [56]
Time to objective response (months)[b]			
Mean (SD)	1.95 (1.23)	2.11 (1.71)	NR
Median (Min, Max)	1.35 (1.2, 6.1)	1.35 (1.2, 10.1)	1.4 (1.2, 10.1)
Progression-free survival (KM) (months)			
Median (95% CI)	6.7 (4.9, 8.1)	6.8 (5.1, 8.2)	6.8 (5.1, 8.2)
Min, Max (+ for censored)	0.3+, 11.5	0.3+, 12.6	NR
Overall survival (KM) (months)			
Median (95% CI)	12.0 (9.5, NE)	12.5 (10.0, NE)	12.5 (10.0, NE)
Min, Max (+ for censored)	1.1, 12.2+	1.1, 15.6+	NR

N = Number of subjects in the analysis set. n = Number of subjects with observed data.
CI = Confidence Interval; KM = Kaplan-Meier; NE = Not Estimable; NR = Not Reported.
Months are derived as days x (12/365.25).
[a] Exact 95% confidence interval was calculated using the Clopper Pearson method.
[b] Time to response and duration of response are calculated among confirmed responders N1.
* At the time of the September 2020 analysis, 123 subjects were deemed to have at least 1 measurable target lesion at baseline and included in the full analysis set. Following central review for the December 2020 data cut, one additional subject was re-evaluated as having had at least 1 measurable target lesion at baseline, and so was included in the December 2020 full analysis set.
¶ Given the high consistency of results between 1st December 2020 and 15th March 2021 but a current lack of patient-level data from the 15th March 2021 data cut, the results of the 1st December 2020 data cut are discussed below, and used in the indirect treatment comparisons and economic model
Reference[52, 54, 56]

B.2.6.2.1 Primary endpoint – objective response rate (updated analysis)

The ORR (95% CI) based on blinded independent central review (BICR) per Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 was 37.1% (95% CI: 28.6, 46.2) with 46 of the 124 evaluable subjects showing response, including 3 subjects (2.4%) with a complete response and 43 subjects (34.7%) with a partial response (Table 8) [54]. The lower bound of the 95% CI for ORR excluded the prespecified benchmark ORR of 23%, signifying the clinical significance of the response with sotorasib, and UK clinical experts at a recent advisory board held by Amgen, February 2021, agreed this ORR was far better than they would expect to see in patients treated with current standard of care docetaxel or nintedanib plus docetaxel [14].

B.2.6.2.2 Secondary endpoints (updated analysis)

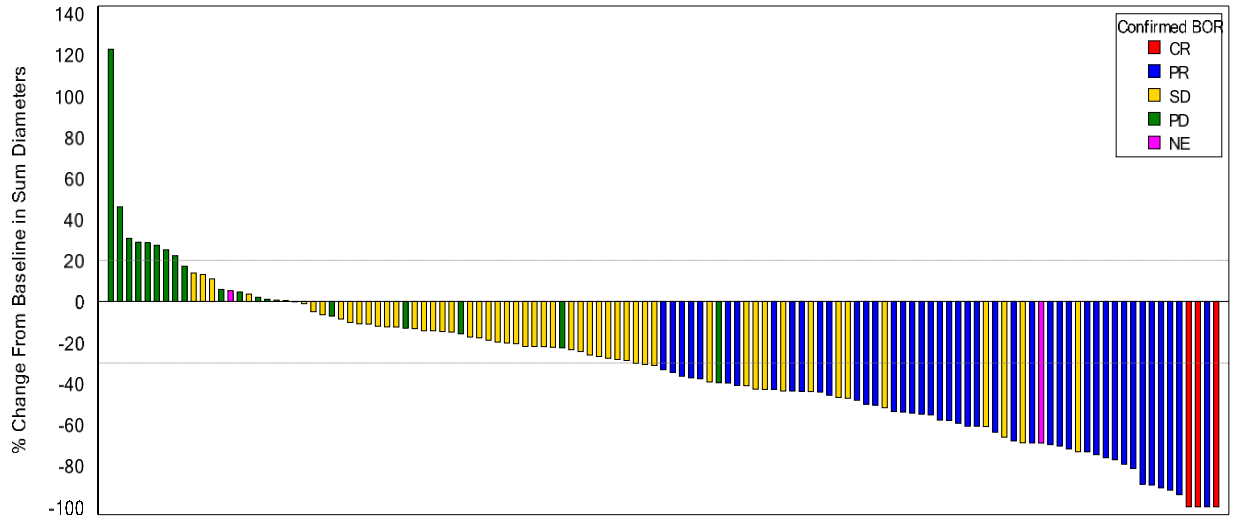
Duration of response

Among the 46 objective responders who had NSCLC, the KM estimate of median (95% CI) DOR was 10.0 (6.9, 11.1) months (Table 8). Twenty-seven subjects (58.7%) were censored, including 20 (45.7%) who were still on study treatment without disease progression [54]. These data indicate that the clinically significant ORR observed with sotorasib is durable.

Disease control rate

The disease control rate, which includes complete response, partial response or stable disease and is a strong predictor of clinical benefit, was high at 80.6% (95% CI: 72.6, 87.2). The percentage of subjects with stable disease was 43.5% [54], which is an important finding since not all patients with advanced NSCLC have tumour shrinkage after cancer therapies [69, 70]. Tumour shrinkage by best overall response to sotorasib is shown in Figure 3.

Figure 3. Waterfall plot of best tumour shrinkage

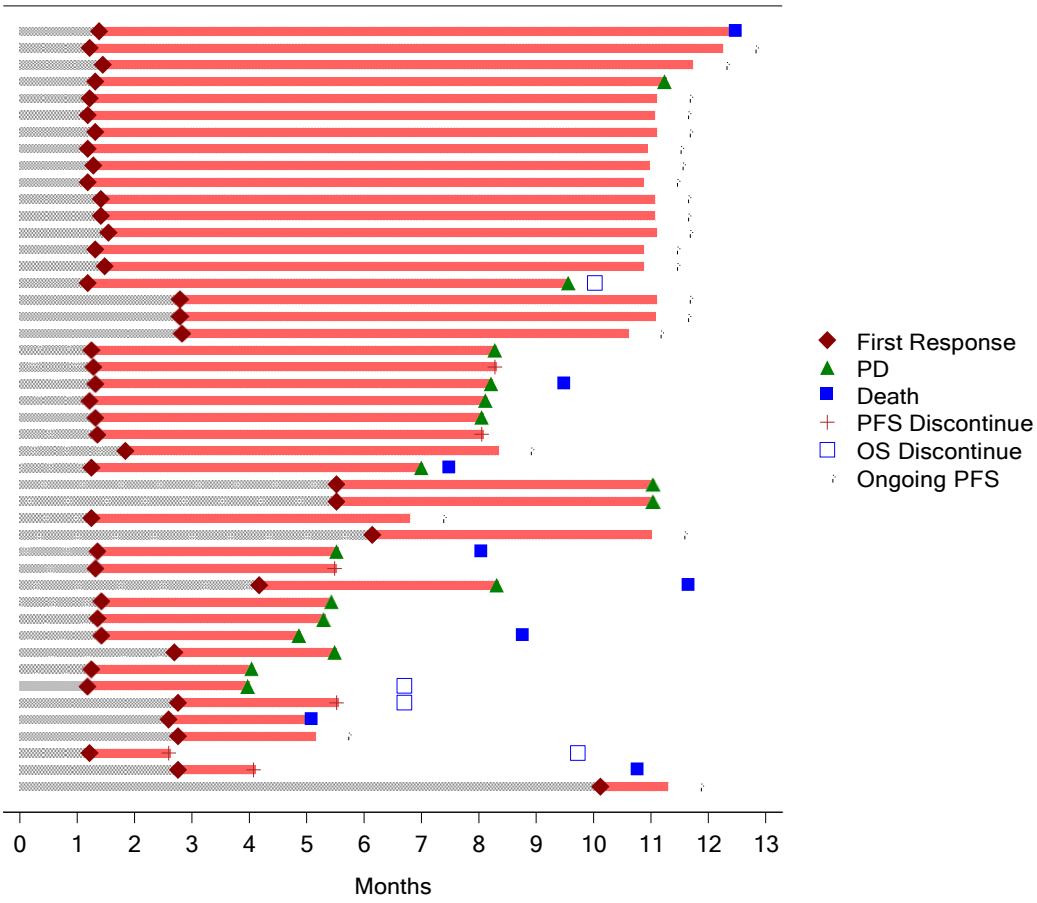


Phase 2 data cut-off date 01DEC2020.
Percent change from baseline in sum of diameters only considers tumor assessments prior to and include the 1st assessment where timepoint response is progressive disease, and prior to start of next anti-cancer therapy.
Two subjects without baseline target lesions and 3 subjects without post-baseline percent changes are not shown.

Time to response

Among the 46 responders in the NSCLC group, the median (range) time to response was only 1.35 (1.25, 2.69) months (Table 8) [54]. Combined with the magnitude and duration of the response, these data indicate that response to sotorasib is clinically significant, rapid and durable, as shown in Figure 4.

Figure 4. Swimmer plot of duration and time to response

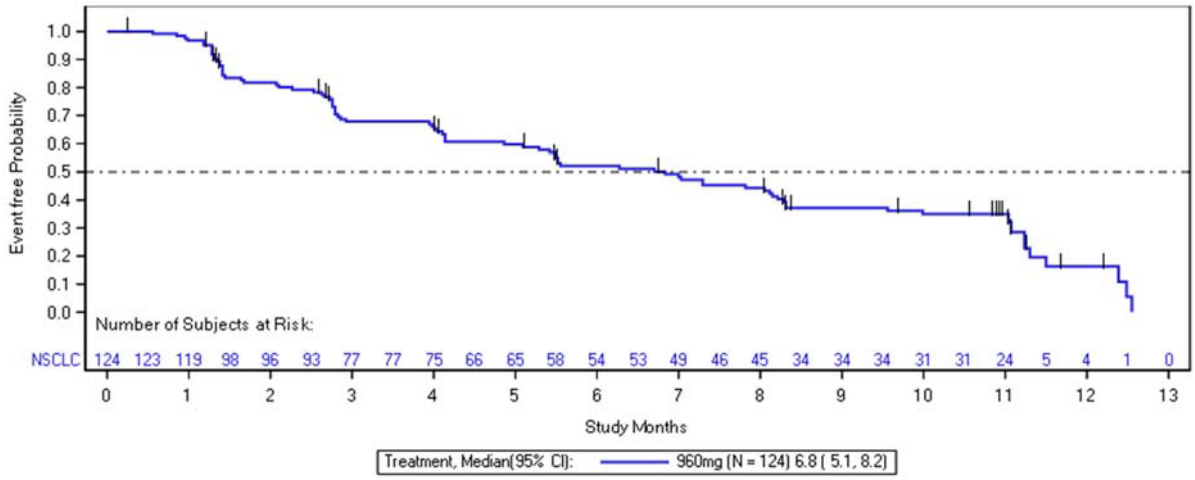


Phase 2 data cut-off date 01DEC2020.
'PFS Discontinue' indicates PFS censor due to no post-baseline assessment, withdrew consent, started of new anti-cancer therapy, missed two or more consecutive tumor assessments, off study due to sponsor decision, or lost to follow-up.
'OS Discontinue' indicate OS censor due to withdrew consent, completed study, off study due to sponsor decision, or lost to follow-up.

Progression-free survival

As of the data cut-off date, median (95% CI) PFS was 6.8 (5.1, 8.2) months (Table 8; Figure 5). The Kaplan-Meier estimate of survival (95% CI) was 52.2% (42.6, 60.9) at 6 months and 16.3% (7.4, 28.2) at 12 months. Seventy patients (56.5%) had experienced disease progression and 13 (10.5%) death events. A total of 41 patients (33.1%) were censored, and of those, 25 (20.2%) were on study without disease progression 7 (5.6%) started new anticancer therapy, 5 (4.0%) missed more than 1 consecutive assessment, and 3 (2.4%) withdrew consent (Appendix E) [54].

Figure 5. Kaplan-Meier Plot of Progression-free Survival



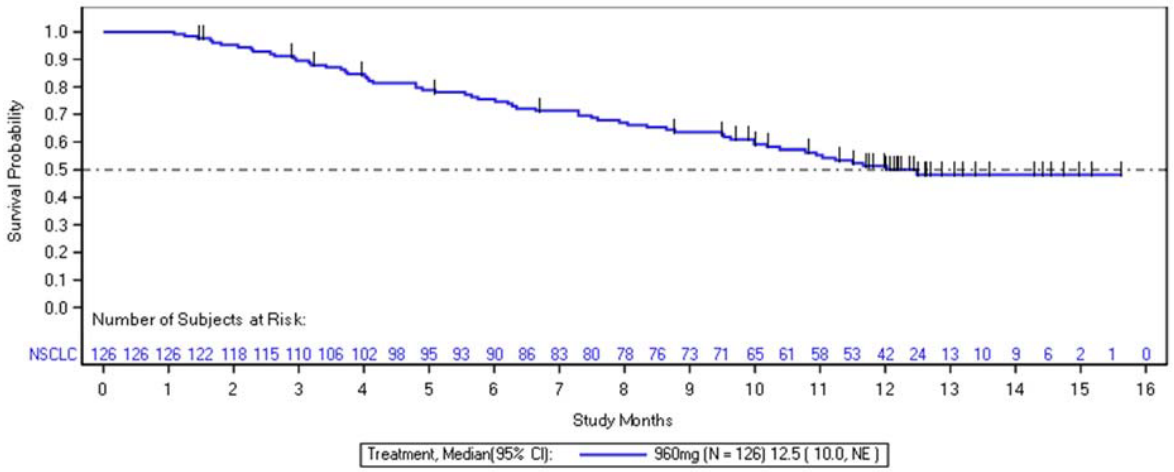
Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Radiological Progression or Death (whichever occurs earlier) is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-002-001-eff-km-pfs-nsclc-p2fas.rtf (Date Generated: 28JAN21:12:38:15) Source: adam.adsl, adam.adtte

Overall survival

Median (95% CI) overall survival (safety analysis set) was 12.5 (10.0, NE) months (Table 8; Figure 6**Figure *5***). The Kaplan Meier estimate of survival (95% CI) was 75.5% (66.8, 82.2) at 6 months and 51.4% (41.9, 60.1) at 12 months. As of the data cut-off date, 59 (46.8%) patients had experienced a death event (Appendix E). A total of 67 patients (53.2%) were censored, and of those, 56 (44.4%) were alive at last follow-up and 9 (7.1%) withdrew consent [54].

Figure 6. Kaplan-Meier Plot of Overall Survival (safety analysis set)



Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Death is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-003-001-eff-km-osm-nsclc-p2saf.rtf (Date Generated: 28JAN21:12:38:17) Source: adam.adsl, adam.adtte

Although the CodeBreaK100 study was not specifically powered for survival outcomes, these PFS and OS data support the clinical benefit of sotorasib predicted by the ORR primary endpoint data. Collectively, the data from the CodeBreaK100 study provide a strong indication of clinically significant efficacy with sotorasib in patients *KRAS p.G12C*-mutated NSCLC who have an urgent need for effective, targeted therapy.

B.2.6.2.3 Exploratory analyses – patient reported outcomes

Patient-reported outcomes (PRO) were assessed as exploratory outcomes using eight different validated tools, including generic and cancer-specific quality of life questionnaires, and symptom and adverse event severity ratings (see Table 5). Results are summarised below based on the latest available data for these descriptive analyses (the primary analysis, 01 September 2020 data cut) [57].

General summary of patient reported outcomes

Based on results across the various PRO tools it is clear that subjects with *KRAS p.G12C*-mutated NSCLC had high symptom burden and impaired quality of life when enrolled in the CodeBreaK 100 trial, as would be expected for this patient group. Whilst on sotorasib treatment, symptom burden and quality of life were stable or marginally improved, and few subjects experienced bothersome side effects. Given that existing non-targeted standard of care therapies are associated with toxicity, intolerance and quality of life impairment (see section 0), these data indicate that targeted therapy with sotorasib may offer important advantages over existing therapy [57]:

- Overall, at baseline, 70-78% completed the PRO questionnaires that were included at study start. For subjects on treatment, compliance rates ranged 69% to 88% from cycles 2 to 13 during the trial.

- Subjects with NSCLC reported high symptomatic burden and impaired physical function and quality of life.
    - Mean baseline scores for physical functioning and global health status/QoL on the QLQ-C30 were 71.7 and 63.9, respectively, in line with the normative values of 78.4 and 58.8 for patients with NSCLC (Scott et al, 2008) and below or comparable to those of 85.1 and 66.1 for the general population (Nolte et al, 2019).
    - Over time, mean global health status/QoL scores were generally sustained or marginally improved compared with baseline, with mean change (for cycles 2 to 13 where there were > 5 subjects) ranging from 1.9 (cycle 3) to -5.3 (cycle 11).

- A trend toward improvement or stabilization was observed in the severity of key lung cancer symptoms of cough, dyspnoea and chest pain.
    - Mean baseline QLQ-LC13 scores for cough, dyspnoea, and chest pain were 36.1, 31.7, and 15.1, respectively, in line with the normative values for patients with NSCLC of 38.4, 29.9, and 19.5, respectively (Scott et al, 2008).
    - Mean scores were decreased (indicating a decrease in symptoms) for dyspnoea, cough, and chest pain at most treatment visits compared with baseline, ranging from mean decrease of 13.3 to mean increase of 4.3.

- Few subjects reported bother due to treatment-related side effects associated with sotorasib. Among those who reported symptom bother, most subjects described their symptoms as mild.
    - At baseline, most subjects (87%) with NSCLC reported that they were “not at all” or “a little bit” bothered by the side effects of treatment.
    - During treatment (up to cycle 13), subjects’ bother with side effects remained about the same over time as at baseline, with about 21-46% reporting some degree of bother. Between cycles 2 and 17, <8% reported being bothered by side effects of treatment “quite a bit” and 0% as “very much.”

EQ-5D-5L

The EQ-5D-5L instrument is a generic health status tool that was administered to subjects in the CodeBreaK 100 trial. It is a validated instrument that with the appropriate cross-walk to the EQ-5D-3L instrument is used to estimate utility values in our economic model (see section B.3.4).

At baseline, most subjects (68-94%) reported that they had no problems or slight problems with the 5 dimensions of health assessed by the EQ-5D-5L, i.e., mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. “No problems” were reported most often for self-care and least often for pain/discomfort. The dimension with the highest percentage of subjects reporting moderate or severe problems or unable to perform the activity was pain/discomfort with 33% reporting one of those responses at baseline [57].

During treatment, subjects’ level of difficulty with the 5 dimensions of health remained about the same. Between baseline and cycle 13 (where n >5), fewer than 14% reported that they had severe problems or were unable to perform the activity [57].

It should be noted that, as a generic health status questionnaire, the EQ-5D-5L is likely to be less sensitive to the impact of NSCLC on patients than the cancer-specific tools that were also employed in the CodeBreaK 100 trial. However, these results, like those for the cancer specific tools, indicate that there was no deterioration in subjects’ health status with sotorasib treatment. Full results for the EQ-5D-5L scales and visual analogue scores are provided in Appendix E.

B.2.7. Subgroup analysis

A range of prespecified subgroup analyses were conducted using the 01 December 2020 data cut to explore the consistency of sotorasib treatment effects across age, race, prior lines of anticancer therapy, prior immunotherapy treatment, ECOG status, histopathology type, disease status, presence of liver or brain or bone metastasis, smoking history, geographical region [54, 58] (see Appendix E). Results were generally consistent across the pre-specified subgroups; however, these analyses are based on small sample sizes that warrant caution in their interpretation. None of the analyses are sufficient to support the preferential use of sotorasib in any of the subgroups that have been considered, and the results therefore support the use of sotorasib in line with the full anticipated licensed indication, as detailed in the proposed positioning of sotorasib in section B.1.3.3.

B.2.7.1 Subgroup analyses for primary endpoint: ORR

In general, no notable treatment effects by subgroup were observed with sotorasib, with the possible exception of prior platinum-based chemotherapy and presence of brain metastases. The ORR was higher for subjects who had not received prior platinum-based chemotherapy (69.2% [9 of 13 subjects]) and was, correspondingly, marginally lower in those who had received it (33.3% [37 of 111 subjects]), compared with the overall subject population (37.1% [46 of 124 subjects]). However, no difference in effects was observed based on prior anti-PD-(L)1 treatment alone, or by number of prior lines of therapy. As is typically observed, ORR was higher for subjects without brain metastasis (42.9%% [42 of 98 subjects) than for those with brain metastasis (15.4% [4 of 26 subjects]) [54]. As noted above, interpretation of these analyses is limited by the small sample sizes in each subgroup (see Appendix E).

B.2.7.2 Subgroup analyses for secondary endpoints: PFS and OS

Analyses of PFS also showed the effects of sotorasib were generally consistent across the same range of subgroups. Median PFS was higher for subjects who had not received prior platinum-based chemotherapy (11.1 months [95%CI: 8.0, not estimable] and was, correspondingly, lower in those who had received it (5.5 months [95% CI 4.1 to 7.3], compared with the overall subject population (6.8 months [95%CI: 5.1, 8.2]). However, these data are based on only 6 progression events in 13 subjects who had not received prior platinum-based chemotherapy. There was little difference in PFS in the 11 subjects who had not received prior anti-PD-(L)1 therapy, and in the 98 subjects without brain metastasis, compared with the overall population. OS data followed a similar pattern but results are more difficult to interpret due to the fact the median OS had not quite been reached in the 01 December 2020 data cut (see Appendix E) [54].

B.2.8. Meta-analysis

As current efficacy data for sotorasib in the treatment of *KRAS p.G12C*-mutated NSCLC are based on a phase 2 single-arm trial, no meta-analyses have been conducted.

B.2.9. Indirect and mixed treatment comparisons

Summary of indirect treatment comparisons

- As detailed in section B.1.3.3, docetaxel monotherapy is the primary comparator for sotorasib; docetaxel in combination with nintedanib is a secondary comparator.
- Given the anticipated conditional approval of sotorasib is based on the single-arm CodeBreaK100 trial, it was necessary to explore unanchored indirect treatment comparison methods to estimate the comparative efficacy of sotorasib and the comparators.
- A systematic literature review and targeted review of previous HTAs identified only one trial (SELECT-1) providing sufficient docetaxel monotherapy data in patients with KRAS-mutated (including *KRAS p.G12C* -mutated) NSCLC. No trials of nintedanib plus docetaxel reporting KRAS status were identified. However, as published observational data and subgroup analyses of SELECT-1 indicate that outcomes with non-targeted treatment are similar irrespective of KRAS-mutant status, it was confirmed by UK clinicians that it would be possible to explore comparative effectiveness using SELECT-1 and other data sources.
- For the primary comparison of sotorasib vs docetaxel monotherapy, a primary Matching Adjusted Indirect Comparison (MAIC) analysis using the CodeBreaK100 trial for sotorasib and SELECT-1 trial for docetaxel monotherapy was determined to be feasible and appropriate.
    - Using available prognostically important covariates, sotorasib was statistically and clinically superior to docetaxel monotherapy for progression-free survival (6.3 months vs 2.8 months; hazard ratio 0.411 [95%CI: 0.315 to 0.537]) and for overall survival (>12.5 months vs 7.9 months; hazard ratio 0.604 [95%CI: 0.443 to 0.826])
    - In a sensitivity analysis using a broader range of covariates, the hazard ratios were reduced even further, and as these analyses do not take into account that patients in CodeBreaK100 were more heavily pre-treated and a higher proportion had brain metastases compared with the SELECT-1 population, the results are plausibly conservative.
- A supplementary proposenity score weighting analysis of CodeBreaK100 and an Amgen-conducted real-world study using the Flatiron database in patients with KRAS-mutated NSCLC treated with a range of chemotherapy regimens was conducted to explore an alternative data source and method of estimating relative treatment effects for the primary comparison of sotorasib vs docetaxel monotherapy.
    - Sotorasib was statistically and clinically superior for overall survival (12.5 months vs 9.0 months; hazard ratio 0.631 [95%CI: 0.438 to 0.910]) and numerically superior for progression-free survival (7.0 months vs 4.5 months; hazard ratio 0.721 [95%CI: 0.519 to 1.001]).
- For the secondary comparison of sotorasib vs nintedanib plus docetaxel, the LUME-Lung 1 trial, used to support the licensing and NICE appraisal of nintedanib plus docetaxel, provided an appropriate data source. Due to missing data and population differences, it was not feasible to conduct an MAIC analysis; however, an alternative approach implemented in the economic model estimates an extrapolated mean (undiscounted) overall survival gain of 6.5 months with sotorasib vs nintedanib plus docetaxel (23.5 months vs 17.0 months).
- These data sources and approaches, which were confirmed as appropriate with UK clinicians, provide plausible early evidence of clinically meaningful improvements in survival outcomes with sotorasib compared with current standard of care, non-targeted therapy.

B.2.9.1 Introduction to the indirect treatment comparisons

As described in Section B.1.3.3, docetaxel monotherapy is the relevant primary comparator for sotorasib in this appraisal, with docetaxel in combination with nintedanib considered as a secondary comparator (in patients with adenocarcinoma). As the anticipated conditional approval of sotorasib is based on the single-arm CodeBreaK100 trial, there are no direct comparative data for sotorasib and no common trial arms with which to conduct anchored indirect treatment comparisons or network meta-analyses. Alternative, unanchored methods for estimating the comparative efficacy of sotorasib and the relevant comparators are therefore needed.

Unanchored methods recognised as potentially appropriate in the *NICE Decision Support Unit report 18 – Methods for population-adjusted indirect comparisons in submissions to NICE* include propensity score weighting, such as matching-adjusted indirect comparison (MAIC), or outcome regression methods, such as simulated treatment comparisons [72]. As there is little in the literature to suggest that one methodology is superior to the other, and as unanchored MAICs have often been employed in NSCLC [73-76], MAIC methods based on propensity score weighting were explored as appropriate primary approaches.

B.2.9.2 Data Sources and Feasibility Assessment for Indirect Comparisons

B.2.9.2.1 Data sources

The exact method to provide comparative effectiveness data for sotorasib versus the primary and secondary comparator is determined by the availability of data for the comparators in the population of interest.

As discussed in section B.1.3.1.2, outcomes data for patients specifically with *KRAS p.G12C* mutated NSCLC are limited. The systematic literature reviews described in Appendix D sought to identify clinical trials of therapies conducted in patients with *KRAS*-mutant NSCLC, and identified only one RCT (SELECT-1) that provided sufficient PFS and OS data for docetaxel monotherapy (the primary comparator) in a population of patients with *KRAS*-mutated NSCLC (including *G12C* and non-*G12C* mutations) [31]. No RCTs of nintedanib plus docetaxel (the secondary comparator) conducted specifically in patients with known/reported *KRAS*-mutated NSCLC were identified. In addition to the SELECT-1 RCT, PFS and OS data in advanced NSCLC patients harbouring the *KRAS p.G12C* mutation and other KRAS mutations in the US are available from large real-world evidence studies undertaken by Amgen using the Flatiron Health - Foundation Medicine Clinico-Genomic Database [37]. These included patients taking a range of therapies, including docetaxel monotherapy (see Appendix D).

As also discussed in section B.1.3.1.2, several published observational studies in Western (European, Australian and US) populations show that survival in patients with *KRAS p.G12C*-mutated NSCLC is similarly poor as that in patients with other *KRAS* mutations or wild type disease who are not eligible for existing targeted therapies [33-35]. Furthermore, the large Flatiron Health real-world evidence study undertaken by Amgen in US patients shows that PFS and OS is similar for patients with *KRAS p.G12C*-mutated NSCLC and those with other *KRAS* mutations who are not eligible for existing targeted therapies [37]

(see Table 3), and the SELECT-1 trial itself showed that survival outcomes were highly consistent for those with *KRAS p.G12C* mutated and those with other *KRAS*-mutated NSCLC [31]. On this basis, the PFS and OS data for docetaxel from the SELECT-1 trial in patients with *KRAS*-mutant NSCLC (including *G12C* and non-*G12C* mutations) is considered to be sufficiently reflective of PFS and OS in patients with *KRAS p.G12C*-mutated NSCLC who receive docetaxel monotherapy. The SELECT-1 trial was therefore considered to be a candidate to provide comparator data for the primary comparison of sotorasib vs docetaxel monotherapy. This approach was agreed as reasonable by the five UK clinical experts attending an Amgen advisory board meeting in February 2021 [14].

Similarly, the PFS and OS data from the large Amgen Flatiron Health real-world evidence study is considered to be sufficiently reflective of PFS and OS in patients with *KRAS p.G12C*-mutated NSCLC; as patients specifically with *KRAS p.G12C*-mutated NSCLC made up around 10% of the overall population of NSCLC patients in this study, but outcomes were very similar irrespective of KRAS-mutant status (Table 3), it was determined that analyses should be explored using the whole data set of patients (n=7,069). In contrast to the SELECT-1 population, approximately 24% of patients in the Flatiron Health dataset had received prior first-line immunotherapy, and patients were treated with a range of non-targeted second and subsequent line therapies that included docetaxel monotherapy, rather than exclusively with docetaxel monotherapy[37]. Therefore, this Amgen Flatiron Health real-world evidence study is considered to be a supplemental, alternative source of data that may be used as confirmatory validation of analyses undertaken using the SELECT-1 trial-based docetaxel monotherapy data. This approach was considered reasonable by the UK clinical experts attending the recent advisory board [14].

Although no RCTs of nintedanib plus docetaxel specifically in patients with *KRAS*-mutated NSCLC were identified, given that nintedanib plus docetaxel is a non-targeted treatment regimen, and outcomes in patients for whom targeted therapies are not a treatment option are similarly poor irrespective of KRAS-mutant status (see section B.1.3.1.2), it is necessary and appropriate to assume that PFS and OS in NSCLC patients treated with nintedanib plus docetaxel will be sufficient to reflect PFS and OS in patients with *KRAS p.G12C*-mutated NSCLC who receive nintedanib plus docetaxel. On this basis, the pivotal trial supporting the licensing and NICE-recommendation of nintedanib plus docetaxel in patients with NSCLC (LUME-Lung 1 [32]) was considered a potential candidate to provide comparator data for the secondary comparison of sotorasib versus nintedanib plus docetaxel. This approach was also considered reasonable by the UK clinical experts attending the recent advisory board [14].

B.2.9.2.2 Compatibility of data sources

Having determined the most appropriate candidate data sources and their ability to reflect the populations and comparators of interest it was necessary to further determine the compatibility of these and the sotorasib CodeBreaK100 trial to determine the feasibility of conducting the indirect treatment comparisons. As the primary comparison is of sotorasib vs docetaxel monotherapy, for which the primary comparator data source is the SELECT-1 trial, and the secondary comparison is of sotorasib vs nintedanib plus docetaxel, for which the comparator data source is the LUME-Lung 1 trial, the assessment of compatibility below relates to the compatibility of CodeBreaK100 vs SELECT-1 and LUME-Lung 1; the

assessment of compatibility of the Amgen real world data set from the Flatiron Health - Foundation Medicine Clinico-Genomic Database, used as a supplemental, alternative source of data for the primary comparison, is provided in Appendix D.

Study Designs and Eligibility Criteria of CodeBreaK100, SELECT-1 and LUME-Lung 1

An overview of the study designs and eligibility criteria of CodeBreaK100, SELECT-1 and LUNE-Lung 1 is provided in Table 9.

Table 9. Overview of study designs of CodeBreaK100, SELECT-1 and LUME-Lung 1 trials

Study characteristics	Sotorasib (CodeBreaK 100) [54]	Docetaxel monotherapy (SELECT-1) [31]	Docetaxel + nintedanib (LUME-Lung 1) [32]
Blinding	Open label	Double-blinded	Double-blinded
Inclusion criteria	Male or female patients (≥ 18 years) Histologically confirmed locally advanced or metastatic NSCLC *KRAS p.G12C* mutation identified through molecular testing ECOG Performance Status 0 – 1 ≥ 1 prior line of systemic anticancer therapy	Male or female patients (≥ 18 years) Histologically confirmed locally advanced or metastatic NSCLC KRAS-mutation identified through molecular testing WHO Performance Status 0 – 1 1 prior line of systemic anticancer therapy	Male or female patients (> 18 years) Histologically confirmed locally advanced or metastatic NSCLC ECOG Performance Status 0 – 1 1 prior line of systemic anticancer therapy
Key exclusion criteria	Active brain metastases Anti-tumour therapy including chemotherapy, antibody therapy, molecular targeted therapy, retinoid therapy within 28 days of study day 1	Brain metastases Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC Prior treatment with a MEK inhibitor or any docetaxel-containing regimen (prior treatment with paclitaxel is acceptable)	Active brain metastases Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC Prior treatment with a VEGFR inhibitor (other than bevacizumab) or docetaxel
Primary endpoint	Centrally-assessed ORR	Investigator-assessed PFS	Centrally-assessed PFS
Key secondary endpoints	Centrally-assessed PFS; Investigator-assessed PFS; OS	OS	OS

Key: ECOG, Eastern cooperative oncology group; KRAS, MEK, mitogen activated protein kinase; NSCLC, non-small cell lung cancer; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; VEGFR, vascular endothelial growth factor receptor; WHO, World Health Organization

CodeBreaK100, SELECT-1 and LUME-Lung 1 were all multicentre studies that recruited patients with confirmed locally advanced or metastatic NSCLC (Stage IIIB – IV) who had failed prior therapy. CodeBreaK100 specifically enrolled patients with *KRAS p.G12C* mutations, whereas SELECT-1 enrolled patients with KRAS mutations at codon 12, 13 or 61 [31]. LUME-Lung 1 did not specify KRAS mutations as an enrolment criterion and did not record KRAS mutations among the participants; however, in the subpopulation of interest (licensed population of patients with adenocarcinoma) the proportion of patients with *KRAS p.G12C* mutations is likely close to the prevalence of *KRAS p.G12C* mutations in the general non-squamous population (~13%). CodeBreaK100 enrolled patients with 1 to 3 prior therapies, whereas SELECT-1 and LUME-Lung 1 included patients with 1 prior therapy. All studies excluded subjects with active brain metastases, although CodeBreak100 and LUME-Lung 1 permitted inclusion of stable brain metastases.

All three studies reported PFS and OS as primary or secondary endpoints. PFS was assessed by investigators in SELECT-1, by both independent central review and by investigator in CodeBreaK100 and by independent central review in LUME-Lung 1.

Patient profiles in CodeBreaK100, SELECT-1 and LUME-Lung 1

A comparison of patient profiles in the CodeBreaK100, SELECT-1 and LUME-Lung 1 trials is presented in Table 10. As LUME-Lung 1 enrolled patients with mixed histology[32], but nintedanib in combination with docetaxel is only licensed for use in patients with adenocarcinoma[77], only the characteristics of the adenocarcinoma subpopulation of LUME-Lung 1 are considered. The distribution of patients between the three trials is similar in terms of age, disease stage and histology, and the majority of patients had ECOG/WHO performance status of 1.

Key characteristics for which there are differences between the trials arise from the different time points at which the trials were conducted. In addition to differences in KRAS mutation status (as discussed in section B.2.9.2.1), CodeBreak100 included patients taking 1-3 prior therapies and a high proportion of patients had prior use of PD(L)-1 inhibitors, reflecting the current treatment pathway for patients with *KRAS p.G12C* -mutated NSCLC in the UK. In contrast, the SELECT-1 and LUME-Lung 1 trials, which were both conducted before the evidence base supported front-line use of immunotherapy, included patients taking 1 prior therapy only and no PD(L)-1 inhibitors. Based on inclusion criteria and/or a lack of recording, it is also not possible to compare for the presence of (non-active) brain metastases in SELECT-1, for the PD-1 expression in LUME-Lung 1, or for the presence of other targetable mutations in either of these comparator trials. It is also of note that LUME-Lung 1 recruited fewer females, fewer prior smokers and patients with fewer brain metastases than CodeBreaK100.

Table 10. Comparison of baseline characteristics in CodeBreaK100, SELECT-1 and LUME-Lung 1 trials

Baseline characteristics [a]	Sotorasib (CodeBreaK100) n=126[54]	Docetaxel monotherapy (SELECT-1) (n=256) [31]	Docetaxel + nintedanib (LUME-Lung 1) (n=322)[j] [32]
Age	62.9 (mean)	60.9 (mean)	58.5 (median)
Gender (% female)	50%	43%	37%
Brain metastases (%)	21%	NR[c]	8%
Performance status (ECOG or WHO; % PS 1 [vs PS 0])	70%	59%	70%
Race (% white)	82%[d]	95%	NR[g]
% *KRAS p.G12C*-mutated	100%	42%[b]	NR[h]
Anti-PD-(L)1 in prior line(s)	91%	0%	0%
Number of prior lines (% with 1/2/3 prior lines)	43%/35%/22%	100%/0%/0%	Mostly 1 prior line[i]
Metastatic disease at baseline	96%	96%	90%
Histology (% Non-squamous)	99%	95%	100%[j]
Smoking status (% ever smoker)	93%[e]	92%	64%
Other targetable mutations (EGFR, ALK, BRAF, ROS-1)	3%	NR[f]	NR
PD-L1 expression at baseline (<5% [vs ≥5%])	48%	58%	NR

Key: ECOG, European Co-operative Oncology Group; NR, not reported

Note:

[a] all reported baseline characteristics in SELECT-1 and other key characteristics

[b] the rest of the population has KRAS mutations other than G12C

[c] not reported for SELECT-1. All studies had exclusion criteria for active brain metastases

[d] 15 percentage points of the 18% remaining correspond to Asian patients

[e] 2 percentage points of the remaining 7% are missing data

[f] probably very low due to KRAS mutant

[g] Race was not reported, the trial was non-US based and run mainly in Europe (71% of patients) as well as Asia

[h] LUME Lung-1 did not enrol by or record genetic mutations; the % of *KRAS p.G12C* is likely close to the prevalence of *KRAS p.G12C* mutations in the general non-squamous population (~13%)

[i] LUME Lung-1 included patients with a prior platinum-based therapy and allowed adjuvant/neoadjuvant as line of therapy

[j] Based on the subpopulation of interest (adenocarcinoma)

B.2.9.2.3 Conclusions on the feasibility of undertaking indirect comparisons

Despite some differences between CodeBreaK100 and SELECT-1 and LUME-Lung 1, UK clinical experts considered these were the best and most relevant sources of data available with which to make indirect comparisons for sotorasib in patients with *KRAS p.G12C*

mutated NSCLC [14]. The data are considered adequate to reflect PFS and OS outcomes with sotorasib, docetaxel monotherapy and nintedanib plus docetaxel treatment following prior therapy in this population of patients.

A propensity score weighted analysis approach such as MAIC requires the matching of prognostic patient characteristics to generate robust comparative treatment effect estimates. Due to missing data or other differences between the trials it would not be possible to match across all trials for *KRAS p.G12C* mutation status, brain metastases, prior lines of therapy or prior use of PD-L1 inhibitors. Given that PFS and OS outcomes are similar in the absence of targeted therapies, irrespective of *KRAS* status (see section B.2.9.2.1), the inability to match by specific KRAS status is unlikely to lead to biased estimates. Patients with brain metastases were excluded from SELECT-1 trial whereas CodeBreaK100 and LUME-Lung 1 permitted enrolment of non-active brain metastases; however, as brain metastases are an important prognostic characteristic, the inclusion of patients with brain metastases in the CodeBreaK100 trial but not in the SELECT-1 trial may lead to conservative estimates of relative treatment effects for sotorasib. Although CodeBreaK100 included patients with 1-3 prior therapies, to match only patients with 1 prior therapy, as per SELECT-1 or LUME-Lung 1 would effectively reduce the available CodeBreaK100 trial population by 57%, which would have significant implications for the precision of any relative treatment effect estimates. The inability to robustly match for number of prior lines of therapy or prior use of immunotherapy is therefore a potential limitation that arises due to limited comparator trial data specifically in KRAS-mutant NSCLC. However, as noted in section B.1.3.1.2, Table 3, PFS and OS outcomes are worse for patients with each successive line of therapy. Given that CodeBreaK100 included 57% of patients with 2 or more prior lines of therapy, a comparison of PFS and OS data from the whole of the CodeBreaK100 NSCLC population against PFS and OS data from patients in SELECT-1 or LUME-Lung 1, who had received only one prior line of therapy, is likely to be conservative.

On balance, in the context of this rare disease with limited available comparator trial data, an indirect comparison using these data sources is feasible and appropriate. The patient population in SELECT-1 appears to be more closely aligned with the CodeBreaK100 trial population than does LUME-Lung 1. On this basis, any formal indirect comparison using propensity score weighting approaches, which requires matching of patient characteristics, would be achieved more robustly for the primary comparison using CodeBreaK100 and SELECT-1 than for secondary comparison using CodeBreaK100 and LUME-Lung 1. To avoid matching CodeBreaK100 patients against a further population defined by the less closely aligned LUME-Lung 1 population, and the associated reduction in effective sample sizes in doing so, it was pragmatically determined that indirect comparisons should be made as follows:

- **Primary comparison of sotorasib vs docetaxel monotherapy**:
    - Primary analysis - formal MAIC for CodeBreaK100 vs SELECT-1
    - Confirmatory validation analysis – propensity score weighted analysis of CodeBreaK100 vs the Amgen Flatiron Health real-world evidence study (see Appendix D for further details of their comparability)

- **Secondary comparison of sotorasib vs nintedanib plus docetaxel**:
    - Alternative approach leveraging the common comparator (docetaxel monotherapy) between SELECT-1 and LUME-Lung 1 using appropriate hazard ratios.

These approaches were supported as reasonable by UK clinical experts [14].

B.2.9.3 Methodology of Indirect Comparisons

B.2.9.3.1 Primary comparison – Sotorasib vs Docetaxel monotherapy

Primary analysis – MAIC using CodeBreaK100 and SELECT-1

An MAIC was used to compare changes in OS and PFS with sotorasib versus docetaxel monotherapy. Patient-level sotorasib data were taken from the CodeBreaK100 trial using the safety analysis set (n=126) and data cut 01 December 2020. Although the primary analysis of PFS in CodeBreaK100 was by blinded independent central review, the investigator-assessed PFS data were highly consistent with these data (see Appendix D) and were used in the MAIC to align with the investigator assessment of PFS in SELECT-1. As patient-level data from the SELECT-1 trial were not available, the Kaplan-Meier curves for OS and PFS from the docetaxel monotherapy arm of the SELECT-1 trial were digitized using WebPlotDigitizer software [78] to create pseudo-patient-level data using the algorithm of Guyot [79].

MAIC requires the matching of prognostic patient characteristics to generate robust comparative treatment effect estimates. A starting list of candidate prognostic covariates was compiled based on literature reviews and informed by discussions with experienced NSCLC physicians. A total of six individual interviews were conducted via teleconferences with two physicians from Canada, and one each from the US, Germany, France and the UK. Pre-read documents (including a questionnaire) were circulated to the physicians; their corresponding responses and individual summary reports were shared with each of them for validation.

This resulted in total of five covariates that were considered as prognostically very important and 13 as somewhat important. Three additional prognostic covariates, related to race, ethnicity and histology at baseline, were added to the list based on their inclusion in recently conducted MAICs in treatment interventions in the NSCLC disease area [73, 75, 76], as noted by the expert clinicians. These covariates are presented in Table 11.

Table 11. Starting list of prognostic covariates

Category	Covariate
Very important	Baseline ECOG (0, 1)
	Presence of brain metastases (Y, N)
	Metastatic at baseline (Y, N)
	PD-L1 protein expression (<5%, >5%)
	Presence of at least one of the following mutations/alterations: EGFR, ALK, BRAF, ROS-1 (Y, N)

Somewhat important	Age
	Smoking status (history of smoking vs no history of smoking)
	Body mass index
	Presence of liver metastases (Y, N)
	Presence of bone metastases (Y, N)
	Number of sites of metastasis (0, 1, 2, 3 or more)
	Number of prior lines of therapies (1, 2, 3)
	Type of therapies administered in prior lines
	Time from prior line initiation to the index date (<3 months, 3 - 6 months, >6 months)
	Albumin at baseline
	Serum LDH
	Liver function (ALT, AST) at baseline
	Renal function (EGFR) at baseline
Additional covariates reported in other MAIC analyses	Sex (F; M)
	Race/Ethnicity (White; Others)
	Histology at baseline (Non-squamous; squamous)

Of these 21 potential covariates, 8 were selected for inclusion in the MAIC analysis based on data availability in SELECT-1, their prognostic importance for patients receiving sotorasib and docetaxel, and the feasibility of matching whilst preserving the effective sample size. As described in section B.2.9.2.3, it is not possible to match CodeBreaK100 patients against SELECT-1 for *KRAS* status or brain metastases, but this is unlikely to introduce significant bias. To match for number of prior lines of therapy would significantly reduce the effective sample size with loss of precision, and excluding this as a co-variate may actually result in conservative relative treatment effect estimates for sotorasib vs docetaxel monotherapy. PD-L1 expression at baseline was listed as potentially very important and these data are available from both the CodeBreaK100 and SELECT-1 trials; however, based on expert clinician feed-back this was clarified as a relevant prognostic factor for PD-L1 treatment but not for treatment with sotorasib and docetaxel.

Two different MAIC models were therefore considered, each comprising a different subset of the 8 selected covariates: a primary MAIC analysis including available covariates identified by physicians as being at least somewhat important; and, a sensitivity analysis including all 8 available covariates (see Table 12).

Table 12. Covariates used in the primary MAIC and in sensitivity analyses

Covariates	All variables pre-specified as important (Primary analysis)	All available variables
ECOG (% PS 1 [vs PS 0])	**X**	X
Age (mean)	**X**	X
Metastatic disease stage at baseline	**X**	X
Smoking status (% ever smoker)	**X**	X

Covariates	All variables pre-specified as important (Primary analysis)	All available variables
PD-L1 expression level		X
Gender (% female)		X
Histology (% Non-squamous)		X
Race (% white)		X
Key: ECOG, Eastern Co-operative Oncology Group; PS, performance status; 2L, second line		

The methods used to estimate weights in the MAIC are detailed in Appendix D, along with the distribution of statistical weights based on the balancing of the four covariates in the primary MAIC analysis. A comparison of the pre- and post-matching for covariates included in the primary MAIC analysis is presented in Table 13. Baseline characteristics post-matching were well balanced, with perfect matching for the four covariates included in the primary MAIC analysis and a difference of less than 5 percentage points for all other characteristics that were able to be compared.

Table 13. Post-matching balanced baseline characteristics

	As reported For docetaxel	Pre-matching For sotorasib	Post-matching (primary analysis)[a] For sotorasib
Covariates	**SELECT-1**	**CodeBreaK 100**	**CodeBreaK 100**
Age (mean)	60.9	62.9	[redacted]
Gender (% female)	43%	50%	[redacted]
ECOG (% PS 1 [vs PS 0])	59%	70%	[redacted]
Race (% white)	95%	82%[b]	[redacted]
Disease stage (% IIIB [vs IV])	4%	4%	[redacted]
Histology (% Non squamous)	95%	99%	[redacted]
Smoking status (% ever smoker)	92%	93%[c]	[redacted]

Key: ECOG, European Co-operative Oncology Group; PD-(L1), programmed death ligand 1; PS, performance status

Note: (a) when adjusting for four covariates; (b) 15 percentage points of the 18% remaining correspond to Asian patients; (c) 2 percentage points of the remaining 7% are missing data

Result of the primary MAIC and sensitivity analyses are provided in section B.2.9.4.

Supplementary analysis – Propensity score weighting analysis using CodeBreaK100 and Amgen Flatiron Health real-world evidence study

As a supplementary analysis to validate the primary MAIC analysis, full details of the propensity score weighting analysis using CodeBreaK100 and the Amgen Flatiron Health real-world evidence study are provided in Appendix D. In summary, this analysis aimed to compared OS and PFS with sotorasib against standard of care therapy observed in a cohort of patients with previously treated KRAS-mutated advanced/metastatic NSCLC in the Flatiron database. The 01 December 2020 data cut-off was used for the CodeBreaK100 trial

[54] and the Flatiron cohort included patients who were treated between 01 January 2011 and 31 March 2020, with follow-up data extending to 30 September 2020 [37].

Only patients in the Flatiron database who were being treated with chemotherapy were included; those receiving immunotherapy were excluded as immunotherapy is not a relevant comparator to sotorasib. Chemotherapy included docetaxel in a minority of patients, but is considered to represent the efficacy of docetaxel monotherapy given the lack of compelling trial data demonstrating a significant survival benefit of other regimens following prior therapy in KRAs-mutated NSCLC. Although the Flatiron database included patients with *KRAS p.G12C* mutations, the analysis is based on the full *KRAS*-mutation population to preserve effective sample size and precision of the relative treatment effect estimates; however, a sensitivity analysis limited to the *p.G12C* population is provided for completeness (see Appendix D).

Candidate prognostic covariates based on those identified in the MAIC (Table 11) were entered into a logistic regression model with sotorasib treatment as the binary response. Those identified as being very important (baseline ECOG performance status, presence of brain metastases, metastatic disease at baseline, presence of mutations/alterations) were fixed in the propensity score model, regardless of their statistical significance. An exception was “PD-L1 expression at baseline”, which was excluded from the model as it had a high proportion of missing values and is not an important predictor for sotorasib or chemotherapy-based treatment. A stepwise variable selection algorithm (complete case analysis) was run on the covariates identified as being somewhat important, resulting in age group, number of prior lines of therapies, prior PD-1 or PD-L1, prior platinum-based chemotherapy, and Albumin at baseline being retained in the model.

A propensity score weighting approach was used for this analysis, as the limited sample size of the two studies prohibits the use of propensity score matching. The average treatment effect of the treated (ATT) weight was used to balance the covariates of the chemotherapy-treated (Flatiron) population to fit the characteristics of the sotorasib-treated population in CodeBreaK100.

Appendix D includes full details of the balance between covariates and assignment of ATT weights. Results of this supplemental analysis and its sensitivity analysis in the *KRAS p.G12C* mutation population are provided in section B.2.9.4.

B.2.9.3.2 Secondary comparison – Sotorasib vs Nintedanib plus docetaxel

For the secondary comparison of sotorasib vs nintedanib plus docetaxel it was not feasible to conduct an MAIC analysis; whilst the patient population and treatment effects in LUME-Lung 1 are sufficient to reflect PFS and OS in *KRAS p.G12C*-mutated NSCLC treated with nintedanib plus docetaxel, the differences in patient characteristics and data availability for matching would present significant challenges, including reducing the effective sample size and precision for relative treatment effect estimates and introducing a further population that is less closely aligned with CodeBreak100 and not aligned with the SELECT-1 trial population to which sotorasib recipients in CodeBreaK100 had already been matched. This conclusion was confirmed by UK clinical experts at an Amgen advisory board [14]. It was therefore determined that an alternative approach would be required.

As the LUME-Lung 1 and SELECT-1 trials both include a docetaxel monotherapy treatment arm [31, 32], and as relative treatment effects of sotorasib vs docetaxel monotherapy in the SELECT-1 population were available from the primary MAIC, the use of a hazard ratio for PFS and OS from LUME-Lung 1 applied to the docetaxel arm of the SELECT-1 trial was explored to provide an estimate of the relative treatment effects of sotorasib vs nintedanib plus docetaxel. However, further examination of the LUME-Lung 1 data indicated that a piecewise approach to hazard ratio estimation would be required, and estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel could only be made following implementation of these data within the economic model (see section B.3.3.5 for full details).

B.2.9.4 Results of the indirect treatment comparisons

B.2.9.4.1 Primary comparison – sotorasib vs docetaxel monotherapy

Results of primary analysis – MAIC using CodeBreaK100 and SELECT-1

Results for the primary MAIC analysis, the sensitivity analysis using all available covariates, and an unadjusted analysis provided for reference, are reported in terms of hazard ratios (HRs) with 95% confidence intervals (CIs) based on robust standard errors (Table 14).

Table 14. Results of MAIC for primary comparison of sotorasib vs docetaxel monotherapy

Analyses	CodeBreaK 100 N (OS / PFS)	CodeBreaK 100 ESS (OS / PFS)	Median OS Sotorasib vs Docetaxel	Median PFS Sotorasib vs Docetaxel
Unadjusted	126	126	[redacted]	[redacted]
MAIC Model: "all variables of prognostic importance" **(Primary analysis)**	**123/ 121**	**108.8/ 106.1**	[redacted]	[redacted]
MAIC Model: "all available covariates" (sensitivity analysis)	98/ 96	53.3/ 53.1	[redacted]	[redacted]

Key: *Median OS not reached, OS was 50.4% at 12.5 months; ¶ Median OS not reached, OS was 52.5% at 12.0 months; CI, confidence interval; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; mths, months; OS, overall survival; PFS, progression-free survival; 2L, second line

The primary MAIC analysis focused on available covariates of prognostic importance. Matching of patients from CodeBreaK100 to SELECT-1 preserved an effective sample size of over 106; a small loss of data compared with the pre-adjusted sample size. The results indicate that sotorasib is statistically and clinically superior to docetaxel monotherapy for both PFS ([redacted]) and OS ([redacted]). For OS, matching adjustment resulted in an improvement such that a median OS was not achieved for sotorasib.

Sotorasib therefore provided at least a 4.6-month gain in median OS compared with the primary comparator docetaxel monotherapy. Kaplan-Meier plots for this adjusted primary analysis and for the unadjusted analysis are shown for OS in Figure 7 and for PFS in Figure 8.

The sensitivity analysis using all available covariates substantially reduced the effective sample size, as expected. However, despite the resulting lower precision, the results indicate an even greater improvement in OS and PFS with sotorasib vs docetaxel monotherapy, confirming that the OS and PFS benefit of sotorasib in the primary MAIC analysis is plausibly conservative. The results of the conservative primary MAIC analysis are used in the base case economic model (see section B.3.3).

Figure 7. Kaplan-Meier plot for primary MAIC analysis of OS for sotorasib and docetaxel monotherapy

Figure 8. Kaplan-Meier plot for primary MAIC analysis of PFS for sotorasib and docetaxel monotherapy

Results of the supplementary primary comparison – Propensity score weighting analysis using CodeBreaK100 and Amgen Flatiron Health real-world evidence study

This supplementary analysis was undertaken to explore an alternative data source and method of estimating relative treatment effects for sotorasib vs docetaxel monotherapy (using the basket of standard of care chemotherapy regimens in the Amgen Flatiron real-world evidence cohort as a proxy for docetaxel monotherapy).

ATT weighting substantially reduced the effective sample size, which impacts on the precision of the relative treatment effect estimates; the 95% confidence intervals around the hazard ratios are therefore very wide. Nonetheless, in the KRAS-mutant population there was clear evidence of a statistically and clinically significant benefit of sotorasib in OS, with a gain in median OS of [redacted] (Table 15). The hazard ratio of [redacted] is of very similar magnitude to that observed in the primary MAIC analysis (Table 14). A clear numerical benefit for PFS in favour of sotorasib was also observed, but due to the wide 95% confidence intervals did not achieve statistical significance.

Similarly, in the subgroup of the Flatiron population with *KRAS p.G12C*-mutated NSCLC the point estimates of the OS and PFS hazard ratios clearly favour sotorasib numerically and suggest it is possible that the hazard ratio estimates based on the wider KRAS mutant population may be conservative; however, the effective sample size is very small.

Table 15. Results of supplementary primary comparison using propensity score weighting analysis

Outcome	Flatiron N before adjustment	*KRAS* mutant		*KRAS-p.G12Cmutant* subgroup	
		ESS	Median HR (95% CI)	ESS	Median HR (95% CI)
Overall survival	206	104.8	[redacted]	17.8	[redacted]
Progression-free survival	206	104.8	[redacted]	17.8	[redacted]

Key: ATT, average treatment effect of the treated; CI, confidence interval; ESS, effective sample size of Flatiron population following adjustment to CodeBreaK100 population; HR, hazard ratio; mths, months; OS, overall survival

Overall, this supplementary analysis, using a different comparative data set and a different, appropriate method of adjustment to align the study populations, supports the results obtained in the primary MAIC analysis. The results in the KRAS mutant Flatiron population, with the larger effective sample size, are explored in a scenario analysis in the economic model (see section B.3.3.4).

B.2.9.4.2 Results for the secondary comparison – sotorasib vs nintedanib plus docetaxel

Estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel was implemented in the economic model (see section B.3.3.5 for full details). This extrapolated estimate over the model time horizon indicates sotorasib provides a gain in mean PFS of 4.2 months (9.2 vs 5.0) and a gain in mean OS of 6.5 months (23.5 vs 17.0) compared with nintedanib plus docetaxel (Table 16).

Table 16. Results of secondary comparison implemented in the economic model

	Sotorasib	Nintedanib plus docetaxel	Increment
Mean OS (months)*	[redacted]	[redacted]	[redacted]
Mean PFS (months)*	[redacted]	[redacted]	[redacted]

Key: *Derived from economic model with 20-year time horizon, undiscounted values (see section B.3.3.5 for how implemented); OS, overall survival; PFS, progression-free survival

B.2.9.5 Uncertainties in the indirect and mixed treatment comparisons

Key areas of uncertainty arise from the fact that trial and observational data in patients with *KRAS p.G12C*-mutated NSCLC are very limited. Whilst there are consistent observational and trial data indicating that survival outcomes in patients with *KRAS p.G12C* mutations are similarly as poor as in patients with other *KRAS* mutation status treated with non-targeted therapies (see section B.1.3.1.2), all of the indirect treatment comparisons, by necessity, have relied on this assumption because comparator data specifically in *KRAS p.G12C* -mutated NSCLC are lacking.

Due to the data limitations with comparators, it has not been possible to match or weight the source trial populations for all known prognostic characteristics. Patients in the CodeBreaK100 trial of sotorasib were more heavily pre-treated than patients in the comparator SELECT-1 and LUME-Lung 1 trials; however, as survival outcomes are generally worse with increasing lines of therapy, any comparison of the CodeBreaK100 trial population against the SELECT-1 and LUME-Lung 1 populations that are less heavily treated will be conservative. In line with the current treatment pathway, the vast majority of patients enrolled in the CodeBreaK100 trial had been pre-treated with immunotherapy, whereas the SELECT-1 and LUME-Lung 1 trials were conducted before immunotherapy had become a widely used front-line therapy. The impact of this difference in populations is unclear, but may be mitigated by the more heavily pre-treated CodeBreaK100 population in the primary MAIC analysis. Furthermore, the cohort of patients in the Flatiron database, used in the supplementary propensity score weighting analysis, included a high proportion of patients previously treated with PD (L)-1 inhibitors, with results supportive of the primary MAIC analysis. Patients with brain metastases were excluded from SELECT-1 trial whereas CodeBreaK100 permitted enrolment of patients with non-active brain metastases; however, as the presence of brain metastases is an important negative prognostic factor, the inclusion of patients with brain metastases in the CodeBreaK100 trial but not in the SELECT-1 trial may also lead to conservative estimates of relative treatment effects.

The primary analysis for the primary comparison of sotorasib against docetaxel monotherapy could be conducted with a robust MAIC approach and the results were generally supported by an alternative approach using a different comparator data set and proposenity score weighting method. However, it was not feasible to conduct an MAIC for the secondary comparison of sotorasib vs nintedanib plus docetaxel due to more pronounced differences in patient characteristics and missing data from the LUME-Lung 1 study. A pragmatic approach was therefore required and adopted, and confirmatory validation with another data source or method was not possible. Despite these limitations, this and all other approaches and data sources adopted were considered by UK clinical experts as appropriate to indirectly demonstrate the comparative efficacy of sotorasib and relevant comparators in this challenging disease area [14].

B.2.9.6 Conclusions on the indirect treatment comparisons

In a primary comparison using the most robust indirect methods possible (MAIC) sotorasib was statistically and clinically superior to docetaxel monotherapy for both PFS ([redacted]) and OS (median [redacted]). The gain in OS with sotorasib was estimated to be at least [redacted] months based on these trial data. Sensitivity analysis, and the fact that the analyses could not be adjusted for the greater negative prognostic factors in the sotorasib trial population, suggest these estimates may be conservative. An alternative propensity score weighting analysis using an alternative comparator data source generally supports these findings.

For the secondary comparison of sotorasib vs nintedanib plus docetaxel, data limitations for the latter necessitated a pragmatic indirect comparison approach to be implemented in the

economic model. This extrapolated estimate over the model time horizon indicates sotorasib provides a gain in mean OS of [redacted] months compared with nintedanib plus docetaxel.

Given the phase 2 single arm trial data currently available in support of sotorasib, and the lack of data for the relevant comparators specifically in *KRAS p.G12C*-mutated NSCLC, every effort has been made to derive the most robust possible indirect estimates of relative efficacy for this innovative therapy. These analyses clearly indicate that sotorasib is a highly effective therapy that plausibly provides clinically meaningful improvements in survival outcomes compared with current, non-targeted standard of care therapies. Sotorasib therefore provides a much-needed targeted treatment option in patients with *KRAS p.G12C*-mutated NSCLC.

B.2.10. Adverse reactions

Summary of safety data for sotorasib

- Safety data from the CodeBreaK100 trial indicate that sotorasib is well tolerated with a very manageable adverse event profile.
- The most common treatment-related adverse events were diarrhoea (31%), nausea (19%), elevations in alanine and aspartate aminotransferase (15%) and fatigue (11%).
    - The majority of these were mild to moderate; 5 (4%) subjects experienced diarrhoea rated as grade 3 or greater, and 7-8 (6%) subjects experienced grade 3 or greater elevations in alanine and aspartate aminotransferase.
    - Of note, only one (0.8%) of the 126 subjects experienced neutropenia by the data cut-off date.
- Nine subjects (7%) discontinued sotorasib due to treatment-related adverse events.
- Patient reported outcomes data indicate that symptom burden and quality of life while on treatment were stable or marginally improved, with few subjects reporting bothersome side effects.
- Given that existing non-targeted standard of care therapies are associated with toxicity, intolerance and quality of life impairment, these data indicate that targeted therapy with sotorasib may offer important safety and tolerability advantages over existing non-targeted therapy.
- Of note, UK clinical experts at a recent advisory board held by Amgen, February 2021, agreed the safety and tolerability of sotorasib appears to be superior to that they would expect to see in patients treated with current standard of care docetaxel or nintedanib plus docetaxel.

B.2.10.1 Exposure data

As of the data cut-off, among the 126 subjects with NSCLC in the safety analysis set, the mean (SD) number of cycles of sotorasib (defined as 21 days of treatment) started was 9.2 (6.1) and the mean (SD) duration of treatment was 28.6 (18.8) weeks. Twenty-four subjects (19%) had treatment for 12 months or more. Twenty-six subjects (20.6%) had a dose change, primarily due to adverse events (22 subjects, 17.5%), and 67 subjects (53.2%) had at least one dose withheld, primarily due to adverse events (46 subjects, 36.5%). Eleven (8.7) subjects discontinued sotorasib due to adverse events. The mean (SD) average dose per day was 856.4mg (193.9) and mean (SD) relative dose intensity was 89.2% (20.2) [55].

B.2.10.2 Overall Summary of Adverse Events

An overall summary of adverse events is provided in Table 17. Although almost all subjects with NSCLC experienced some form of treatment-emergent adverse events (TEAE, adverse events that began between the first dose and 30 days after the last dose of sotorasib), the incidence of adverse events judged by investigating clinicians to be related to sotorasib treatment (treatment-related adverse events, TREAs) was much lower, with the majority judged to be mild to moderate. Twenty (20.6%) subjects experienced grade 3 or greater severity TRAE, and only 10 (7.9%) subjects experienced serious TRAEs. Nine (7.1%) subjects discontinued sotorasib due to TRAE, and there were no deaths due to TRAEs [55].

The overall incidence of TEAEs and TRAEs was similar in the non-NSCLC subjects enrolled on the CodeBreaK100 study (see Appendix F).

Table 17. Summary of overall adverse events in NSCLC subjects in CodeBreaK100

	Sotorasib 960 mg daily (N = 126) n (%)
All treatment-emergent adverse events	125 (99.2)
Grade ≥ 2	110 (87.3)
Grade ≥ 3	75 (59.5)
Grade ≥ 4	23 (18.3)
Serious adverse events	63 (50.0)
Leading to discontinuation of sotorasib	11 (8.7)
Serious	7 (5.6)
Non-serious	5 (4.0)
Fatal adverse events	20 (15.9)
Treatment-related treatment-emergent adverse events	88 (69.8)
Grade ≥ 2	49 (38.9)
Grade ≥ 3	26 (20.6)
Grade ≥ 4	1 (0.8)
Serious adverse events	10 (7.9)
Leading to discontinuation of sotorasib	9 (7.1)
Serious	4 (3.2)
Non-serious	5 (4.0)
Fatal adverse events	0 (0.0)

N = Number of subjects in the analysis set, n = Number of subjects with observed data.
Coded using MedDRA version 23.1. Severity graded using CTCAE version 5.0
Reference: [55]

B.2.10.3 Treatment-Emergent and Treatment-Related Adverse Events

Treatment-emergent adverse events (TEAE) of any severity occurring in ≥ 10% of NSCLC patients are presented in Appendix F. These included diarrhoea, nausea, fatigue, elevations in alanine and aspartate aminotransferase, and also dyspnoea and cough, which may be symptoms of the underlying disease.

Treatment-related adverse events (TRAEs) of any severity occurring in ≥ 5% of NSCLC, and their incidence of Grade 3 or greater severity, are presented in Table 18. The most common TRAEs were diarrhoea (31%), nausea (19%), elevations in alanine and aspartate aminotransferase (15%) and fatigue (11%). Importantly, the majority of these were mild to moderate; 5 (4%) subjects experienced diarrhoea rated as grade 3 or greater, and 7-8 (6%) subjects experienced grade 3 or greater elevations in alanine and aspartate aminotransferase. Of note, only one (0.8%) of the 126 subjects experienced neutropenia by the data cut-off date [55].

The adverse event profile of sotorasib therefore appears to be very manageable and UK clinical experts at a recent advisory board held by Amgen, February 2021, agreed the safety

and tolerability of sotorasib appears to be superior to that they would expect to see in patients treated with current standard of care docetaxel or nintedanib plus docetaxel [14].

Table 18. Treatment-related adverse events occurring in ≥5% of NSCLC subjects in CodeBreaK100

Treatment-related adverse events (TRAEs) occurring in ≥ 5%, n (%)	Any Grade N = 126	Grade 3+ N = 126
Any event	**88 (69.8)**	**25 (19.8)**
Diarrhoea	39 (31.0)	5 (4.0)
Nausea	24 (19.0)	0
ALT increase	19 (15.1)	8 (6.3)
AST increase	19 (15.1)	7 (5.6)
Fatigue	14 (11.1)	0
Vomiting	10 (7.9)	0
Blood alkaline phosphatase increase	9 (7.1)	1 (0.8)
Maculopapular rash	7 (5.6)	0

ALT, alanine aminotransferase; AST, aspartate aminotransferase
Reference:[55]

B.2.10.4 Events of interest and other safety findings

B.2.10.4.1 Events of interest

Prespecified adverse events of interest were hepatotoxicity and renal toxicity events. As of the 01 December 2020 data cut, forty-seven (37.3%) of subjects with NSCLC experienced these events, including 40 (31.7%) experiencing a hepatotoxicity and 21 (16.7%) reporting a renal toxicity event. The majority were mild to moderate in severity and did not require dose interruption or discontinuation.

For hepatotoxicity events, grade 3 or greater severity occurred in 22 (17.5%) subjects, sotorasib dose interruption was required in 17 (13.5%) subjects, and discontinuation was required in 7 (5.6%) subjects [55]. For renal toxicity events, grade 3 or greater severity occurred in 3 (2.4%) subjects, dose interruption was required in 1 (0.8%), but no sotorasib discontinuation was required [55]. These adverse events of interest were therefore manageable and rarely resulted in treatment discontinuation.

B.2.10.4.2 Other safety findings

Electrocardiogram and QTc analysis was prespecified among the safety endpoints in the trial. QT prolongation (>450 msec) was recorded as a treatment-emergent adverse event in 3 subjects (2.4%) and was classed as treatment-related in only 1 (0.8%) subject. There were no instances of T-wave abnormalities reported [55].

B.2.11. Ongoing studies

Data collection to address areas of uncertainty in the evidence base for sotorasib and potential comparators is ongoing in clinical trials and real-world evidence studies. Key amongst these is the confirmatory phase 3 RCT, which will provide directly comparative data for sotorasib vs the primary comparator in this appraisal, docetaxel monotherapy, within the next 2 years.

A summary of these ongoing studies and the current estimated timing of their analyses (subject to change) is provided in Table 19.

Table 19. Ongoing studies that may inform the evidence base for sotorasib and comparators

Study	Design	Outcomes assessed	Estimated* date for analyses
CodeBreaK100 (NCT03600883)	Phase 2, multi-country, single-arm trial	ORR; PFS; OS	Estimated study completion date Feb 24, 2025
CodeBreaK200 (NCT04303780)	Phase 3, multi-country, randomised, open-label trial of sotorasib vs docetaxel in ~330 *KRAS p.G12C*-mutated advance/metastatic NSCLC patients with ECOG PS 0-1 who have failed at least 1 prior systemic therapy	**Primary endpoint:** PFS **Key secondary endpoints:** OS ORR PROs	**PFS:** • Primary analysis cut off: January 2022 **OS:** • [redacted]
Expanded access study (NCT04667234)	Expanded access study in US, Brazil and Israel	Safety profile of sotorasib in patients with previously treated locally advanced/ unresectable/ metastatic *KRAS p.G12C* mutated NSCLC in a real-world setting	Ongoing; TBD
UK retrospective chart review	Retrospective cohort study to describe the characteristics, treatment patterns, outcomes, and healthcare resource use in KRAS mutant or KRAS wild-type NSCLC patients in the UK between Jan 2018-Dec 2019	Duration of treatment (DoT); real-world physician-assessed response; Time to next treatment (TTNT), real-world time to progression (rw-TTP); real-world progression-free survival (rw-PFS), overall survival (OS)	Q1/2 2022
PRO cross-sectional and retrospective chart review	HRQoL in KRAS mutant or KRAS wild-type NSCLC patients in the UK, France and Germany between Oct 2020-Dec 2021	HRQoL using EORTC QLQ-C30; EORTC QLQ-LC13; EQ-5D-5L, incl. EQ-VAS Clinical characteristics Treatment patterns	Q1/2 2022

Study	Design	Outcomes assessed	Estimated* date for analyses
Key: ECOG PS, ECOG performance status; EORTC, European Organization for Research and Treatment of Cancer; EORTC QLQ-C30, EORTC Quality of life questionnaire 30-item core module; EORTC QLQ-LC13, EORTC Quality of life questionnaire lung cancer module; EQ-5D-5L, EuroQol-5 dimension-5 level; EQ-VAS, EuroQol visual analogue scale HRQoL, health-related quality of life; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PRO, patient-reported outcomes (e.g., quality of life and symptom burden assessments)			

B.2.12. Innovation

Sotorasib is a highly innovative and effective therapy that addresses a significant unmet need in patients with *KRAS p.G12C* -mutated NSCLC. It provides a true step-change in current therapy, with additional benefits to patients, clinicians and the healthcare system that are unlikely to be adequately captured in its modelled estimates of cost effectiveness:

- *KRAS p.G12C* -mutated NSCLC occurs in around 13% of NSCLC cases. Patients harbouring this mutation have a very poor prognosis, particularly when they have progressed to second- or subsequent lines of therapy (see B.1.3.1.2).
- In contrast to other, far less common mutation-driven NSCLCs, there are currently no targeted therapies for *KRAS p.G12C* -mutated NSCLC. Treatment options for patients and clinicians are therefore currently limited to non-targeted therapies that are minimally effective and are associated with off-target toxicities that further impair of quality of life (see section B.1.3.1.2). There is therefore a significant unmet need for a highly targeted, effective, tolerable, and convenient treatment that improves clinical outcomes and quality of life for patients with *KRAS p.G12C*-mutated NSCLC.
- Sotorasib is an oral, once daily therapy targeted specifically to inhibit the *KRAS p.G12C*-mutated protein in NSCLC. It is the first $KRAS^{G12C}$ inhibitor to have progressed to regulatory filing since the discovery that KRAS mutations are a key driver of cancer tumours over 40 years ago [6].
- Survival outcomes with sotorasib are significantly superior to those of current non-targeted standard of care therapies (see section B.2.9). UK clinical experts agreed sotorasib provides is a significant improvement in efficacy and also tolerability compared with over current standard of care therapy following progression on front-line therapies.
- As on oral therapy, the administration of sotorasib is more convenient for patients and their carers compared with intravenous administration of the existing standard of care therapies that require hospital visits and several hours away from home. UK clinical experts have noted that a high proportion of patients with *KRAS p.G12C* -mutated NSCLC patients will have been heavy smokers and may have high co-morbidities that impact performance status, and so ease of administration is a consideration in deciding who would be eligible for second- or subsequent-line therapy. Oral administration would also help to free up capacity in hospital intravenous units and services [14].
- Access to sotorasib, as a mutation-targeted therapy, will support the UK Government's commitment to incorporate the latest genomics advances into routine healthcare to improve the diagnosis, stratification and treatment of illness[47], and will support the NHS Long Term Plan in improving cancer outcomes, particularly in this area with such poor prognosis [48].

    - As screening for *KRAS* mutations is now routine [49], sotorasib does not require additional testing beyond the usual diagnostic work-up of NSCLC patients.
- Collectively, sotorasib provides a true step-change in the treatment of patients with *KRAS p.G12C* -mutated NSCLC. It is a highly innovative therapy, and is recognised as such in the UK:
    - Sotorasib was granted an Innovation Passport under the recently introduced Innovative Licensing and Access Pathway [9].
    - Sotorasib has been designated as a Promising Innovative Medicine under the Early Access to Medicines Scheme [10].
    - [redacted].

B.2.13. Interpretation of clinical effectiveness and safety evidence

B.2.13.1 Overview of *KRAS p.G12C*-mutated NSCLC

KRAS p.G12C -mutated NSCLC occurs in around 13% of NSCLC cases. Patients harbouring this mutation have a very poor prognosis, particularly when they progress to second- or subsequent lines of therapy (see section B.1.3.1.2). KRAS mutations are now included in the National Test Directory and moving forward will be routinely screened for as part of the diagnostic work up of patients presenting with suspected NSCLC (see section B.1.3.2.1). However, in contrast to other, far less common mutation-driven NSCLCs, there are currently no targeted therapies for *KRAS p.G12C* -mutated NSCLC. Treatment options for patients and clinicians are therefore currently limited to non-targeted, intravenous cytotoxic therapy regimens that are minimally effective and are associated with off-target toxicities that further impair quality of life (see section B.1.3.1.2). There is therefore a significant unmet need for a highly targeted, effective, tolerable, and convenient treatment that improves clinical outcomes and preserves quality of life for patients with *KRAS p.G12C*-mutated NSCLC.

B.2.13.2 Summary of clinical evidence base

B.2.13.2.1 Efficacy and safety

The CodeBreaK 100 trial is a phase 2, single-arm trial that is accepted as sufficiently robust to demonstrate the efficacy and safety of sotorasib in support of its early, conditional approval by the MHRA.

The primary and secondary endpoints of the CodeBreaK 100 trail provide early evidence that sotorasib is highly effective when used in line with the full anticipated licensed indication as a second- or subsequent line therapy.

- ORR was 37.1%, which UK clinical experts at a recent advisory board held by Amgen, February 2021, considered to be much better than they would expect to see

with existing standard of care docetaxel monotherapy or nintedanib plus docetaxel [14].

- Disease control rate, which is a strong predictor of clinical benefit, was high at 80.6%.
- With a median time to response of only 1.35 months, and duration of response of 10 months, the response to sotorasib was rapid and durable.
- Median PFS was 6.8 months, and median OS was 12.5 months (see section B.2.6.2), which is somewhat greater than typically seen with existing non-targeted standard of care therapies used in these lines of therapy (and similar to that observed with first-line therapies in the Amgen real world studies in the US) (see section B.1.3.1.2.1).
- These data therefore suggest that sotorasib provides a meaningful improvement in efficacy compared with existing non-targeted standard of care therapies.
- As treatment effects were generally similar across pre-specified subgroups (see section B.2.7), the results support the use of sotorasib in line with the full anticipated licensed indication.

Safety data from the CodeBreaK 100 trial indicate that sotorasib is well tolerated, with a very manageable adverse event profile.

- The majority of treatment-related adverse events with sotorasib were mild to moderate; 5 (4%) subjects experienced diarrhoea rated as grade 3 or greater, and 7-8 (6%) subjects experienced grade 3 or greater elevations in alanine and aspartate aminotransferase.
- Nine subjects (7%) discontinued sotorasib due to treatment-related adverse events (see section B.2.10.3).
- Patient reported outcomes data indicate that symptom burden and quality of life while on treatment were stable or marginally improved, with few subjects reporting bothersome side effects (see section B.2.6.2.3.1).
- Given that existing non-targeted standard of care therapies are associated with toxicity, intolerance and quality of life impairment (see section 0), these data indicate that targeted therapy with sotorasib may offer important safety and tolerability advantages over existing non-targeted therapy.
- Of note, UK clinical experts at a recent advisory board held by Amgen, February 2021, agreed the safety and tolerability of sotorasib appears to be superior to that they would expect to see in patients treated with current standard of care docetaxel or nintedanib plus docetaxel [14].

B.2.13.2.2 Comparative evidence vs relevant comparators

The relevant primary comparator for sotorasib based on the existing UK clinical pathway for patients with *KRAS p.G12C* -mutated NSCLC is docetaxel monotherapy; nintedanib plus docetaxel is a secondary comparator (see section B.1.3.3).

Robust indirect treatment comparisons provide plausible early evidence that sotorasib provides clinically meaningful improvements in survival outcomes compared with the relevant comparators

- A formal indirect comparison of relevant sotorasib and docetaxel monotherapy trial data, using the most robust methods possible (MAIC), indicates that sotorasib is significantly superior to docetaxel monotherapy for both PFS ([redacted]) and OS ([redacted]) (see section B.2.9.4.1).
- A supplementary analysis using an alternative approach based on real-world outcomes with standard of care therapies (propensity score weighting analysis) generally confirms these findings (see section B.2.9.4.1).
- In a secondary indirect comparison implemented in the economic model the undiscounted mean OS with sotorasib was [redacted] vs [redacted] with nintedanib plus docetaxel (see section B.2.9.4.2).
- These indirect comparisons are as robust as possible given the available data for sotorasib and the comparators, and support UK clinical experts' opinions that the response rates observed with sotorasib appear to be much better than would be expected with existing standard of care therapies [14].
- There is therefore plausible early evidence that sotorasib provides clinically meaningful improvements in outcomes compared with existing non-targeted standard of care therapies.

B.2.13.3 End of Life criteria

Sotorasib in its full anticipated licensed indication as a second- or subsequent line therapy meets the NICE criteria for an end of life medicine, as demonstrated in Table 20.

Table 20. Sotorasib meets the NICE end of life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	• Large real world evidence studies indicate that that OS with non-targeted 2nd line therapies is <10 months, and with 3rd line therapies is <7 months. • OS with 2nd line docetaxel monotherapy in the SELECT-1 study was 7.9 months [31]. • OS with 2nd line nintedanib plus docetaxel in the LUME-Lung 1 study was 10.9 months [32].	• Section B.1.3.1.2, page 19-21
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	• A robust MAIC indicates sotorasib provides at least an additional [redacted] in median OS compared with docetaxel monotherapy based on available trial data. • The economic model estimates that sotorasib plausibly provides an additional undiscounted mean OS of [redacted] months compared with docetaxel monotherapy and [redacted] months compared with nintedanib plus docetaxel*.	• Document B, section B.2.9.4.1, page 57 • Document B, section B.2.9.4.2, page 60

Key: *Derived from economic model with 20-year time horizon, values undiscounted (see section B.3.3.5 for how comparison of sotorasib vs nintedanib plus docetaxel is implemented); OS, overall survival; PFS, progression-free survival

B.2.13.4 Generalisability and relevance of the clinical evidence base

The efficacy and safety of sotorasib observed in the CodeBreak 100 trial, and its comparative effectiveness in the indirect comparisons, are generalisable to UK clinical practice. These early data support the use of sotorasib in line with its full licensed indication within the UK clinical pathway.

B.2.13.4.1 Patient populations

CodeBreaK 100 recruited adult patients with a mean age of 63 years, confirmed *KRAS p.G12C* -mutated NSCLC, ECOG performance status (PS) 0-1, and with one to three prior lines of therapy. The majority (81%) had prior treatment with combined immunotherapy/chemotherapy (see Table 6), which is reflective of front-line therapies in current UK practice (see section B.1.3.3). UK clinical experts at an Amgen Advisory board considered the population was reflective of patients in practice. Although patients with ECOG PS 2 were not enrolled in the trial, UK clinical experts considered that these patients could be good candidates for treatment with sotorasib based on its favourable adverse event and tolerability profile, and its easier administration, relative to the current non-targeted, intravenous cytotoxic standard of care therapy (docetaxel monotherapy or nintedanib plus docetaxel, where relevant). Patients with ECOG PS 2 may therefore have particular unmet

needs. Patients with squamous histology were under-represented, reflecting the fact that the majority of patients with *KRAS p.G12C* -mutated NSCLC have non-squamous histology. Nonetheless, provided the licensed indication does not preclude its use in patients with ECOG PS 2 or squamous histology, sotorasib should be made available to these patients. As results were generally similar across pre-specified subgroups (see section B.2.7), the results support the use of sotorasib in line with the full anticipated licensed population.

The SELECT-1 trial for docetaxel monotherapy [31] and the LUME-Lung 1 trial of nintedanib plus docetaxel [32], used in the indirect comparisons, recruited patients with similar characteristics, although they were not required to be exclusively KRAS-G12C-mutated patients. It should be noted that, as sotorasib is the first $KRAS^{G12C}$ inhibitor to be developed for regulatory approval, there is a lack of other RCT data in patients with *KRAS p.G12C* -mutated NSCLC. Given that survival is similarly very poor for patients with *KRAS p.G12C*-mutated NSCLC and other patients who are also not candidates for existing targeted therapies (see section B.1.3.1.2.1 and Table 3), these data are the most relevant and suitable to use in the indirect comparisons. Use of immunotherapy front-line was not established at the time of these studies; however, as these studies recruited patients with only one prior line of therapy, compared with one to three prior therapies in CodeBreaK 100 (57% had 2-3 prior therapies – see Table 6), and as survival outcomes clearly worsen with successive lines of existing therapies (see Table 3), the estimated survival benefits of sotorasib in the indirect comparisons may be conservative.

B.2.13.4.2 Intervention

Sotorasib was dosed in the CodeBreaK 100 trial in the same dose regimen as anticipated to be licensed. There is no reason to doubt that the efficacy observed with the dose regimen used in the trial will differ in patients meeting the licensed indication in practice.

B.2.13.4.3 Comparators

The relevant primary comparator for sotorasib based on the existing UK clinical pathway for patients with *KRAS p.G12C* -mutated NSCLC is docetaxel monotherapy; nintedanib plus docetaxel is a secondary comparator (see section B.1.3.3). UK clinicians confirmed that these are the most relevant comparators for sotorasib in its anticipated licensed indication [14].

The CodeBreaK 100 trial was a single-arm trial, which provides no direct comparative data. The indirect comparisons, providing estimates of comparative effectiveness against these comparators, are therefore appropriate and reflect comparisons against the existing non-targeted standard of care therapies that will be displaced by sotorasib in clinical practice.

B.2.13.4.4 Outcomes

The primary endpoint of CodeBreaK 100 to support conditional marketing approval is ORR assessed by blinded independent central review to reduce the risk of bias in this single arm study. This is a compelling measure of antitumor activity showing the proportion of subjects with a response and is reasonably likely to predict clinical benefit. It is of note that response rate was the primary endpoint in studies for several other targeted therapies that received positive recommendations by NICE for use in the Cancer Drugs Fund [46, 59, 60] and for routine commissioning [61, 62].

Key secondary endpoints included time to and duration of response, and also PFS and OS. The indirect comparisons appropriately used the PFS and OS endpoints, as these are key outcomes of concern for patients, and were required as inputs for the economic model (see section B.3.2). As the indirect comparison needed to match the CodeBreaK 100 trial to the SELECT-1 trial, which used investigator assessed PFS, the investigator PFS data from CodeBreak 100 were used in the analysis. As the investigator assessed PFS data were highly consistent with the independent central review assessed PFS data in CodeBreak 100 (see Appendix E), this is an appropriate approach.

B.2.13.4.5 Study design

This single-arm design of the phase 2 CodeBreaK 100 trial was considered appropriate to balance the need for sufficiently robust evidence generation against the need to ensure timely availability of sotorasib in patients with profound unmet clinical needs (see section B.1.3.1.3). Data from this trial have been accepted as sufficient for early, conditional regulatory approval. Evidence of anti-tumour activity from single-arm trials has been sufficient for NICE to make positive recommendations for use of several targeted therapies within the Cancer Drugs Fund [46, 59, 60] and for routine commissioning [61, 62].

Notwithstanding the single-arm, non-randomised design, quality assessment of the CodeBreaK100 trial indicates this to be at low risk of bias, with good external validity. Combined with the evidence from the indirect treatment comparison, using the most robust methods possible to generate comparative effectiveness estimates, there is no reason to doubt that the results of CodeBreaK100 are generalisable to the anticipated use of sotorasib in *KRAS .G12C*-mutated NSCLC patients in UK clinical practice in the UK.

B.2.13.5 Strengths and Limitations of evidence base

B.2.13.5.1 Strengths

The CodeBreak100 trial provides an early indication of the benefits and safety of targeted therapy with sotorasib. The trial recruited patients who are reflective of those in UK clinical practice, and assessed relevant outcomes using the dose regimen anticipated to be licensed in the UK. These data are supplemented with data from indirect comparisons using the most robust methods possible to derive comparative effectiveness estimates against the only

relevant comparators. Collectively, in the context of a disease with no other targeted therapies, high and urgent unmet needs and with limited available data with which to make comparisons, the early evidence in support of sotorasib is compelling, and indicates that sotorasib has a superior efficacy and safety profile compared with the current non-targeted standard of care therapies.

B.2.13.5.2 Limitations

Given the context of a disease with no other targeted therapies, high and urgent unmet needs and with limited available data with which to make comparisons, there are limitations to the current evidence base. CodeBreaK 100 is a single-arm trial with a relatively low sample size, which is associated with uncertainty and the potential for bias. However, the trial employed blinded independent central review to minimise the risk of bias in subjective endpoint assessments, and the trial design was accepted by the MHRA as sufficient to support conditional marketing approval. The trial was not specifically powered for PFS and OS, which are the gold standard assessments of treatment effects in cancer; however, it was powered to show a clinically meaningful ORR, which is a reasonable predictor of clinical benefit, and the PFS and OS data are sufficiently mature and informative to confirm that clinical benefit.

By necessity, the indirect treatment comparison of sotorasib and the relevant primary comparator is based on an unanchored MAIC, which is also associated with uncertainty; however, given the consistency of the results across different MAIC models and an alternative approach to derive an indirect estimate of comparative effectiveness, it seems likely that the uncertainty in the comparative effectiveness estimates relates more to the precision of the estimates rather than their magnitude or direction.

There is limited data for sotorasib or the comparators in patients with squamous histology, and the trial excluded patients with ECOG PS 2, who could particularly benefit from its superior safety profile compared with existing non-targeted standard of care therapies. Given these patients have particular high unmet needs, and provided the licensed indication does not preclude its use, sotorasib should be made available to these patients.

B.2.13.5.3 Forthcoming data to address uncertainty

A number of ongoing studies are being implemented which will address areas of uncertainty in the evidence base for sotorasib and potential comparators is ongoing in clinical trials and real-world evidence studies. Key amongst these is the confirmatory phase 3 RCT, which will provide directly comparative data for sotorasib vs the primary comparator in this appraisal, docetaxel monotherapy, within the next 2 years (see Section B.2.11).

B.2.13.6 Conclusions

KRAS p.G12C -mutated NSCLC is associated with very poor survival. Current treatment is limited to non-targeted intravenous cytotoxic therapies that are minimally effective and associated with significant toxicities. Sotorasib is a highly innovative, first-in-class targeted

therapy, anticipated to receive a conditional marketing authorisation for the treatment of patients with *KRAS p.G12C* -mutated NSCLC following failure of prior therapy. Although data are currently limited to a phase 2 single-arm trial and indirect treatment comparisons, this collective early evidence indicates clearly that sotorasib is highly effective and well tolerated, and provides superior PFS and OS compared with the current non-targeted standard of care comparator therapies. Sotorasib should therefore be made available as early as possible to address the high and urgent unmet needs of these patients.

Sotorasib meets the NICE criteria for consideration as an end-of-life medicine. With confirmatory phase 3 trial data anticipated to be available in the next 2 years, sotorasib may be a candidate for use via the Cancer Drugs Fund.

B.3. Cost effectiveness

Summary of economics

- Sotorasib is a novel and clinically effective treatment option for KRAS p.G12c NSCLC which significantly improves life-years and QALYs compared with docetaxel and nintedanib + docetaxel
- The primary analysis is well-aligned to the decision problem and reflective of UK clinical practice. Results are generated based on an MAIC adjusted comparison via the docetaxel arm of SELECT-1 and an indirect comparison of nintedanib plus docetaxel via LUME LUNG 1
- The most clinically plausible extrapolations of PFS and OS data were selected for the base case analyses and extensive scenario analyses were presented with only a small impact to the results of the analyses
- The base case modelling approach, including structure, costs included, and utility values applied, is consistent with those accepted in previous TAs for treatments of NSCLC
- In the primary comparison base case analysis the ICER for sotorasib versus docetaxel was £47,146 per QALY gained; results of the alternative analysis using real-world data versus a basket of chemotherapy comparators reduced to £39,773 per QALY gained
- In the secondary comparison, using the list price of nintedanib, the ICER for sotorasib versus nintedanib + docetaxel reduced to £35,779 per QALY gained
- Sotorasib meets the end-of-life criteria and likely represents a cost-effective use of NHS resources.

B.3.1. Published cost-effectiveness studies

A systematic literature review (SLR) was conducted to identify published studies for evaluating cost-effectiveness, costs and resource use and health-related quality of life for treatments in non-small-cell lung cancer (NSCLC) relevant to the decision problem.

Full details on the methodology and findings of the SLR, including search terms, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram and outcomes are detailed in the appendices to this report for cost-effectiveness studies (Appendix G), costs and resource use (Appendix I) and health-related quality of life (HRQOL; Appendix H: Health-related quality of life studies).

No published economic studies were identified in the SLR which examined the cost-effectiveness of interventions for the management of patients with *KRAS p.G12C* mutation-positive locally advanced or metastatic NSCLC or for KRAS mutation in general.

B.3.2. Economic analysis

Given that there were no identified cost-effectiveness studies published in NSCLC with the *KRAS p.G12C* mutation (or KRAS mutation in general) following the SLR, a *de novo* economic model was constructed for this submission.

To inform development of the *de novo* economic evaluation a targeted search of economic models submitted to NICE for previously treated NSCLC consistent with the relevant treatment pathway for sotorasib was conducted. The key features of these economic models are presented in Appendix G (Table 1). All previous economic evaluations used a partitioned survival structure, with a time horizon ranging from 12 to 25 years and utilities derived from EQ-5D data obtained from the relevant clinical trials. Costs and resource use were typically obtained from standard UK cost sources including British National Formulary (BNF), electronic market information tool (eMIT), Monthly Index of Medical Specialities (MIMS) and National Health Service (NHS) Reference Costs [80-82]. The model cycle length used varied between 7 to 30 days with the most frequently used cycle length of 7 days. All models applied health effects using quality-adjusted life years (QALYs), applied a 3.5% discount rate to costs and QALYs and used an NHS and Personal Social Services (PSS) perspective, as per NICE guidelines [83]. A half-cycle correction was applied in all models.

B.3.2.1 Patient population

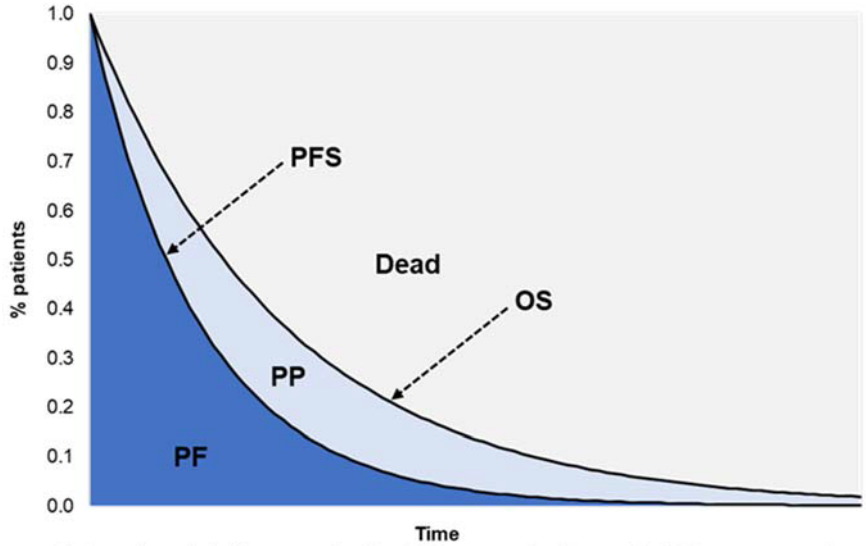
In line with the final scope of this appraisal and the anticipated marketing authorisation, the patient population for the economic analysis is adult patients with *KRAS p.G12C*-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy [12, 84]. The population matches the Phase II NSCLC cohort of the clinical trial CodeBreaK100 described in Section B.2.3.

Subgroups analyses were not included in the cost-effectiveness analysis given there were no subgroups observed with substantially different efficacy compared to the whole population (Section B.2.7). Further, the population enrolled in CodeBreaK100 is relatively small (N = 126) which would limit the sample size available and interpretability of any subgroup analyses.

B.3.2.2 Model structure

The economic evaluation was developed using a cost-utility framework in Microsoft Excel®. A partitioned survival analysis was used based on three distinct health states (Figure 9): progression-free, progressed disease and dead. All patients entered the model in the progression-free state and were at risk of progression of disease or death. Transitions to the death state occurred from either the progression-free or progressed disease health states. Death was an ‘absorbing state’, where once entered, patients reside for the remainder of the model time horizon.

Figure 9: Partitioned survival analysis model



OS, Overall survival; PF, progression free; PFS, progression-free survival; PP, post progression

This model structure is fully aligned with the primary objectives of treatment in oncology and NSCLC, namely avoiding disease progression and prolonging life. Furthermore, the structure and health states selected are typical of modelling in oncology and have been used in previous NSCLC technology appraisals conducted by NICE (Appendix G: Table 1) [51, 60, 62, 85-88].

The model contains the three most relevant disease related health states from a patient, clinician, and NHS perspective:

- *Progression free:* Patient disease is in a stable or responding state and not actively progressing. Patients in this state are assumed to incur costs associated with treatment, administration, medical management of the condition and the management of Grade 3/4 adverse events.

- *Progressed:* Patient disease has progressed, and patients are assumed to receive subsequent treatment. This health state is associated with acquisition and administration costs of subsequent therapies as well as costs of disease management.

- *Death:* This is an absorbing state; a cost for palliative care is assumed upon death.

The proportions of patients in each health state at the beginning of each model cycle are calculated from the progression-free survival (PFS) and overall survival (OS) survival functions from relevant clinical trials as follows, where PF(t) is the proportion of patients who are progression-free at time (t), Dead(t) is the proportion of patients who are not alive at time (t) (1 – Overall survival) and PP is the proportion of patients who are not progression-free and who are still alive at time (t). In the model, all patients start treatment in the progression free health state and on treatment:

$$PF(t) = PFS(t)$$

$$Dead(t) = 1 - Overall\ Survival(t)$$

$$PP(t) = OS(t) - PF(t)$$

The estimated time on treatment for each treatment in the analysis was used to inform acquisition costs and related administration costs. Additional costs included in the analysis include disease management costs per health state, subsequent treatment costs and terminal care costs. As discussed in Section B.1.3.2.1, costs associated with genetic mutation testing are not required to be captured in the model as *KRAS p.G12C* testing is routinely funded as part of panel testing at diagnosis. It is therefore assumed that all patients entering the model have a *KRAS p.G12C* mutation-positive status.

The progression-free health state typically reflects a relatively higher HRQOL associated with disease before progression, where patients are receiving benefit from an active treatment, whereas the progressed disease state is designed to capture the relatively poor HRQOL following disease progression. However, as previous studies have shown NSCLC patients to have markedly decreased utilities towards the end of life, the measurements included in the model base case was informed by a time-to-death analysis [89]. This approach has been used in previous NICE TAs [85, 87, 88, 90] and was considered by UK clinical experts to better reflect the experience of their patients with NSCLC.

Time-to-death sub-health states were therefore implemented to capture patients' quality of life as a function of how much lifetime patients had left until they eventually died as predicted in the model. The use of time-to-death sub-health states was implemented considering four health states: less than 1 month before death, 1–3 months before, 3–6 months before and more than 6 months before death.

Features of the de novo analysis

The analyses were conducted from an NHS and PSS perspective in England and Wales and are consistent with NICE guidelines [83] The model uses a 7-day cycle length, with a half-cycle correction applied and a time horizon of 20 years. This aligns with the maximum life expectancy of the cohort predicted by parametric survival analysis and was considered appropriate by clinically experts given that it is highly unlikely for patients with NSCLC with the *KRAS p.G12C* mutation with advanced or metastatic disease to survive beyond this time period. The impact of the selection of the time horizon on results is explored in sensitivity analysis. A discount rate of 3.5% per annum was applied for costs and benefits. The perspective chosen, time horizon assessed, and the discount rates used are all in line with the NICE reference case. A summary of the key features of the economic analysis is presented in Table 21.

Table 21: Features of the economic analysis

	Current appraisal	
Factor	**Chosen values**	**Justification**
Model structure	Partitioned survival analysis with 3 health states: progression-free, progressed and death	Reflects the three most relevant disease health states which capture the clinical events experienced by patients with NSCLC. The structure is typical of NSCLC modelling and has been used in several previous NICE appraisals (See Appendix G: Table 1).
Time horizon	20 years	20 years is considered sufficiently long so that most patients (> 99%) would have died by the end of the model time horizon. The model is therefore able to reflect all differences in costs and outcomes in line with the NICE reference case.[83]
Cycle length and half-cycle correction	1 week	The weekly model cycle length was based on clinical trial measurement points and was judged short enough to ensure accuracy in model calculations.
Comparator(s)	Primary comparator: Docetaxel monotherapy Secondary comparator: Nintedanib + docetaxel	Docetaxel monotherapy is recognised in the NICE lung cancer guideline (NG122)[16] and pathway [13] as a key second- and subsequent-line option in NSCLC across non-squamous and squamous disease and across PD-L1 expression and tumour proportion scores. In addition, nintedanib + docetaxel may also be considered a relevant comparator for patients with adenocarcinoma histology. A UK advisory board conducted by Amgen confirmed the suitability of comparators assessed in this appraisal (Section B.1.3.3) [14].
Source of utilities	CodeBreaK100 and published literature	Utility values were derived from EQ-5D-5L data collected in CodeBreaK100, which were mapped to EQ-5D-3L using the UK cross walk tariff, in line with the NICE reference case [91]. Where not available, utility values were sourced from relevant literature and previous NICE appraisals in NSCLC (TA428 and TA484)[86, 87]
Source of costs	UK standard costs sources: NHS Reference costs and eMIT/MIMs	Drug costs: Public list price of treatments were used, in line with the NICE reference case.[80-82] Other costs: EMIT, MIMS, PSSRU, NHS Reference Costs; consistent with NICE reference case.[80, 81, 83]
Health effects measure	QALYs	Consistent with NICE reference case[83]
Discount rates	3.5%	Consistent with NICE reference case[83]. A scenario is explored using reduced discount rates (1.5%) consistent with ongoing proposals in the NICE methods review.[92]
Perspective	NHS/PSS	Consistent with NICE reference case[83]
Key: eMIT, electronic market information tool; MIMS, Monthly Index of Medical Specialities; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; NSCLC, non-small-cell lung cancer; PSS, Personal Social Services; QALY, quality-adjusted life year.		

B.3.2.3 Intervention technology and comparators

Intervention

Sotorasib (LUMYKRAS™) is the first *KRAS p.G12C* inhibitor to be submitted for marketing authorisation. It is a once-daily oral therapy (960 mg, administered as eight 120 mg tablets), anticipated to be licensed in the UK by the Medicines and Healthcare products Regulatory Agency (MHRA) [redacted] for use as *monotherapy the treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic non-small cell lung cancer (NSCLC) previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated* (Section B.1.2).

According to the draft Summary of Product Characteristics [12], sotorasib is administered until disease progression or until no further clinical benefit is expected, which is aligned with the phase 2 registrational trial, CodeBreaK100 [52, 56].

Comparators

The current NHS clinical pathway for treating patients with *KRAS p.G12C*-mutated NSCLC is detailed in Section B.1.3.2 and B.1.3.3. The main comparator for the economic analysis is docetaxel monotherapy, with the rationale for selection presented in Section B.1.1, Table 1. The economic analysis also includes a comparison of sotorasib with nintedanib + docetaxel combination treatment, on the basis that some patients eligible for docetaxel monotherapy in the UK may also receive the combination treatment. Clinical expert opinion obtained at a UK Advisory Board in February 2021 indicated that the use of combination therapy was highly variable across the UK and is thus considered as a secondary comparator in the economic analysis (Table 21) [14].

The dosing and administration frequencies for these comparators were implemented in the model in line with their marketing authorisations and UK clinical practice.

In the base case analysis, the clinical effectiveness of sotorasib versus docetaxel is based on an MAIC using CodeBreaK100 trial data and the reference arm of the SELECT-1 randomised controlled trial (docetaxel monotherapy). Full details of the MAIC methodology and outcomes are presented in Section B.2.9 and Appendix D. A comparison of sotorasib versus nintedanib plus docetaxel is also conducted using relative efficacy measures for the adenocarcinoma subgroup of the NSCLC trial LUME LUNG-1, using the methodology described in Section B.3.3.5).

In addition to these analyses, a supplemental, alternative approach was explored using real-world data from the Flatiron Health database, which may be used as confirmatory validation of the primary MAIC of sotorasib vs docetaxel monotherapy. As described in Section B.2.9.3.1 and Appendix D, a propensity score analysis was conducted to compare the benefit of sotorasib versus the current standard of care treatments in this dataset.

B.3.3. Clinical parameters and variables

B.3.3.1 Population baseline characteristics

The core characteristics used in the model to define the patient cohort were obtained directly from CodeBreaK100 and are presented in Table 22. These were considered appropriately generalisable and reflective a population anticipated to be treated with sotorasib in UK clinical practice.

Table 22: Patient baseline characteristics

Patient characteristic	CodeBreaK100	Source
Age at baseline (years)	62.9	CodeBreaK100 CSR, Table 9.2 [55]
Gender (female)	50%	CodeBreaK100 CSR, Table 9.2 [55]
Weight (kg)	71.1	CodeBreaK100 CSR, Section 9.3 [55]
Body surface area (m^2)	1.81	Calculation [Mosteller formula [93]]
Key: CSR, clinical study report.		

B.3.3.2 Overview of modelled efficacy

Efficacy inputs for the progression-free survival and overall survival of sotorasib were based on outcomes from CodeBreaK100 (Data cut: 1 December 2020) [55]. As CodeBreaK100 is a single-arm trial, an SLR was conducted to identify studies which reported the efficacy of comparator treatments in the relevant population (see Appendix D). The SLR identified only one RCT (SELECT-1 [31]) that provided sufficient PFS and OS data for docetaxel monotherapy (the primary comparator) in a population of patients with *KRAS*-mutated NSCLC (including *G12C* and non-*G12C* mutations).

An indirect treatment comparison was necessary to compare the efficacy outcomes from CodeBreaK100 and the SELECT-1 clinical trial and an MAIC analysis was used to adjust for differences in prognostic patient characteristics between the two studies (Section B.2.9.3.1). The Kaplan–Meier plots for CodeBreaK100 were adjusted using the weights derived from the MAIC applied to the docetaxel outcomes from SELECT-1.

As the maximum available follow-up from the CodeBreaK100 clinical trial was shorter than the modelled time horizon it as necessary to extrapolate OS and PFS outcomes to populate the partitioned survival analysis. The approach taken was informed by and is ultimately consistent with recommendations from the NICE Decision Support Unit (DSU; Technical Support Document [TSD] 14) [94].

Parametric curves were fitted to the MAIC weighted time-to-event data using an extensive analysis framework which assessed multiple distribution functions, including restricted versus unrestricted joint parametric models and independent survival models. The parametric functions were assessed based on goodness-of-fit statistics Akaike information criterion (AIC) and Bayesian information criterion (BIC) and visual inspection of the fitted curves to the observed clinical trial data. Furthermore, for the given MAIC weighted data, the proportional hazards assumption and the presence of accelerated failure time was assessed

using log cumulative hazard plots, Schoenfeld residuals and quantile-quantile (Q-Q) plots. No one aspect of the analysis per se was considered more important that the other and final curve selection was based on the totality of the evidence.

Finally, the extrapolations were assessed for clinical plausibility of long-term outcomes based on clinical expert feedback and alternative data sources.

As described above, a supplemental, alternative approach was explored using real-world data from the Flatiron Health database (see Section B.2.9.3.1 and Appendix D) and which may be used as confirmatory validation and to support the robustness of the results generated in the MAIC.

For the secondary comparison of sotorasib versus nintedanib plus docetaxel, no direct comparative data was available for nintedanib plus docetaxel in a KRAS mutation-positive population. The relative treatment effect of nintedanib plus docetaxel versus docetaxel monotherapy was therefore obtained from the adenocarcinoma subgroup of the LUME Lung 1 trial [32] which was identified as the pivotal Phase 3 trial informing the efficacy of nintedanib plus docetaxel and was used as the primary evidence source in NICE TA347 [51]. To facilitate a comparison of sotorasib versus nintedanib plus docetaxel in the economic evaluation, the relative treatment effect of nintedanib plus docetaxel vs docetaxel was applied to the SELECT-1 modelled docetaxel curve (Section B.3.3.5).

B.3.3.3 Sotorasib versus docetaxel

The hazard ratio and corresponding 95% confidence interval (CI) based on the MAIC (presented in Section B.2.9.4.1) are summarised in Table 23. As discussed previously. when selecting the variables for MAIC adjustment, both the precision and the absence of bias were considered. Whereas in terms of bias it would have been desirable to adjust for all available variables, the impact on the effective sample size was significant: maximal variable selection would have been limited in terms of robust parametric curve fitting. To balance precision versus bias, for the base case analysis, solely variables which were considered important for prognosis (informed by clinical expert opinion) were included. Therefore, PD-L1 expression at baseline was not considered, as despite being a relevant predictor for treatment with PD-L1, it was not an important predictive factor for treatment with sotorasib or docetaxel. In the selected base case model, the effective sample size was 108.8 and 106.1 for OS and PFS, respectively. The unadjusted comparison and the MAIC with all available covariates were explored in sensitivity analysis.

Across all models explored, the results demonstrated a statistically and clinically significant decrease in event rates for both OS and PFS with sotorasib vs docetaxel (Table 23).

Table 23: Matching-adjusted indirect comparison hazard ratio estimates

Analyses	N (OS - PFS)	ESS (OS - PFS)	HR for OS (95% CI)	HR for PFS (95% CI)
Unadjusted	126	126	[redacted]	[redacted]
MAIC Model: 'all variables of prognostic importance' (base-case)	123/ 121	108.8/ 106.1	[redacted]	[redacted]
MAIC Model: 'all available covariates' (scenario analysis)	98/ 96	53.3/ 53.1	[redacted]	[redacted]

Key: CI, confidence interval; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival; PFS, progression-free survival

Overall survival

The weighted Kaplan-Meier (KM) plot for OS from the base case MAIC analysis is reproduced for convenience in Figure 10. Consistent with expectations, this plot demonstrates the relative improvement of sotorasib OS when balancing the prognostic baseline characteristics.

Figure 10: Kaplan-Meier OS plot for docetaxel (SELECT-1) and sotorasib (CodeBreaK100) using adjusted and unadjusted MAIC

Standard parametric distributions were fitted to the adjusted time-to-event data with joint fitting (restricted and unrestricted models) and independent fitting conducted using the statistical software R (ver. 4.0.3) with the flexurv packages. The parametric distributions included exponential, Weibull, Gompertz, generalized gamma, lognormal and loglogistic.

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 24). For individually fitted curves, the lognormal distribution was the best statistically fitting curve with the lowest AIC and BIC across both sotorasib and docetaxel and the relative performance of each distribution was similar between arms. As a result, jointly fitted survival models (either restricted or unrestricted) were considered the most appropriate since they can reduce uncertainty due to the estimation of fewer parameters and the use of a larger data set.

For the jointly fitted curves, AIC and BIC indicate that the best fitting curve for both the restricted and unrestricted models was the lognormal followed by the generalised gamma and log-logistic models. There was a notable deterioration in the performance of other distributions based on the statistical AIC and BIC criteria. For the best-fitting distributions, AIC and BIC consistently favoured the restricted versus unrestricted joint fits.

Table 24: Goodness of fit statistics for independent and jointly fitted models

Model	Independent fit – sotorasib		Independent fit - docetaxel		Joint fit (unrestricted)		Joint fit (restricted)	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	417.7	420.5	1209.7	1213.2	1627.4	1635.2	1627.4	1635.2
Gompertz	417.5	423.1	1211.4	1218.5	1628.9	1644.6	1629.4	1641.1
Weibull	414.1	419.7	1209.6	1216.7	1623.7	1639.5	1624.1	1635.9
Generalized Gamma	411.5	419.9	1194.6	1205.2	1606.0	1629.7	1602.5	1618.3
Loglogistic	412.4	418.0	1196.3	1203.4	1608.7	1624.4	1607.0	1618.8
Lognormal	410.1	415.8	1192.8	1199.9	1602.9	1618.7	1601.1	1612.9

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion
Note: Underlined values indicate the best statistically fitting parametric distribution

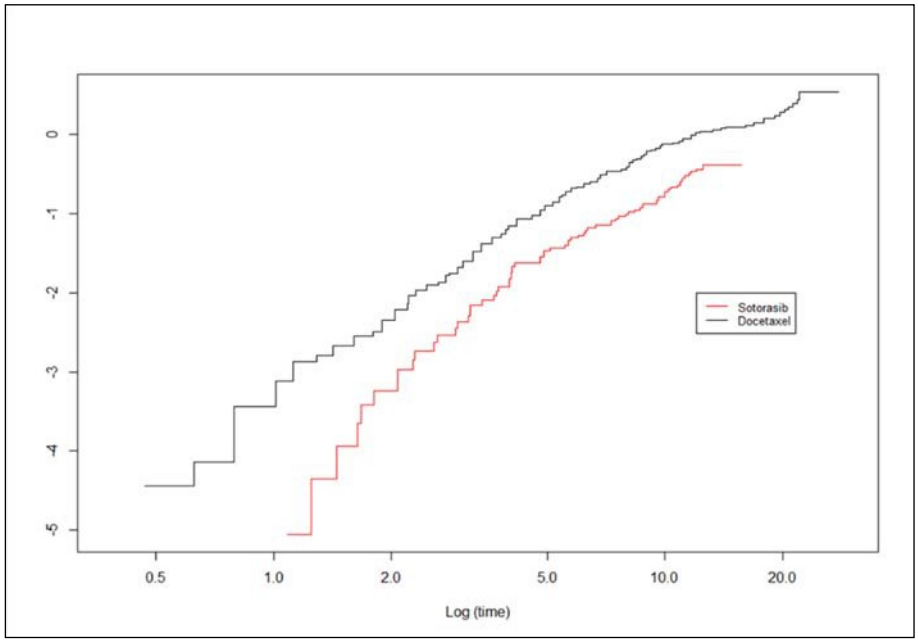
Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

To further inform the most appropriate distribution to extrapolate OS, the proportional hazards assumption and the presence of accelerated failure time was assessed using log cumulative plots, Schoenfeld residuals and QQ plots.

The assumption of proportional hazards between the two datasets was assessed using the log-cumulative hazards plot (Figure 11) and the Schoenfeld residuals plot (Figure 12). The log-cumulative hazards and the Schoenfeld residuals plot for sotorasib and docetaxel indicated that the proportional hazards assumption may be considered valid. However, a slight convergence in the first few months of the log-cumulative hazard plot was observed with a kink at the 5 months mark and a different slope apparent thereafter (Figure 11). Likewise, with respect to the Shoenfeld residuals, at most times the point-wise confidence intervals included zero with the exception at around two months, where the non-significance was borderline (Figure 12). Considering the slope observed and the sample size used in the analysis, the validity of the proportional hazard's assumption was considered uncertain but not strongly violated.

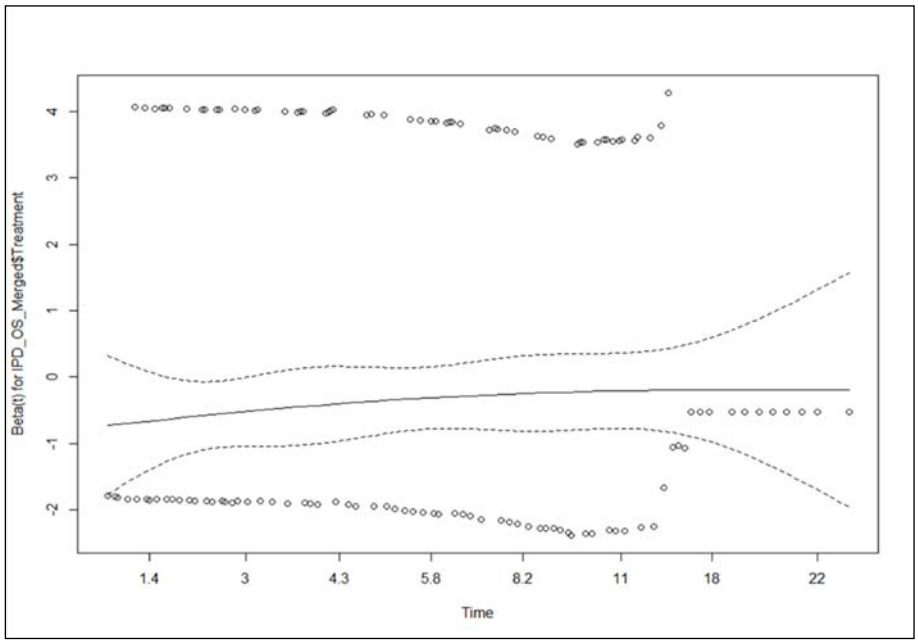
The presence of accelerated failure time was assessed using a QQ plot. This demonstrated an almost perfect straight line, indicating the use of accelerated failure time was valid (Figure 13).

Figure 11: Log-cumulative hazards plot for OS using base case MAIC



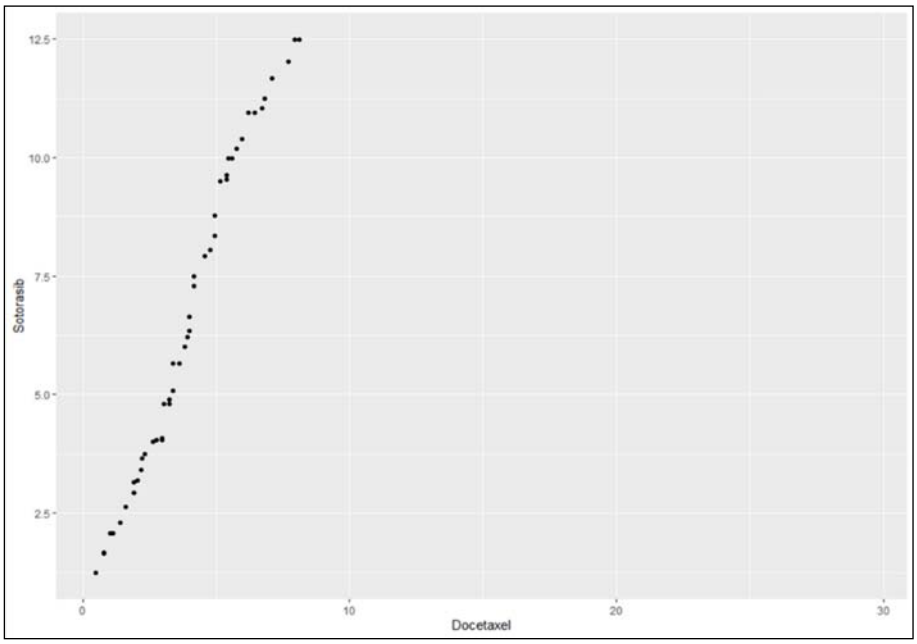
Key: MAIC, matching-adjusted indirect comparison; OS, overall survival.

Figure 12: Schoenfeld residuals plot for OS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; OS, overall survival.

Figure 13: Q-Q plot for OS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; OS, overall survival; Q-Q, quartile-quartile.

Visual Inspection to Observed Data

A plot of jointly fitted parametric distributions fitted to sotorasib and docetaxel using the base case MAIC adjusted Kaplan–Meier curves is shown below (Figure 14). The restricted joint models are presented given the superior statistical fits observed previously. Visual inspection of the docetaxel plot indicated that the Weibull, Gompertz and exponential distributions overestimated OS in the early periods (up to 14 months) with consistent underestimation of OS after this timepoint (Figure 14). Similarly, the sotorasib plot indicated that the Weibull and Gompertz plots underestimated OS up to 2 months and were the most conservative OS estimates for the long-term projections (Figure 14). These findings are consistent with the AIC and BIC results previously presented.

Visual inspection of the best statistically fitting distributions generally indicated that the extrapolated data matched the Kaplan–Meier plots well and captured the longer term shape of the survival function (Figure 14). In all cases sotorasib was shown to improve long-term OS versus docetaxel.

Figure 14: OS KM for sotorasib and docetaxel from base case MAIC with parametric functions fitted

Key: KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival.

Base Case Distribution Selection

In summary, the goodness of fit statistics indicated that the lognormal approach to OS extrapolation was the best statistically fitting distribution, with the Weibull, Exponential and Gompertz performing relatively poorly. There was some evidence to suggest that the proportional hazards assumption did not hold true and, therefore, the application of a treatment effect was not considered appropriate. However, the QQ diagnostic plot clearly demonstrated that an accelerated failure time model was valid, which was supported by the performance of these restricted models and the visual inspection versus the observed data. Based on these conclusions the jointly fitted (restricted) log-normal was considered to be the most appropriate approach for the base case analysis. The 2nd and 3rd best performing distributions (jointly fitted [restricted] generalised gamma and log-logistic, respectively) were considered for sensitivity analyses.

Clinical Plausibility

The clinical plausibility of the parametric models used in the economic analysis were evaluated by considering the predicted OS landmark results at timepoints of 1-year, 5-years and 10-years, and the shape of the underlying hazard function was assessed. In addition, the analyses based on the external Flatiron real-world data (Section B.3.3.4) were also used as confirmatory validation (Section B.3.9).

The OS predictions for the joint fitting (restricted) models are presented in Table 25 below.

Table 25: OS predictions for models using joint fitted (restricted model)

	Exp	Gompertz	Weibull	GG	Loglogistic	Lognormal
Sotorasib						
1 Year	55.1%	55.1%	54.4%	54.1%	53.7%	54.2%
5 Years	5.1%	5.0%	2.2%	11.3%	9.3%	9.6%
10 Years	0.3%	0.2%	0.0%	4.0%	3.5%	2.8%
Docetaxel						
1 Year	40.2%	40.2%	40.0%	37.9%	36.7%	37.8%
5 Years	1.0%	1.0%	0.3%	5.6%	4.9%	4.2%
10 Years	0.0%	0.0%	0.0%	1.7%	1.8%	1.0%
Key: Exp, exponential; GG, generalized gamma; OS, overall survival						

Clinical experts consulted by Amgen considered docetaxel survival predictions at 5-years of approximately 5% to be reasonable in this population and would expect a small proportion of patients to remain alive at the 10-year landmark. Although it was acknowledged that patients in clinical practice could perform slightly worse, the more pessimistic curves presented (exponential, Gompertz and Weibull) were considered to underestimate the long-term survival and did not reflect clinical experience.

The base case log-normal model was determined to provide clinically valid projections of docetaxel and was well-aligned with clinical expectation at the 5-year (4.2%) and 10-year (1.0%) landmarks. Furthermore, the projections of sotorasib at 5-years (9.6%) and 10-years (2.8%) were considered reasonable given the observed response rate, duration of response and survival data available from CodeBreaK100, as well as the ability to receive more effective subsequent therapies.

Finally, the clinical plausibility of the hazard function shape was assessed. The exponential (constant hazard), Weibull (logarithmic increase) and Gompertz (exponential increase) were not considered to reflect hazard of the population in NSCLC whereas the log-normal, generalised gamma and log-logistic (increase to peak within 6-9months with subsequent decline over time) were considered appropriate by clinical experts. This was rationalised based on the relatively high (and increasing) risk reflecting patients with a poor prognosis and non-responders early in the modelled time horizon, followed by a decreasing risk for patients who respond to treatment and have an improved relative prognosis as the time horizon progresses.

In conclusion, the base case selection of the jointly fitted (restricted) log-normal distribution was considered to be clinically valid and reflects the expected survival of the population under consideration.

Progression-free survival

The same approach as used for OS was repeated to determine the appropriate distribution to assess PFS in the portioned survival model. The weighted KM plot for OS from the base case MAIC analysis is presented in Figure 15.

Figure 15: Kaplan-Meier PFS plot for docetaxel (SELECT-1) and sotorasib (CodeBreaK100) using adjusted and unadjusted MAIC

Standard parametric distributions were fitted to the adjusted time to event data with joint fitting (restricted and unrestricted models) and independent fitting conducted using the statistical software R (ver. 4.0.3) with the flexurv packages. The parametric distributions included exponential, Weibull, Gompertz, generalized gamma, lognormal and loglogistic.

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 26). For individually fitted curves, the AIC and BIC both indicated that the lognormal distribution provided the best statistical fit for sotorasib, whereas the generalised gamma performed the best for docetaxel. However, across both distributions the AIC and BIC were not meaningfully different with little separating the two. Given this, and consistent with the approach taken for OS, jointly fitted survival models (either restricted or unrestricted) were considered more appropriate to reduce uncertainty through the estimation of fewer parameters and the use of a larger data set.

For the jointly fitted models, the AIC indicates that the generalised gamma distribution is the best performing, whereas the BIC indicates that the lognormal provides the best statistical fit to the observed data, although again differences are minor. In this instance, the BIC statistic was preferred as its use mitigates the risk of overfitting statistical noise in the tails of the observed distributions. Similar to the conclusions from the OS survival analysis, there was a notable deterioration in the performance of the Exponential, Weibull and Gompertz distributions. For the best-fitting distributions, BIC consistently favoured the restricted versus unrestricted joint fits.

Table 26: Goodness of fit statistics for independent and jointly fitted models

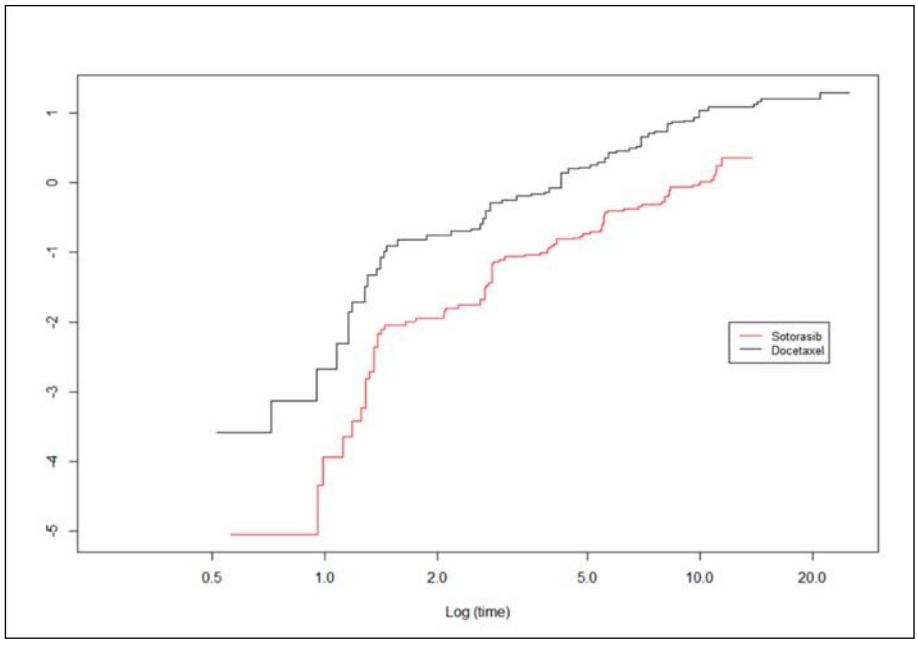
Model	Independent fit – sotorasib		Independent fit - docetaxel		Joint fit (unrestricted)		Joint fit (restricted)	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	503.9	506.7	1166.5	1170.0	1670.4	1678.3	1670.4	1678.3
Gompertz	503.5	509.1	1166.9	1174.0	1670.4	1686.1	1672.1	1683.9
Weibull	499.5	505.2	1160.6	1167.7	1660.1	1675.9	1659.3	1671.1
Generalized Gamma	494.3	502.7	1099.5	1110.1	1593.8	1617.4	1595.5	1611.3
Loglogistic	496.6	502.2	1113.5	1120.6	1610.1	1625.8	1610.3	1622.1
Lognormal	492.9	498.5	1105.7	1112.8	1598.6	1614.4	1598.8	1610.6

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion
Note: Underlined values indicate the best statistically fitting parametric distribution

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

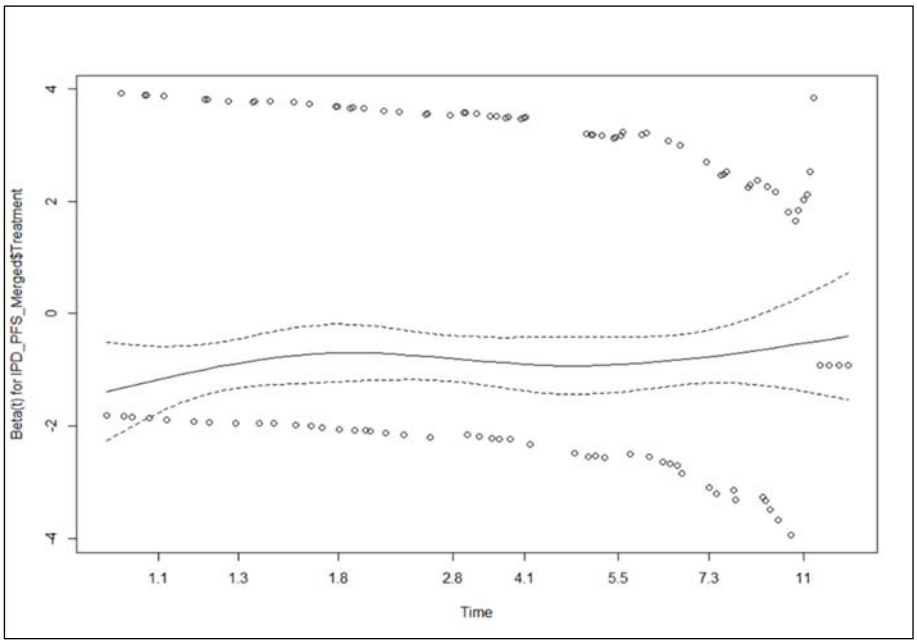
Similar to OS, the assumption of proportional hazards between the two datasets was assessed using the log-cumulative hazards plot (Figure 16) and the Schoenfeld residuals plot (Figure 17). The log-cumulative hazards and the Schoenfeld residuals plot for sotorasib and docetaxel indicated that the proportional hazards assumption is unlikely to hold: the log-cumulative hazards plot demonstrated the convergence of both curves in the first 2 months, which diverges before 3 months and then remains parallel beyond 4 months. Likewise, the confidence bands of the scaled Schoenfeld residuals did not include zero for the majority of the time horizon.

Figure 16: Log-cumulative hazards plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival.

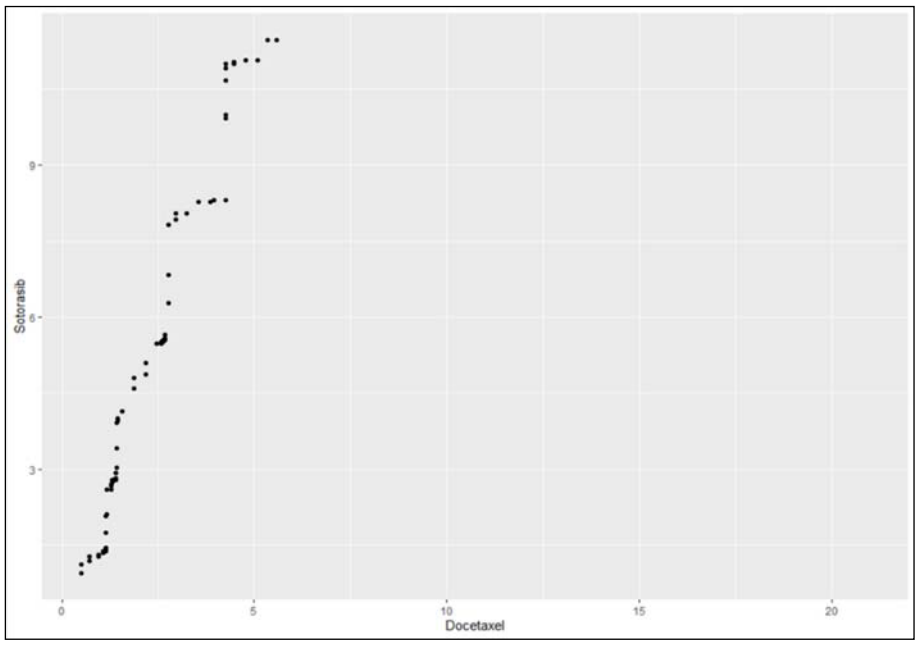
Figure 17: Schoenfeld residuals plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival.

The Q-Q plot, however, indicated that accelerated failure time assumption was sufficiently valid with a straight-line trend clearly discernible (Figure 18).

Figure 18: Q-Q plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; Q-Q, quartile-quartile.

Visual Inspection to Observed Data

A plot of jointly fitted parametric distributions fitted to sotorasib and docetaxel using the base case MAIC adjusted Kaplan–Meier curves is shown below (Figure 19). Visual inspection of the docetaxel plot indicated that the Weibull, Gompertz and Exponential distributions overestimated PFS in the early periods (up to 12 months) with underestimation of PFS after this timepoint (Figure 19). Both the lognormal and log-logistic model fit the data well and the generalised gamma, although performing well on statistical tests, shows a slight underestimation for docetaxel between 6 and 12 months. Generally. these findings are consistent with the AIC and BIC results previously presented.

Figure 19: PFS KM for sotorasib and docetaxel from base case MAIC with parametric functions fitted

Key: KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, Progression Free Survival

Visual inspection of the best statistically fitting distributions indicated that the extrapolated data matched the Kaplan–Meier plots well and captured the longer-term shape of the survival function (Figure 19). In all cases sotorasib was shown to improve long-term PFS versus docetaxel.

Base Case Distribution Selection

In summary, although selecting the most appropriate distribution to model PFS was less clear than OS, the goodness of fit statistics indicated that the lognormal approach to PFS extrapolation was the best statistically fitting distribution, with the Weibull, Exponential and Gompertz performing relatively poorly. Further, the diagnostic plots suggest that the proportional hazards assumption is likely to be violated and that an accelerated failure time model is appropriate.

One factor adding to the difficulty of fitting a parametric curve to PFS was that PFS data were not being collected at a truly continuous level. As progression was not continuously assessed, but only measured at certain points of times, the Kaplan–Meier curves were less smooth than those for OS. However, the Kaplan–Meier curves (Figure 7 and Figure 11), did not provide justification for assuming different parametric functions or for fitting curves independently.

Therefore, PFS was modelled based on a jointly fitted restricted model with log-normal distribution which is consistent with the distribution selected for OS and supported by the visual inspection versus the observed data. The 2nd and 3rd best performing distributions (jointly fitted [restricted] generalised gamma and log-logistic, respectively) were considered for sensitivity analyses.

Clinical Plausibility

The clinical plausibility of the parametric models used in the economic analysis were evaluated by considering the predicted PFS landmark results at timepoints of 1-year, 3-years and 5-years, and the shape of the underlying hazard function was assessed.

UK clinical experts consulted by Amgen considered docetaxel and sotorasib projections based on the selected log-normal distribution to be appropriate, clinically valid and reflect expected survival of the population under consideration. Furthermore, similar to the conclusions on the hazard function shape for OS, the clinical experts considered a non-monotonic hazard function was appropriate to model long-term PFS.

Sensitivity analysis for survival outcomes

Based on the above analyses, the selected base case parametric model for both OS and PFS was the jointly fitted (restricted) lognormal distribution utilising the MAIC model accounting for all variables of prognostic importance. Multiple sensitivity analyses were conducted to assess alternative survival modelling and explore the impact of uncertainty related to long-term survival estimation. In addition to this, an alternative approach based on real-world data from the Flatiron Health database was explored to provide confirmatory validation and to support the robustness of the results generated in the MAIC (Section B.3.3.4).

The additional analyses included the following:

- Survival analysis based on unadjusted comparison of CodeBreaK100 and SELECT-1 (i.e., naïve comparison)
- Alternative MAIC model using all available covariates (see Section B.2.9.4.1)
- Exploration of 2nd and 3rd best performing distributions for OS and PFS (i.e., generalised gamma; log-logistic joint [restricted] distributions)
- Exploration of joint (unrestricted) lognormal PFS parametric distribution (best fitting unrestricted model based on BIC criteria)

Details of the alternative approaches being used in the sensitivity analysis are presented in Appendix N.

B.3.3.4 Alternative comparative efficacy based on Flatiron real-world data

Given the limitations of the SELECT-1-based MAIC approach discussed in Section B.2.9.5, an alternative approach to quantify comparative effectiveness of sotorasib versus docetaxel was conducted. Here, patients with a *KRAS* mutation of the Flatiron real-world cohort being treated with a basket of standard of care chemotherapy were used as a proxy for patients with a *KRAS p.G12C* mutation undergoing treatment with docetaxel.

Restricting the population to patients treated with docetaxel was not feasible due to the significant reduction in sample-size: of the 206 patients included, only 21 patients were on docetaxel monotherapy. Further, limiting the population to specifically patients with a *KRAS p.G12C* mutation was equally not feasible given that of the 206 patients identified with a *KRAS* mutation, only 85 patients had a *KRAS p.G12C* mutation. Nevertheless, the prognosis of *KRAS p.G12C* and *KRAS* mutant populations are considered to be comparable and although not directly used in this alternative analysis, summary-statistics regarding the *KRAS p.G12C* populations, in particular HRs for OS and PFS, are presented to support this assumption.

Full details of the propensity score analysis conducted is presented in Section B.2.9.3.1.

The OS hazard ratio (95% CI) based on the Cox model for the *KRAS* mutant population was estimated at [redacted] (Table 27), with an effective sample size of 104.8 in the control arm. The hazard ratio calculated for *KRAS p.G12C* was [redacted] and was estimated from an effective sample size of 17.8.

The PFS hazard ratio (95% CI) based on the Cox model for the *KRAS* mutant population was estimated at [redacted] (Table 27). The hazard ratio calculated for *KRAS p.G12C* was [redacted]

Table 27: Flatiron based HRs of OS and PFS (PSW-adjusted, ATT weights, sotorasib vs. chemotherapy)

	***KRAS* mutant**		***KRASp.G12C* mutant**	
Outcome	**ESS**	**HR (95% CI)**	**ESS**	**HR (95% CI)**
Overall survival	104.8	[redacted]	17.8	[redacted]
Progression-free survival	104.8	[redacted]	17.8	[redacted]
Key: ATT, average treatment effect of the treated; CI, confidence interval; ESS, effective sample size (refers to comparator arm); HR, hazard ratio.				

Overall survival

The weighted KM plot for OS from the propensity score weighting analysis is presented in Figure 20.

Figure 20. Kaplan-Meier plot of OS from the propensity score weighting analysis

As per the approach taken previously, standard parametric distributions were fitted to the ATT propensity-score weighted time-to-event data using the statistical software R (ver. 4.0.3) using *'flexurv'* and *'survminer'* packages. The parametric distributions included exponential, Weibull, Gompertz, generalized gamma, lognormal and loglogistic.

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 28). For individually fitted curves, the lognormal distribution was the best statistically fitting curve with the lowest AIC and BIC across both sotorasib and chemotherapy, with the exception of the BIC for chemotherapy which marginally favoured the exponential distribution. Consistent with the approach taken previously, jointly fitted survival models (either restricted or unrestricted) were considered the most appropriate for this analysis.

For the jointly fitted curves, AIC and BIC indicate that the best fitting curve for both the restricted and unrestricted models was consistently the lognormal distribution with the restricted model demonstrating the best statistical fit based on BIC criteria. This conclusion is consistent with the goodness of fit statistics conducted on the SELECT-1 MAIC analysis.

Table 28: Goodness of fit statistics for jointly fitted OS models for KRAS mutant ATT

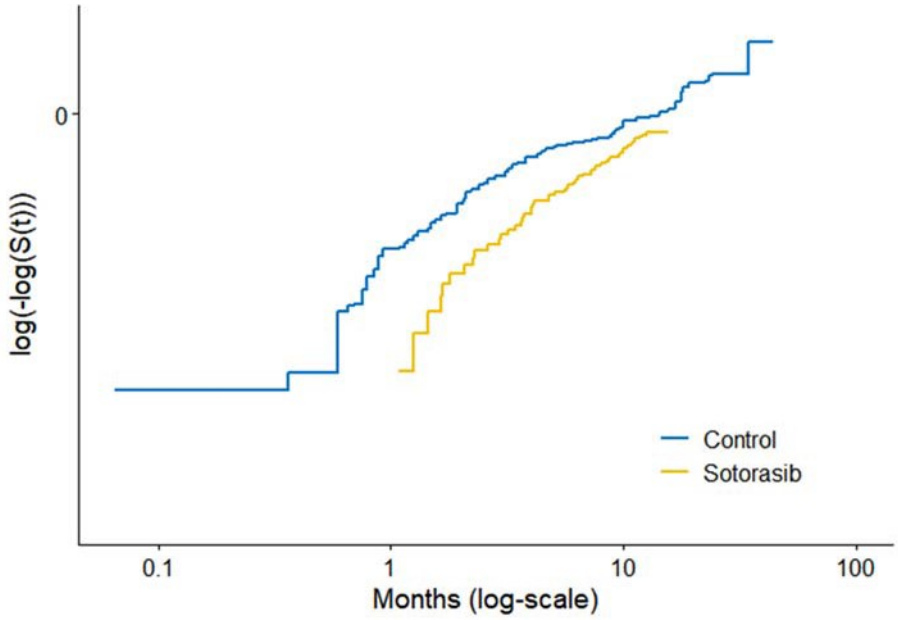
Model	Independent fit – sotorasib		Independent fit - chemo		Joint fit (unrestricted)		Joint fit (restricted)	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	438.2	441.0	605.1	608.4	1043.3	1050.8	1043.3	1050.8
Gompertz	437.6	443.1	605.7	612.4	1043.3	1058.4	1045.2	1056.5
Weibull	434.0	439.6	606.6	613.2	1040.6	1055.7	1044.8	1056.1
Generalised Gamma	432.1	440.4	604.1	614.1	1036.2	1058.9	1035.8	1050.9
Loglogistic	432.4	438.0	604.6	611.3	1037.0	1052.2	1038.2	1049.5
Lognormal	430.4	436.0	602.4	609.0	1032.8	1047.9	1033.8	1045.2

Key: AIC, Akaike information criterion; ATT, average treatment effect of the treated; BIC, Bayesian information criterion.
Note: Underlined values indicate the best statistically fitting parametric distribution.

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

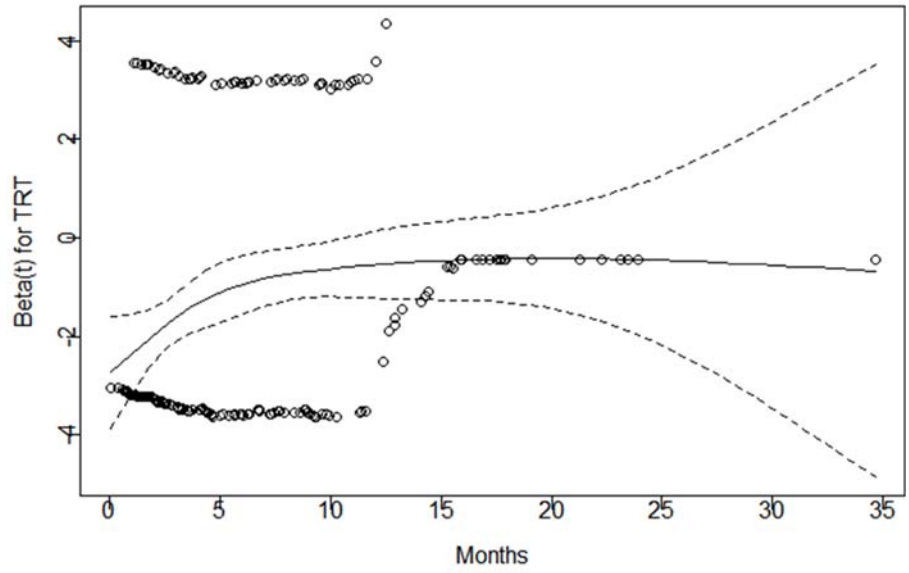
The proportional hazards assumption was evaluated for OS using the log cumulative hazards plot (Figure 21) and the Schoenfeld residuals plot (Figure 22). These plots indicated the proportional hazards assumption was unlikely to be valid. Accelerated time failure for OS was assessed using a Q-Q plot (Figure 23). The plot indicated that despite some deviation either side of the from the fitted line the assumption of accelerated failure time appears reasonable.

Figure 21: OS log cumulative hazards plot for sotorasib and control



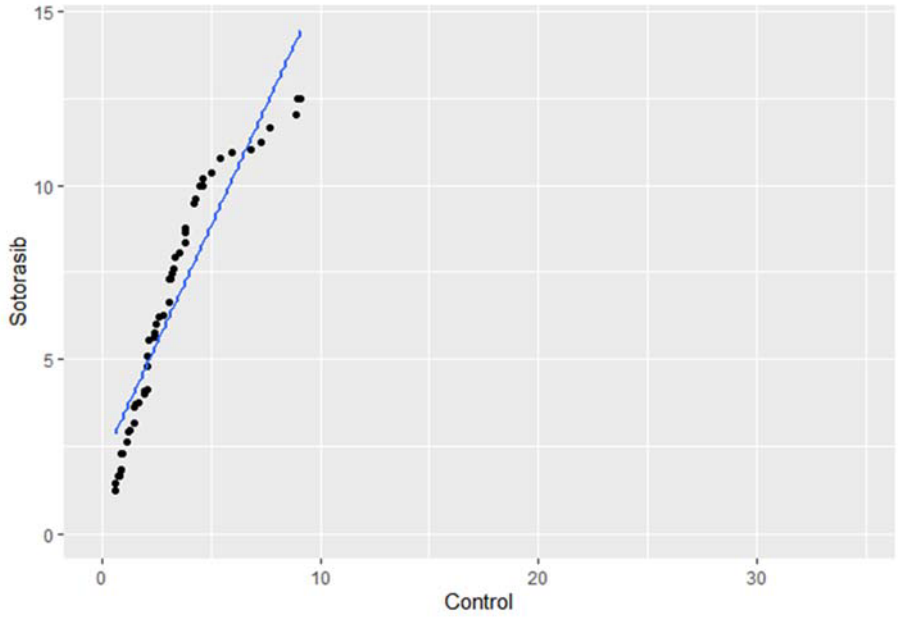
Key: OS, overall survival.

Figure 22: OS Schoenfeld residuals plot for sotorasib and control



Key: OS, overall survival.

Figure 23: OS Q-Q plot for sotorasib and control



Key: OS, overall survival; Q-Q, quartile-quartile.

Distribution Selection

Given that the joint fit (restricted) lognormal provides the best statistical fit to the observed ATT propensity adjusted data and the assumption of accelerated failure time appears to hold, this curve was selected to inform this sensitivity analysis. The visual fit of the ATT propensity KM data to the lognormal distribution is presented in Figure 24.

The use of this analysis as a supplementary, external validation for the primary MAIC-based analysis is discussed in Section B.2.9.3.1.

Figure 24: ATT OS KM versus fitted lognormal model using restricted model

Key: ATT, average treatment effect of the treated ; KM, Kaplan–Meier, OS, overall survival.

Progression-free survival

The weighted KM plot for PFS from the propensity score analysis is presented in Figure 25.

Figure 25. Kaplan-Meier plot of PFS from the propensity score weighting analysis

As per the approach taken previously, standard parametric distributions were fitted to the ATT propensity-score weighted time-to-event data using the statistical software R (ver. 4.0.3) using *'flexurv'* and *'survminer'* packages. The parametric distributions included exponential, Weibull, Gompertz, generalized gamma, lognormal and loglogistic.

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 29). For individually fitted curves, the lognormal distribution was the consistently best statistically fitting curve with the lowest AIC and BIC across both sotorasib and chemotherapy. As a result, jointly fitted survival models (either restricted or unrestricted) were considered the most appropriate since they can reduce uncertainty due to the estimation of fewer parameters and the use of a larger data set.

For the jointly fitted curves, AIC and BIC indicate that the best fitting curve for both the restricted and unrestricted models was consistently the lognormal distribution with the restricted model demonstrating the best statistical fit based on BIC criteria. This conclusion is consistent with the goodness of fit statistics conducted on the SELECT-1 MAIC analysis.

Table 29: Goodness of fit statistics for jointly fitted PFS models for KRAS-mutant ATT

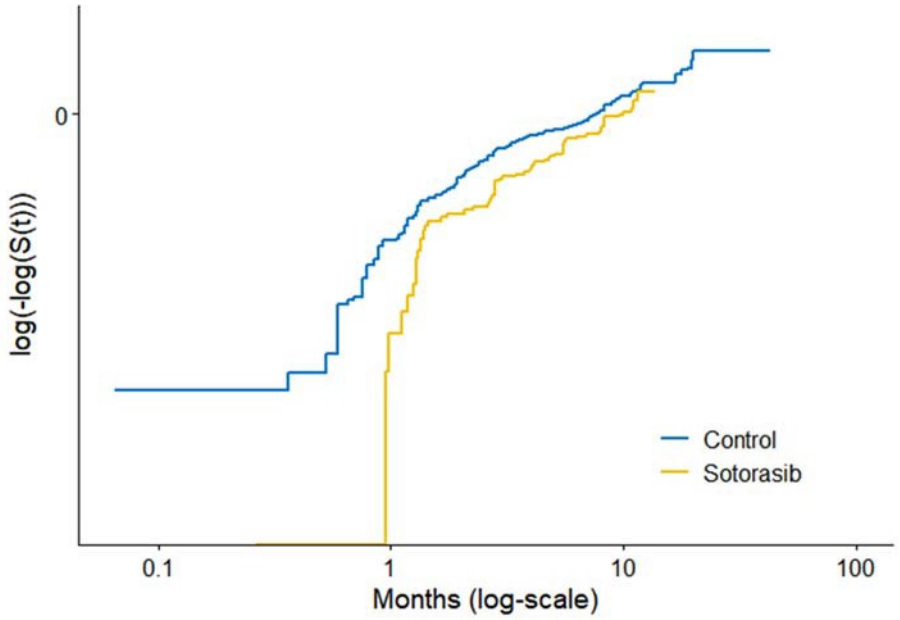
Distribution	**Independent fit – sotorasib**		**Independent fit - chemotherapy**		**Joint fit (unrestricted)**		**Joint fit (restricted)**	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	506.8	509.6	601.5	604.8	1108.3	1115.9	1108.3	1115.9
Gompertz	505.4	511.0	603.0	609.7	1108.4	1123.6	1110.2	1121.6
Weibull	501.6	507.2	603.4	610.1	1105.0	1120.1	1107.0	1118.4
Generalized Gamma	497.7	506.0	597.7	607.6	1095.4	1118.1	1093.4	1108.6
Loglogistic	499.8	505.4	598.8	605.4	1098.6	1113.7	1097.6	1108.9
Lognormal	496.2	501.7	595.9	602.5	1092.0	1107.1	1091.4	1102.8

Key: ATT, average treatment effect of the treated; AIC, Akaike information criterion; BIC, Bayesian information criterion; PFS, progression-free survival.
Note: Underlined values indicate the best statistically-fitting parametric distribution.

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

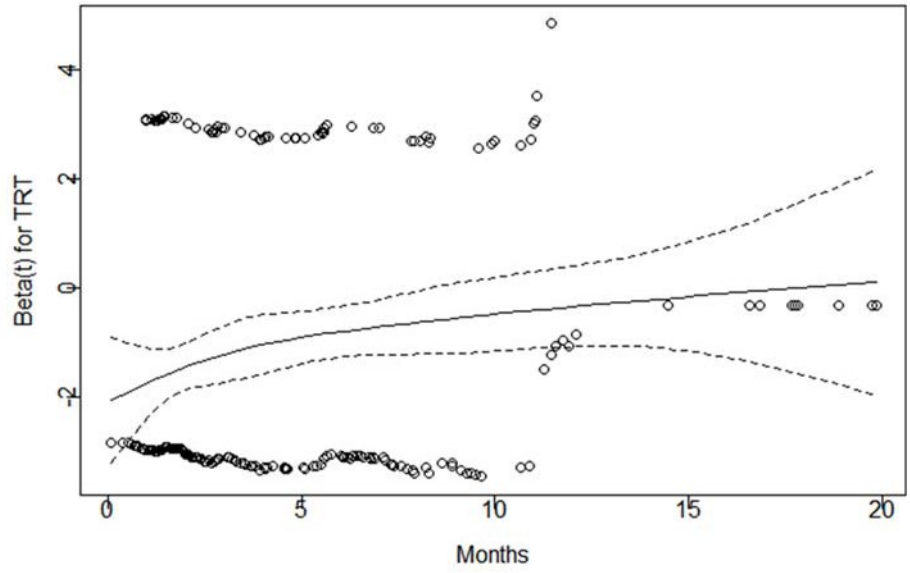
The proportional hazards assumption was evaluated for PFS using the log cumulative hazards plot (Figure 26) and the Schoenfeld residuals plot (Figure 27). These plots indicated the proportional hazards assumption was not valid. Accelerated time failure for PFS was assessed using a Q-Q plot (Figure 28). The plot indicated some deviation either side of the from the fitted line. However, overall the assumption of an accelerated failure time model appeared acceptable.

Figure 26: PFS log cumulative hazards plot for sotorasib and control



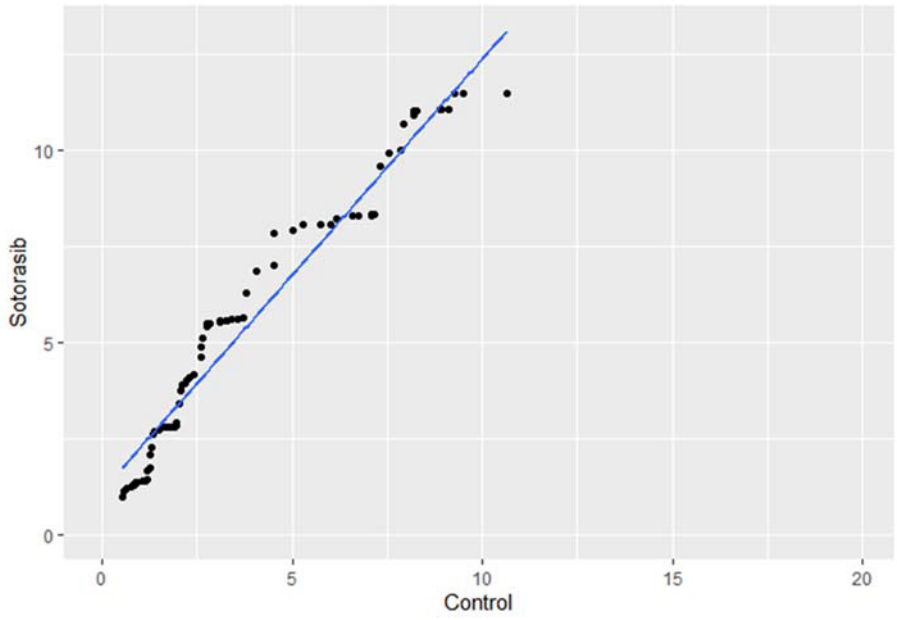
Key: PFS, progression-free survival.

Figure 27: PFS Schoenfeld residuals plot for sotorasib and control



Key: PFS, progression-free survival.

Figure 28: PFS Q-Q plot for sotorasib and control



Key: PFS, progression-free survival; Q-Q, quartile-quartile.

Distribution Selection

Given that the joint fit (restricted) lognormal provides the best statistical fit to the observed ATT propensity adjusted data and the assumption of accelerated failure time appears to hold, this curve was to inform this sensitivity analysis. The visual fit of the ATT propensity KM data to the lognormal distribution is presented in Figure 29 and shows a close visual fit of the extrapolation to the Kaplan–Meier data.

Figure 29: ATT PFS KM versus fitted lognormal model using restricted model

[redacted]

Key: ATT, average treatment effect of the treated; KM, Kaplan–Meier, PFS, progression-free survival.

B.3.3.5 Nintedanib + docetaxel

No direct comparative evidence was identified to inform the efficacy for the secondary comparison between the nintedanib + docetaxel combination and sotorasib. As such, efficacy for nintedanib plus docetaxel was obtained from the LUME Lung-1 trial, which compared the nintedanib plus docetaxel combination treatment with docetaxel monotherapy in patients with advanced NSCLC [32]. This trial was selected as it was the pivotal Phase III study investigating nintedanib plus docetaxel versus docetaxel, was recommended by UK clinicians as the most appropriate and relevant data to the decision problem, and formed the core evidence base in NICE TA347 [51]. Specifically, patients with adenocarcinoma histology (N=322) were selected for this analysis as this reflects current use of nintedanib plus docetaxel in UK clinical practice and more closely aligns with patient characteristics in CodeBreaK100. Further details of the LUME Lung-1 and CodeBreaK100 populations are presented and discussed in Section B.2.9.2.2.

A UK advisory board conducted by Amgen confirmed that an MAIC was unlikely to be appropriate for this comparison [14]. The main obstacles to conducting an MAIC included the fact the KRAS status of patients in LUME Lung-1 was unknown, there were fewer smokers and patients with brain metastases in LUME Lung-1 compared to CodeBreaK100, and the distribution and type of prior lines or treatments were different. Furthermore, it was considered more appropriate to assess the comparative effectiveness in a population consistent with SELECT-1 rather than introduce another population into the analysis.

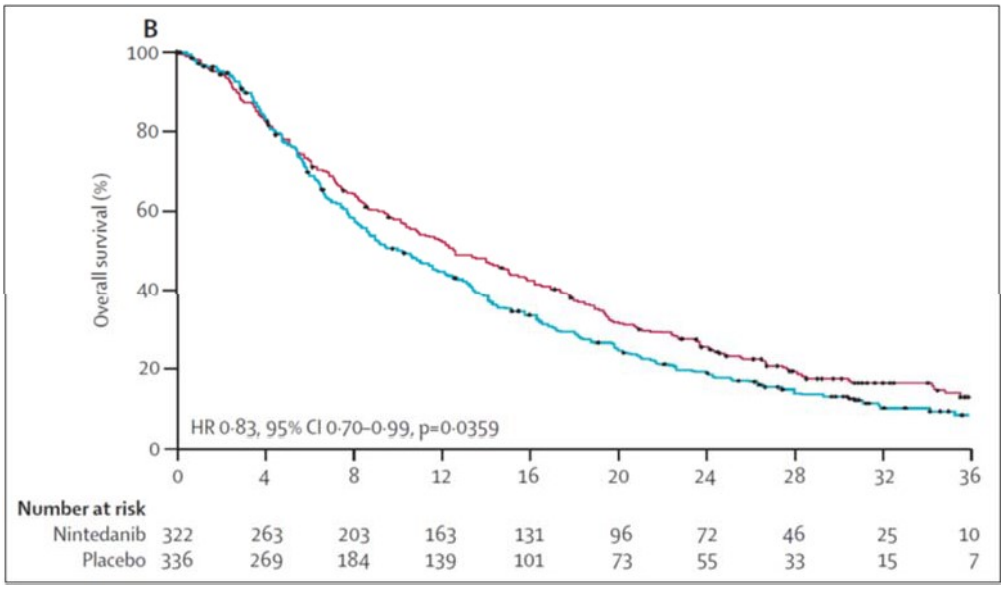
The PFS and OS Kaplan Meier plots for the adenocarcinoma population were digitised, and pseudo-patient level data was generated using the Guyot method [79].

Consistent with the preferred approach in NICE TA347, the modelling of nintedanib and docetaxel was carried out by applying time-dependent HRs to the docetaxel monotherapy arm of SELECT-1. The time-dependent HR's were derived from the pseudo-patent level data generated from digitized OS and PFS data for docetaxel from SELECT-1 and nintedanib and docetaxel from LUME Lung-1 study. Further details are provided in the OS and PFS sections below.

<u>Overall survival</u>

Visual inspection of the OS KM plots for nintedanib and docetaxel and placebo and docetaxel suggested the proportional hazards assumption may not hold, with little difference between outcomes at the start of follow-up (up to 4 months), and the plots appearing to converge towards the end of follow-up (after 26 months) (Figure 30).

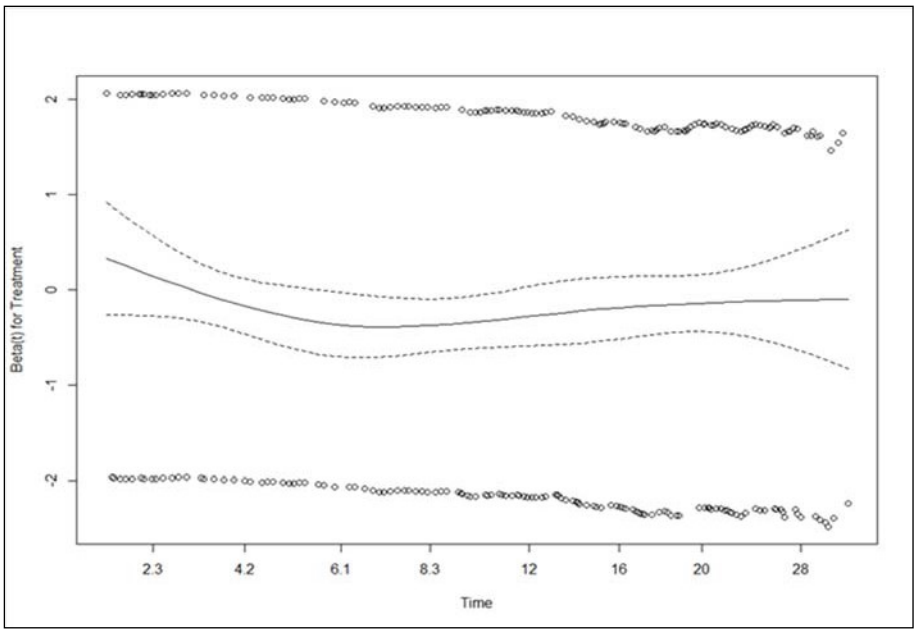
Figure 30: OS Kaplan–Meier plot from LUME Lung-1 for nintedanib versus placebo



Key: CI, confidence interval; HR, hazard ratio; OS, overall survival.

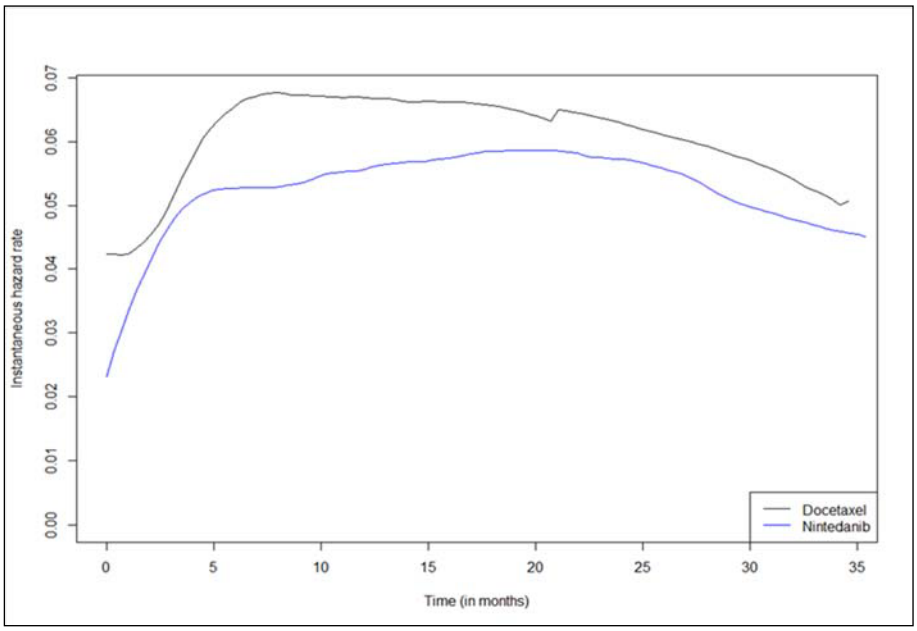
The data were assessed for proportional hazards using Schoenfeld residuals plot (Figure 31) and the instantaneous hazards plot (Figure 32). The plots indicated that the proportional hazards assumption did not hold. Based on the shape of the instantaneous hazards plot, an inflection point was observed at 6 months in which the direction of the hazards changed direction (Figure 32). Similarly, the direction of the hazards was judged to change (albeit gradually) at the landmark 26-month time point (Figure 32). These intervals were confirmed using the Kaplan–Meier plot (Figure 30) and were used for a piecewise hazard ratio calculation.

Figure 31: OS Schoenfeld residuals from LUME Lung-1 for nintedanib and docetaxel



Key: OS, overall survival.

Figure 32: OS instantaneous hazards plot from LUME Lung-1 for nintedanib and docetaxel



Key: OS, overall survival.

A separate Cox model was applied to each interval. The observation time started at the beginning of each interval and was censored at the end of each interval. The OS piecewise hazard ratios are presented (Table 30).

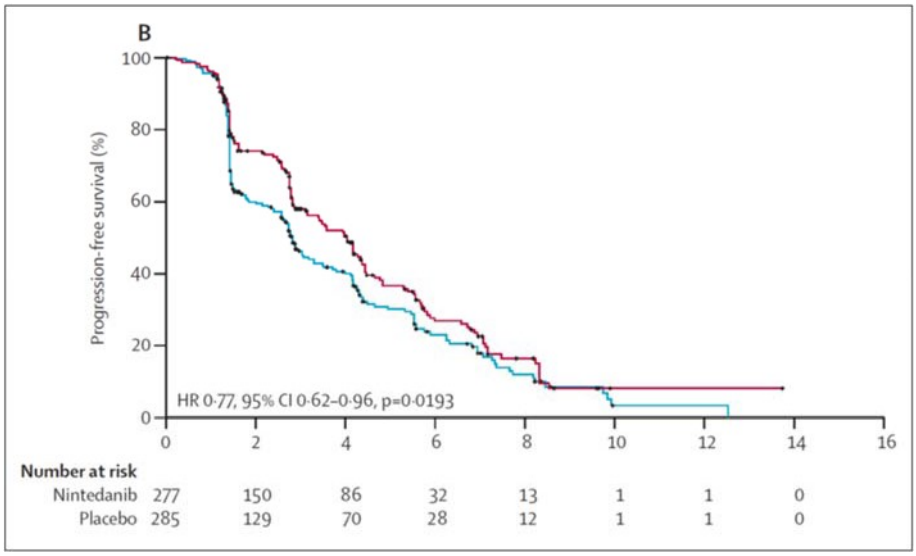
Table 30: OS piecewise hazard ratios for nintedanib + docetaxel versus sotorasib

Analyses	HR (0 – 6 months) (95% CI)	HR (6 – 26 months) (95% CI)	HR (26+ months) (95 % CI)
Hazard ratio for overall survival	[redacted]	[redacted]	[redacted]
Key: CI, confidence interval; HR, hazard ratio; OS, overall survival.			

Progression-free survival

Visual inspection of the PFS Kaplan–Meier plots for nintedanib and placebo suggested the proportional hazards assumption may not hold, with little difference between outcomes at the start of follow-up (up to 2 months) and the plots appearing to converge towards the end of follow-up (after 6 months; Figure 33).

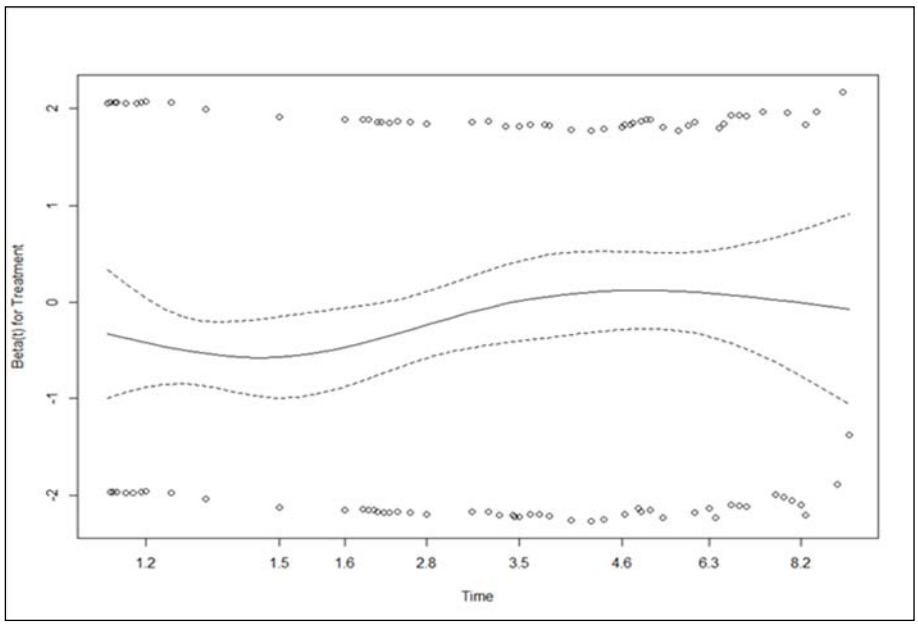
Figure 33: PFS Kaplan–Meier plot from LUME Lung-1 for nintedanib versus placebo



Key: CI, confidence interval; HR, hazard ratio, PFS, progression-free survival.

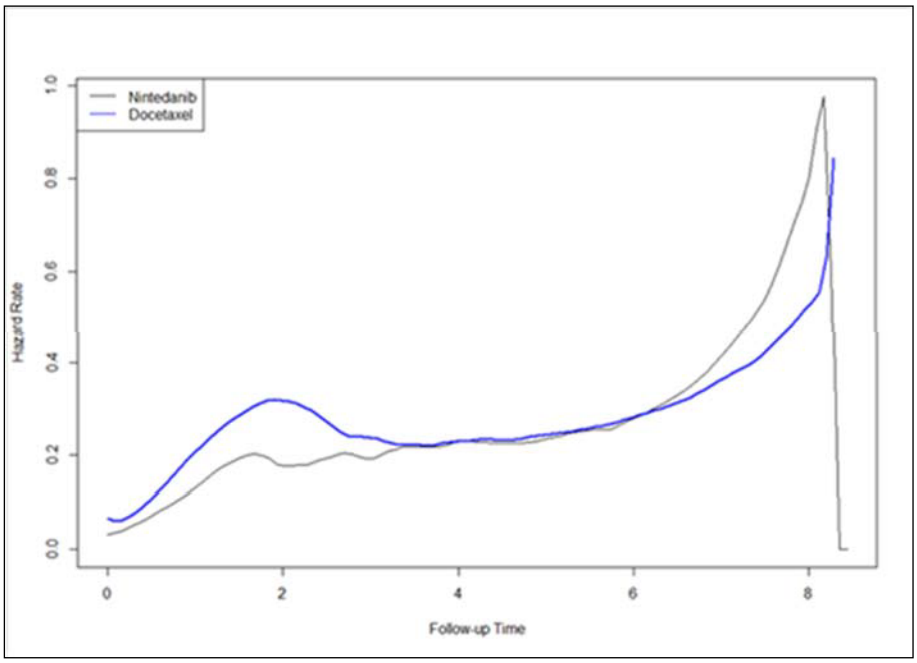
The data was assessed for proportional hazards using Schoenfeld residuals plot (Figure 34) and the instantaneous hazards plot (Figure 35). The plots indicated that the proportional hazards assumption did not hold. Based on the shape of the instantaneous hazards plot, an inflection point was observed at 2 months in which the direction of the hazards changed direction (Figure 35). Similarly, the direction of the hazards was judged to change (albeit gradually) at the landmark 6-month time point (Figure 35). These intervals were confirmed using the Kaplan–Meier plot (Figure 33) and were used for a piecewise hazard ratio calculation.

Figure 34: PFS Schoenfeld residuals from LUME Lung-1 for nintedanib and docetaxel



Key: PFS, progression-free survival.

Figure 35: PFS instantaneous hazards plot from LUME Lung-1 for nintedanib and docetaxel



Key: PFS, progression-free survival

A separate Cox model was applied to each interval. The observation time started at the beginning of each interval and was censored at the end of each interval. The PFS piecewise hazard ratios are presented in Table 31.

Table 31: PFS piecewise hazard ratios for nintedanib + docetaxel versus sotorasib

Analyses	HR (0 – 2 months) (95% CI)	HR (2 – 6 months) (95% CI)	HR (6+ months) (95 % CI)
Hazard ratio for progression-free survival	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; HR, hazard ratio; PFS, progression-free survival.

The hazard ratios for each interval were validated by comparing to the hazard ratios calculated for the full period and cross referencing to the full period hazards in the publication, comparing the nintedanib arm with the reference docetaxel arm in the LUME Lung-1 trial [32]. The OS hazard ratio for the full period was calculated as 0.833 (95% CI: 0.701, 0.992), which was similar to the published hazard ratio of 0.83 (Figure 30). The PFS hazard ratio for the full period was 0.772 (95% CI: 0.621, 0.958), which compared to the reported hazard of 0.770 (Figure 33).

B.3.3.6 Treatment duration

Sotorasib

Sotorasib treatment duration was modelled using a hazard ratio applied to PFS from CodeBreaK100. The hazard ratio was estimated using a Cox model with the effect estimated between time to treatment discontinuation and progression-free survival ([redacted], 95% CI: [redacted]). The modelled treatment duration curve was compared to the unadjusted time-to-treatment-discontinuation curve and indicated a close match, with a marginal over estimation up to 6 months and under estimation after 6 months of modelled treatment duration to Kaplan–Meier curve (Figure 36).

Figure 36: Time to treatment discontinuation for sotorasib

Key: KM, Kaplan–Meier; TTD, time-to-treatment discontinuation

An alternative approach to modelling sotorasib treatment duration was explored in a sensitivity analysis where the weights generated from the MAIC analysis were applied to the CodeBreaK100 discontinuation data and parametric models were fitted to extrapolate the treatment duration. This approach was considered to be more complex and ultimately dependent on the variable selection in the MAIC analysis. Furthermore, as this approach does not meaningfully reduce any uncertainty associated with extrapolating treatment duration it was not preferred in the base case analysis.

Docetaxel

There was no robust data to inform treatment duration for docetaxel. However, as the cost associated with docetaxel is relatively small, the effect of docetaxel treatment duration on the results was expected to be negligible. Docetaxel treatment duration was therefore assumed to be equal to progression-free survival in SELECT-1. A plot showing treatment duration for docetaxel and sotorasib used in the economic model is presented below (Figure 37).

Figure 37: Time to treatment discontinuation for docetaxel and sotorasib

Key: PFS, progression-free survival; TTD, time-to-treatment discontinuation.

Nintedanib

The approach of modelling nintedanib treatment duration was based on a review of time to treatment discontinuation and progression-free survival reported in the previous NSCLC nintedanib NICE submission for TA347 [51].

Nintedanib per-cycle discontinuation was reported as 0.125 (NICE TA347 Committee Papers, Table 112) [95]. This corresponds to a monthly discontinuation rate of 0.193. In contrast, the median PFS was reported as 3.4 months (NICE TA347, Table 18[51]), this corresponds to a monthly rate of 0.204. Therefore, the nintedanib discontinuation rate in TA347 was lower than the progression rate.

In the cost-effectiveness model we assume that subjects discontinue nintedanib at disease progression. As based on TA347 the discontinuation rate was lower than the progression rate, this modelling assumption can be considered as conservative.

B.3.3.7 Adverse events

Grade 3+ adverse events with an incidence of ≥ 5% in any of the comparator arms (sotorasib, docetaxel, nintedanib and docetaxel) are included in the model. Sotorasib adverse events are informed by the CodeBreaK100 Clinical Study Report for the 01 December 2020 data cut [55]. Docetaxel and nintedanib plus docetaxel adverse events are informed by the SELECT-1 and LUME-Lung 1 clinical trials [31, 32]. The 95% confidence intervals for adverse event rates are calculated using the method by Clopper and Pearson (1934) [96].

A table of the adverse events and incidence used in the model is shown in Table 32. In the base case analysis treatment-related adverse events are utilised for the sotorasib and docetaxel treatment arms; however, only treatment emergent AEs were available from LUME Lung-1 to inform the secondary comparison with nintedanib plus docetaxel. A scenario analysis assessing the impact of using treatment-emergent AEs for both sotorasib and docetaxel is also presented. Nevertheless, treatment-related AEs were preferred to minimise bias given the absence of randomised data and the fact that some AEs may be driven by the underlying disease.

Table 32: Adverse events incidence

Adverse event	Sotorasib [a]	Docetaxel [b]	Nintedanib + docetaxel [c]
Decreased neutrophils	0.8%	0.0%	32.1%
Decreased white blood cell count	0.0%	0.0%	16.4%
Diarrhoea	4.0%	2.4%	6.4%
Dyspnea [d]	0.0%	0.0%	0.0%
Fatigue	0.0%	1.6%	5.5%
Febrile neutropenia	0.0%	0.0%	7.1%
Increased ALT	6.3%	0.0%	7.8%
Increased AST	5.6%	0.0%	3.4%
Neutropenia	0.8%	1.6%	12.1%
Pleural Effusion [d]	0.0%	0.0%	0.0%
Pneumonia [d]	0.0%	0.0%	0.0%

Key: ALT, alanine aminotransferase, AST, aspartate aminotransferase.
Note:
[a], CodeBreaK100 phase 2 NSCLC cohort, safety analysis set Table 14b-6.8.2 Treatment-related Adverse Events. December 1, 2020 data cut-off [55]
[b]Janne 2017 eTable 1 Most Frequently Reported Adverse Events Causally Related to treatment [31]
[c], Reck 2014 Table 3 Treatment-Emergent Adverse Events [32]
[d] Treatment emergent adverse events included alongside base case AEs as scenario analysis

B.3.3.8 Mortality

Survival in the model is capped by age-sex matched general population mortality based on published UK life tables for 2017–19 [97].

B.3.4. Measurement and valuation of health effects

As described in Section B.3.1 an SLR was conducted to identify published studies for evaluating cost-effectiveness, costs and resource use and health-related quality of life for treatments in NSCLC relevant to the decision problem. Full details on the methodology and findings of the SLR, including search terms, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram and outcomes are detailed in Appendix H.

The SLR did not identify any studies which provided HRQoL estimates for NSCLC with *KRAS p.G12C* mutation or more broadly with any *KRAS* mutation. As such, the primary source of HRQoL values used for the NSCLC *KRAS p.G12C* mutation in the model was

CodeBreaK100, using 01 September 2020 data cut-off. The analytical approaches for utility estimation included both estimation by time to death and by health state and the methods used in the following sections. In the base case analysis, the time to death utility approach was used to better reflect the deterioration of patient quality of life at the end of life based on clinical expert feedback. Utilities based on health state (progression-free or progressed) was explored in a scenario analysis.

B.3.4.1 Health-related quality-of-life data from CodeBreaK100

HRQoL was collected in CodeBreaK100 using the EQ-5D-5L instrument, a generic patient reported outcome measure commonly used to assess patient quality of life over time [98]. Questionnaires were completed on day 1 of cycle 1 and on every first day of subsequent cycles up to cycle 7, and then on the first day of every second cycle until end of treatment.

The analysis conducted included a descriptive analysis by visit, progression status and time before death. A mixed-models with repeated measures were fitted to estimate the impact of disease progression, time before death category (> 6 months, 3–6 months, 1–3 months and < 1 month before death) which were used to inform the model.

The descriptive analysis included two approaches:

- The analyses involving nominal utility index scores were performed on the dataset which included all patients from the CodeBreaK100 NSCLC safety analyses set (N = 126) who completed at least one EQ-5D-5L questionnaire in line with study protocol with all fields of the questionnaires completed (AN01 analysis set).

- Utility change from baseline or including baseline utility value as a covariate was performed on the analysis set of patients who have completed the EQ-5D-5L questionnaire at baseline and who had at least another visit completed with no missing data (AN02 analysis set). For analyses involving categorization by progression, three patients were excluded from the full analysis set (N = 123).

Mixed models with repeated measures (MMRM) were fitted to the data to capture the impact of (i) progression category and (ii) time before death category on quality of life, while accounting for inter- and intra-subject variability in EQ-5D questionnaires.

For each health state of interest, each of the following three model specifications tested were:

- Utility index score based on health state and random effects on subject id
- Utility index score based on health state, Eastern Cooperative Oncology Group Performance Status (ECOG-PS), brain metastases, number of prior lines of treatments, an interaction covariate for health state and random effects on subject id. This accounted for the potential effect and interaction of baseline covariates on the health state utility
- Utility index score based on health state, utility index baseline score, health state utility * utility index baseline score and random effects on subject id. This approach

assumed utility baseline score was a good proxy for the impact of other patient baseline characteristics and was fitted to the AN02 analysis set

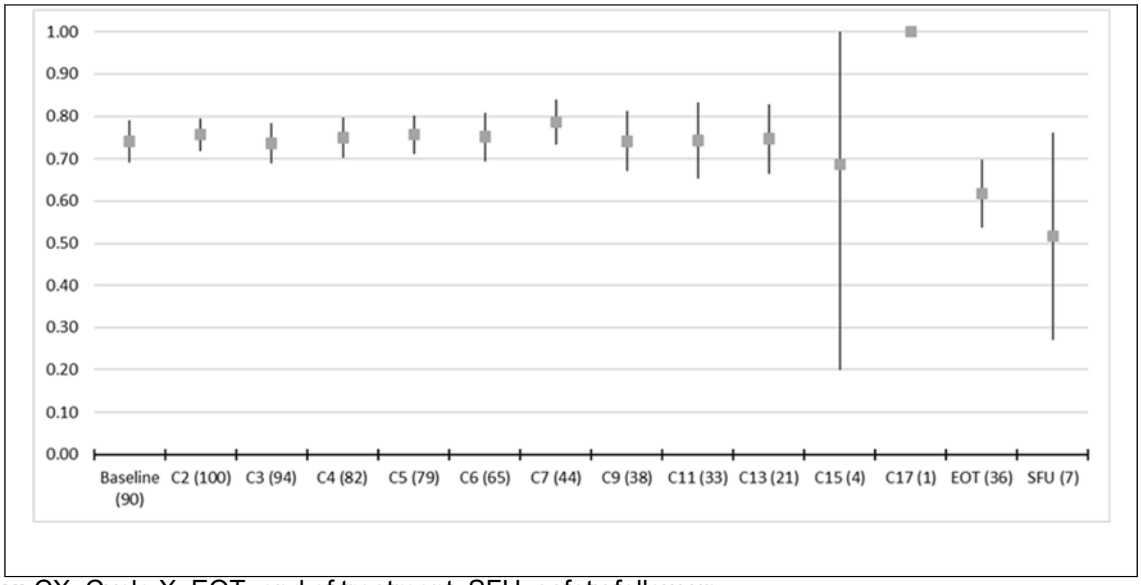
Sensitivity analysis was undertaken for progression status, where models were fitted to the full analysis set * AN01 analysis set to account for the exclusion of three patients from the full analysis set.

Descriptive analysis

The number of patients with at least one completed EQ-5D-5L per visit per protocol and with all components of the EQ-5D-5L questionnaire completed for the Safety Analysis Set was 122 out of 126, and for the Full Analysis Set was 119 out of 123. The number of patients with a completed EQ-5D-5L at baseline visit per protocol and at least one completed EQ-5D visit was 86 in the Safety Analysis Set and 84 in the Full Analysis Set.

The results indicated that while on treatment, quality of life was maintained to the level of baseline over time (Figure 38).

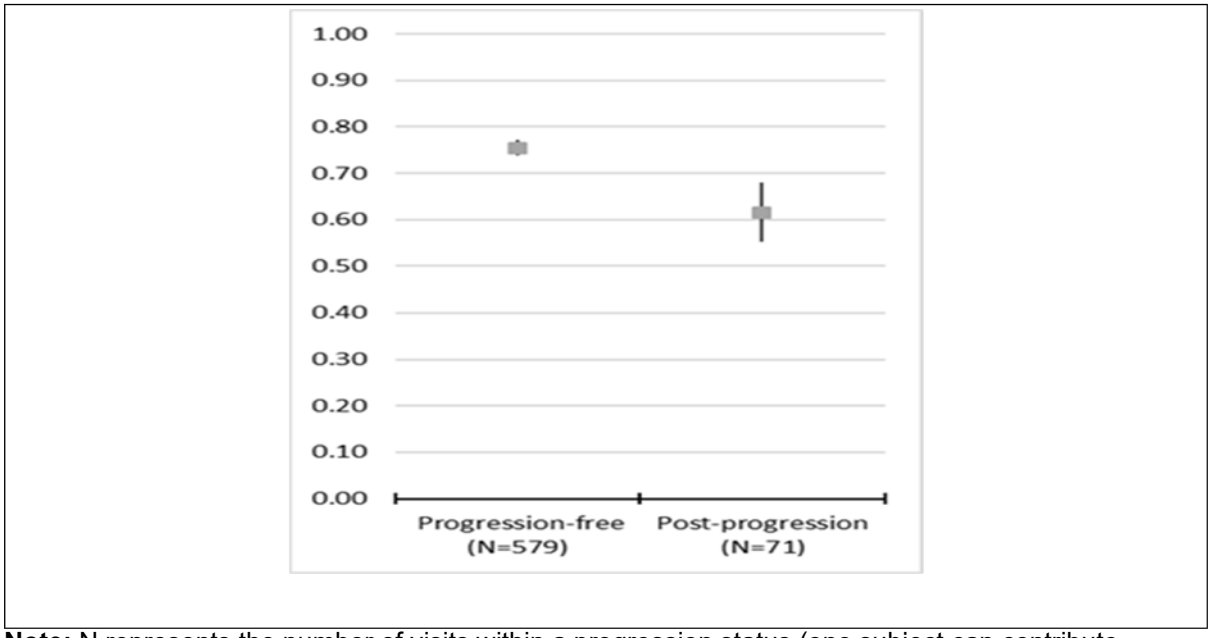
Figure 38: Mean EQ-5D-5L utility index score by visit (Safety Analysis Set)



Key: CX, Cycle X; EOT: end of treatment; SFU: safety follow up.
Note: The number of subjects by visit are given between brackets after the visit label. Graph shows mean EQ-5D-5L utility index score along with 95% confidence intervals calculated with normal distribution approximation

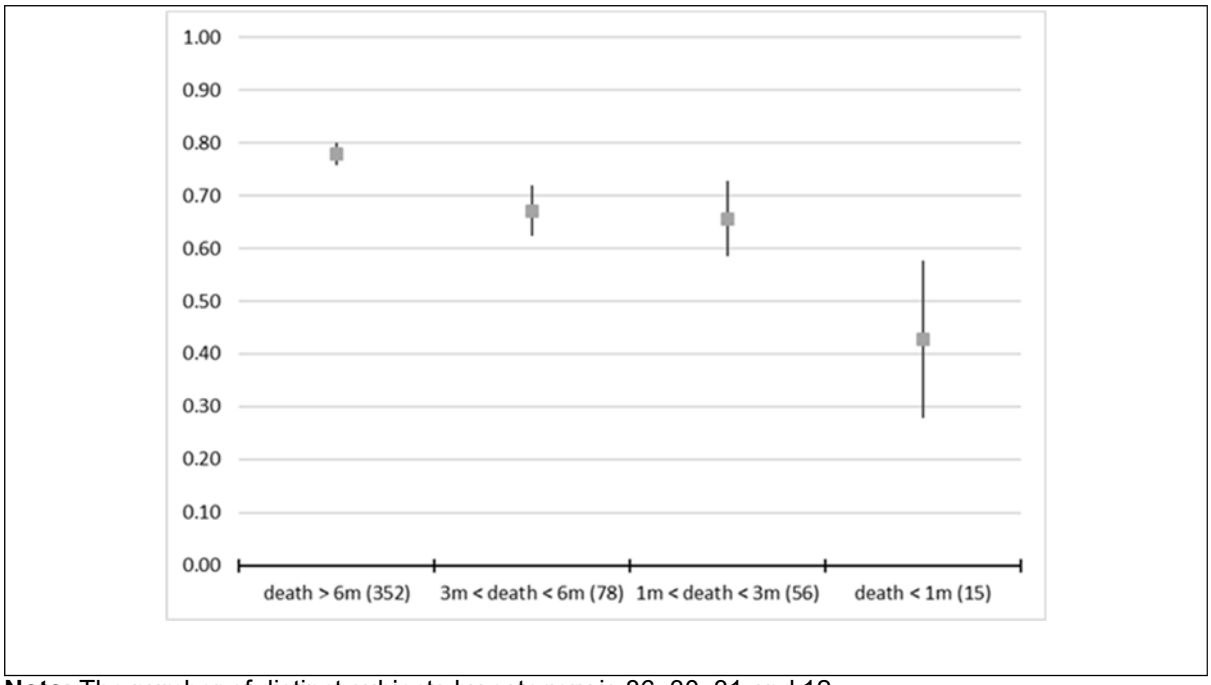
The mean EQ-5D-5L utility index score was calculated by progression status for the Safety Analysis Set and the Full Analysis Set (AN01) using UK tariffs (Figure 39). The results indicate the negative impact of progression on quality of life. A clear trend for decreasing quality of life as patients approach death was noted, with a similar value reported at 3 months prior and 1 month prior to death (Figure 40). There were fewer observations for the last category (< 1 month before death), which could be accounted for by a bias in response where patients were too sick to answer the questionnaire.

Figure 39: Mean EQ-5D-5L utility index score by progression status (Safety Analysis Set for AN01).



Note: N represents the number of visits within a progression status (one subject can contribute several visits). Graph shows mean EQ-5D-5L utility index score along with 95% confidence intervals calculated with normal distribution approximation. 120 unique subjects contributed to the progression-free estimates and 41 unique subjects contributed to the post-progression estimates.

Figure 40: Mean EQ-5D-5L utility index score by time to death (Safety Analysis Set for AN01) using all subjects



Note: The number of distinct subjects by category is 86, 30, 31 and 12.

Mixed models with repeated measures analysis

The MMRM approach assumed that cycle visit had no impact on the utility index score. This assumption was justified on the basis that visit cycle did not demonstrate a clear trend in the results over time when fitted using MMRM. The consequence of this assumption is that visits for individual subjects were not ordered by time in the models.

The MMRMs demonstrated a clear and consistent progression effect on subject quality of life in both datasets (AN01 and AN02). When using the subset with lower baseline values (AN02), the utility decrement was lower (-0.082 versus -0.071 respectively), but the progression-free utility was higher using AN02 versus AN01 (0.740 versus 0.745 respectively). Removing the three patients who were not in the full analysis set (FAS) had no impact on the estimates or uncertainty ranges, so the FAS was used in the EQ-5D analyses presented below.

All models using AN02 used the baseline utility value as an independent covariate. The estimated impact of progression did not vary substantially across models (Table 33).

Table 33: Mixed models and repeated measures progression status adding baseline utility as covariate.

Model	Progression disutility estimate (95% CI)	Dataset
Utility index score ~ progression status	[redacted]	AN01 FAS
Utility index score ~ progression status	[redacted]	AN02 FAS
Utility index score ~ progression status and baseline value	[redacted]	AN02 FAS
Key: CI, confidence interval; FAS, full analysis set.		

Other models were tested, with additional covariates in addition to progression status and EQ-5D utility score at baseline. This included models with interactions between progression status and important covariates (utility index at baseline, ECOG-PS at baseline, brain metastasis at baseline etc.). No interaction between progression status and other covariates were found to have a significant impact. These models were not preferred as they resulted in higher AIC than the model with only progression status and baseline utility as covariates.

To confirm that the model approach using progression status and baseline utility was the best fitting approach, a stepwise selection based on AIC criteria was conducted using progression status and specifying the following covariates: baseline utility index score, age group (≤ 64, 65–74 and ≥ 75), gender, race, region, ECOG-PS, brain metastases at baseline, liver metastases, prior treatment with platinum chemotherapy, prior treatment with anti-PD-(L)1 and number of prior lines of therapy. The stepwise analysis confirmed the selection of the model with baseline utility as the covariate.

B.3.4.2 Mapping

The NICE reference case stipulates the EQ-5D-3L instrument [83]. It was therefore necessary to map the trial outcomes from the 5-level instrument to the 3-level instrument to

align with the NICE reference case. This was done using the NICE recommended cross-walk algorithm using the UK tariff, published by van Hout (2012) [91].

B.3.4.3 Health state utilities

The inclusion of baseline utility in the MMRM did not result in a substantial difference to the estimate of the impact of progression on utility index score. To account for all information available the estimates of the MMRM with only progression status as covariate (plus the random effect of subject) fitted to the AN01 using FAS (N = 119) was used. The health state utility for progression free and the estimate for disutility due to progression are given below, along with their 95% confidence intervals (Table 34).

Table 34: Summary of health state utility values

Health state	Mean (95% CI)	Reference
Progression-free	0.739 (0.704, 0.774)	CodeBreaK100 EQ-5D-5L analyses [a] [57] using UK crosswalk tariffs [91]
Disutility in progressed disease	0.084 (0.044, 0.123)	
Post-progression	0.655	Calculation

Key: CI, confidence interval.
Note: [a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-4.7.702 and subsequent analyses

B.3.4.4 Time to death utilities

The same process was replicated for time-to-death categories (less than 1 month before death, 1–3 months before, 3–6 months before and more than 6 months before) as for progression status. The Safety Analysis Set was used to perform time to death analyses, as OS in CodeBreaK100 was estimated based on this analysis set. Observations that occurred within a 6-month window before a death censoring have been excluded from the analyses, as it was unknown in which time-to-death category the observation would finally occur.

There was a clear trend for a deteriorating utility as the subject approached death, across the tested models and fitted analysis sets. The change in quality of life in the last month of life was very clear. Marked differences in point estimated across all categories were observed, however not all differences were statistically significant due to the limited sample size.

The AN01 dataset was used to inform the model, similar to progression status. A stepwise selection based on AIC was performed which indicated that baseline utility was the only covariate. The estimates for time to death used in the economic model were calculated, as shown in Table 35.

Table 35: Time to death utilities

Health state	Mean (95% CI)	Reference
Utility more than 6 months to death	0.765 (0.728, 0.803)	CodeBreaK100 EQ-5D-5L analyses using UK crosswalk tariffs (1 September 2020 data cut-off)[91]
Disutility between 3 and 6 months to death (versus. more than 6 months)	0.040 (-0.011, 0.090)	
Disutility between 1 and 3 months to death (versus. more than 6 months)	0.120 (0.061, 0.179)	
Disutility less than 1 month to death (versus. more than 6 months)	0.250 (0.161, 0.339)	
Utility between 3 and 6 months to death	0.725	Calculated
Utility between 1 and 3 months to death	0.645	Calculated
Utility in last month of life	0.515	Calculated
Key: CI, confidence interval.		

B.3.4.5 Health-related quality-of-life studies

A search for studies reporting HRQoL in patients with the *KRAS p.G12C* mutation were conducted as part of the economic SLR (see Appendix H). One study was identified (Yang, 2016 [abstract]), which reported HRQoL in patients with advanced NSCLC which investigated the effect of including biomarker testing to guide individual therapy. However, the scale used in the HRQoL score was not reported [99]. A second study (Schuette, 2015) investigated the inclusion of panitumumab with chemotherapy in inoperable stage IIIB or IV primary pulmonary non-squamous NSCLC with KRAS wild type [100]. No studies were identified which reported HRQOL in patients with NSCLC and the *KRAS p.G12C* mutation.

B.3.4.6 Adverse reactions

The economic model includes the quality of life impact of AEs of Grade 3+ adverse events with an incidence of ≥ 5% in any of the comparator arms. Table 36 presents the disutility per episode for each of the included AEs consistent with sources used in previous NICE appraisals in this disease area. As disutility values could not be identified for all AEs, a disutility value of 0 was assumed in these cases. This assumption could potentially be conservative given the generally increased frequency of these AEs in the comparator arms versus sotorasib.

For each included AE the disutility was applied in the first model and the duration of each adverse event was assumed to be 4 weeks, with a lower bound of 3.2 weeks and upper bound of 4.8 weeks.

Table 36: Adverse event disutilities

Adverse event	Mean (95% CI)*	Source
Decreased neutrophils [a]	0.000	NICE TA484 [assumption][86]
Decreased white blood cell count [a]	0.050 (0.040, 0.060)	NICE TA 347, 520, and 484 [assumption] [51, 86, 88]
Diarrhoea [a]	0.047 (0.016, 0.077)	Nafees 2008[101]
Dyspnoea [b]	0.050 (0.026, 0.074)	Doyle 2008[102]
Fatigue [a]	0.073 (0.037, 0.110)	Nafees 2008[101]
Febrile neutropenia	0.090 (0.058, 0.122)	Nafees 2008[101]
Increased ALT [a]	0.050 (0.040, 0.060)	NICE TA 347, 520, and 484 [assumption] [51, 86, 88]
Increased AST [a]	0.000	NICE TA484 [assumption][86]
Neutropenia [a]	0.090 (0.059, 0.120)	Nafees 2008[101]
Pleural effusion [b]	0.000	NICE TA484 [assumption][86]
Pneumonia [b]	0.008 (0.006, 0.010)	Marti 2013[103]

Key: ALT, alanine aminotransferase, AST, aspartate aminotransferase.
Note:
*, confidence intervals calculated using normal distribution
[a], adverse events included in base case analysis
[b], treatment emergent adverse events included alongside base case AEs as scenario analysis

B.3.4.7 Health-related quality of life data used in the cost-effectiveness analysis

As CodeBreaK100 is a single-arm trial, there are no comparative data available on the quality of life associated with sotorasib versus the relevant comparators in this appraisal. Direct use of reported utility levels associated with comparators in other previously treated advanced NSCLC patients may introduce bias due to potential differences in patient characteristics.

In the base case analysis, utility values were estimated directly from CodeBreaK100 using a time to death approach for all comparators (Table 35). This approach was preferred to using utility values estimated based on disease health states as it reflects the findings of studies which have shown NSCLC patients to have markedly decreased utilities towards the end of life [89]. Furthermore, health state utility values may not be appropriate given utility levels while progression-free were observed to be stable in CodeBreaK100, whilst evidence suggests quality of life for patients treated with docetaxel decreases over time while in a progression-free health state [51]. Using the time to death approach is consistent with previous NICE TAs in NSCLC [85, 87, 88, 90] and was considered by UK clinical experts to better reflect the experience of their patients with NSCLC. The use of utilities based on health state occupation, as reported in Table 19, are explored in a scenario analysis.

The impact of treatment-related Grade 3+ AEs with an incidence of ≥ 5% in any of the comparator arms (sotorasib, docetaxel, nintedanib and docetaxel) were also included in the model based on the values presented in Table 36. The impact of considering treatment-emergent adverse events are presented in a scenario analysis.

Finally, direct use of reported utility data from CodeBreaK100 likely underestimates the true utility decrement associated with docetaxel and nintedanib plus docetaxel given increased cytotoxicity of these agents and the implications of an hospital-based IV administration, compared with a targeted oral therapy such as sotorasib. UK clinical experts consulted by Amgen verified that a treatment-specific disutility for both docetaxel and nintedanib plus docetaxel would be appropriate to capture in the base case analysis. To inform this, a previous study in advanced NSCLC was used which identified a 0.025 utility decrement associated with IV versus oral administration [104].

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

As described in Section B.3.1 an SLR was conducted to identify published studies for evaluating cost-effectiveness, costs and resource use and health-related quality of life for treatments in NSCLC relevant to the decision problem. Full details on the methodology and findings of the SLR, including search terms, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram and outcomes are detailed in Appendix I.

Studies reporting costs and healthcare resource use in patients with NSCLC and a *KRAS p.G12C* mutation were included in the economic section of SLR. Of the 14 studies identified, 13 studies examined the costs associated with biomarker testing thus were not considered relevant for this appraisal. The topics varied between studies and covered comparison of costs of the existing sequencing techniques, cost of techniques to acquire biological samples or economic evaluations. No evidence was identified for the economic burden of *KRAS p.G12C* mutation or any other KRAS mutation.

B.3.5.1 Intervention and comparators' costs and resource use

The drug acquisition cost per treatment are presented in Table 37 below, with the unit costs for comparators sourced from the electronic market information tool (eMIT) and the British National Formulary (BNF).[80, 82] The sotorasib dose of 960mg per day is consistent with the anticipated license (Appendix C) and the dosing regimen in CodeBreaK100 [52, 56, 58]. The dose of docetaxel and nintedanib plus docetaxel is aligned with UK clinical practice and informed by NHS treatment protocols [105].

Estimation of the monthly cost of treatment are inclusive of the relative dose intensity observed in the respective clinical trial programmes and utilised in the previous NICE appraisal for nintedanib plus docetaxel (NICE TA347) [31, 51, 52, 54, 105]. This ensures that efficacy estimates remain internally consistent with drug utilisation assumptions. Furthermore, with respect to sotorasib, feedback from UK NHS pharmacists confirmed that inclusion of RDI in drug utilisation calculations best reflects clinical practice given the ability to implement dose reductions and the single-strength formulation of sotorasib packs.

Table 37: Unit drug costs

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)[e]
Sotorasib	240 x 120 mg tablets	[redacted]	[redacted]	960mg per day	89.2%[b]	[redacted]
Docetaxel	160 mg per vial	17.95	eMIT[80]	75 mg/m^2 on day of treatment	90.3%[c]	19.93[e]
Nintedanib	120 x 100 mg tablets	2,151.10	BNF[82]	400 mg per day (21-day cycle)[a]	92.1%[d]	1,926.28

Key: BNF, British National Formulary; eMIT, electronic market access tool.
Note:
[a] Nintedanib administered on days when docetaxel is not taken, i.e., 20 days per 21 day cycle,
[a] CodeBreaK100 CSR (01DEC2020), Table 14b-5.1, Exposure to sotorasib (AMG510)
[b] Jänne, 2017[31]
[c] Reck 2014[32]
[d] Docetaxel cycle cost is based on cost per mg x dose per administration (75 mg/m^2) x body surface area (1.81 m2)
[e] calculated from CEM

B.3.5.2 Health-state unit costs and resource use

Administration Costs

The costs of treatment administration for sotorasib, docetaxel and nintedanib plus docetaxel are shown in Table 38. No additional costs are assumed for sotorasib or nintedanib as these are administered orally. According to the SmPC and NHS treatment protocols, the time required per administration of docetaxel is 60 minutes every 3 weeks so is assumed to occur costs associated with the delivery of simple parenteral chemotherapy (SB12Z) [106].

Table 38: Administration costs

Drug	Cost (£)	Source
Sotorasib	0.00	Assumption
Docetaxel (per admin)	241.06	NHS Reference Costs 2018/2019 [SB12Z – OP][106]
Nintedanib	0.00	Assumption

Key: admin, administration; NHS, National Health Service.

Monitoring and Disease Management Costs

Given the limited published literature that explores the resource use associated with previously treated locally advanced or metastatic NSCLC, monitoring and disease management costs are largely informed from assumptions used and accepted in previous NICE TAs and validated with UK clinical experts. In particular, the previous NICE appraisal for nintedanib plus docetaxel (NICE TA347) [51] was used as a primary source.

Disease monitoring and management costs were aligned with the model structure and reflect resource utilisation in both progression-free and post-progression health states, as well as one-off costs associated with treatment initiation and at the point of progression. All costs

were inflated to the 2018/19 cost year to remain consistent with the latest available NHS Reference Costs using the PSSRU HCHS and NHSCII inflation indices [107].

A summary of the costs used in the economic model are presented in Table 39. Full details of the resource assumptions and unit costs are provided in Appendix L.

Table 39: Disease management costs per model cycle

Health state	**Cost per Cycle (£)**	**Source**
Progression-free	77.04	NHS Reference Costs 2018/2019 [106]; PSSRU[107]; aligned with NICE TA347 and TA428
Post-progression	39.98	
Event	**Cost (£)**	**Source**
At treatment initiation	834.25	NHS Reference Costs 2018/2019 [106]; PSSRU[107]; aligned with NICE TA347 and TA428
At progression	116.53	
Key: NHS, National Health Service; PSSRU: Personal Social Services Research Unit		

B.3.5.3 Biomarker testing

KRAS testing is routinely commissioned by NHS in NSCLC [49] and no additional tests beyond those used in the routine diagnostic work up and management of patients with NSCLC are required.

B.3.5.4 Adverse reaction unit costs and resource use

The AEs included in the economic model are previously described in Section B.3.4.6. The unit costs related to the management of AEs events were mainly derived from a previous NICE MTA (TA374) [108] and are broadly consistent with other TAs in this disease area [86-88]. All costs were inflated to the 2018/19 cost year to remain consistent with the latest available NHS Reference Costs using the PSSRU HCHS and NHSCII inflation indices [107].

AE costs used in the base case analysis are summarised in Table 40, below.

Table 40: Adverse events and associated costs

Adverse event	Cost (£)	Source [a]
Decreased neutrophils	204.20	Consistent with NICE TA428 [assumed to be the same as neutropenia].[87]
Decreased white blood cell count	483.38	NICE TA347. Consistent with NICE TA484, TA428, TA520, TA525.[51, 86-88, 109]
Diarrhoea	1,237.97	NICE MTA374. Consistent with NICE TA428.[87, 108]
Dyspnoea	0.00	Assumption from NICE STA319. Consistent with TA484.[86, 110]
Fatigue	2,631.32	NICE MTA374. Consistent with NICE TA428.[87, 108]
Febrile neutropenia	8,325.68	NICE MTA374. Consistent with NICE TA428, TA 484, TA520.[86-88, 108]
Increased ALT	670.79	NICE TA347. Consistent with NICE TA484.[51, 86]
Increased AST	383.96	NICE TA347. Consistent with NICE TA484. [51, 86]
Neutropenia	204.20	NICE MTA347. Consistent with NICE TA520, TA428, and TA484.[86-88, 108]
Pleural effusion	631.94	NICE TA347 [51].
Pneumonia	1,530.40	NHS Reference Costs 2018/2019 [DZ11T Lobar, Atypical or Viral Pneumonia, without Interventions, with CC Score 7-9].[106]

Key: ALT, alanine aminotransferase; AST, aspartate aminotransferase; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.
Note: [a], all costs derived from original sources were inflated to 2018/2019 using the HCS/NHSCII index from PSSRU [107].

B.3.5.5 Subsequent treatment costs

The costs of subsequent treatment were included in the economic model as a one-off cost at disease progression (Table 41).

The treatment regimens and the distributions following progression on docetaxel or nintedanib plus docetaxel were informed by previously accepted assumptions used in NICE TA347 for patients on docetaxel monotherapy and nintedanib plus docetaxel [51]. The distribution of subsequent treatments for patients who progress on sotorasib were informed by UK clinical experts who advised that a greater proportion of patients are likely to be able to tolerate active treatment and that subsequent docetaxel would be preferred over re-challenge with platinum-based chemotherapy regimens.

The duration of subsequent treatments used in the economic model was 3.3 months and is consistent with previous assumptions used in NICE TA347 and TA428 [51, 87]. Detailed costs used to derive subsequent treatments were sourced from eMIT [80] and are presented in Appendix L. Best supportive care was assumed to have no cost in the model.

Table 41: Subsequent treatment costs

Subsequent treatment	BSC	Platinum-based	Docetaxel	Source
Sotorasib (%)	50%	10%	40%	Assumption based on clinical expert feedback
Docetaxel (%)	70%	30%	0%	NICE TA347 - assumption[51]
Nintedanib + docetaxel (%)	70%	30%	0%	NICE TA347 - assumption[51]
Treatment duration (weeks)	14	14	14	NICE TA347, TA428[51, 87]
Cost of subsequent treatment (£) [a]	0	2,835	1,219	Calculation (Appendix L)

Key: BSC, best supportive care; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.
Note: [a], all costs derived from original sources were inflated to 2018/2019 using the HCS/NHSCII index from PSSRU [107].

B.3.5.6 Terminal care costs

A one-off cost is applied to those patients at the moment of dying to reflect the cost of terminal care. The resource consumption reflects treatment received in various care settings and is based on the values used in the NICE MTA for erlotinib and gefitinib (TA374) [108]. These costs are assumed to be the same for all treatments.

The terminal care cost used in the model is summarised in Table 42 and detailed calculations are presented in Appendix L. All costs were inflated to the 2018/19 cost year to remain consistent with the latest available NHS Reference Costs using the PSSRU HCHS and NHSCII inflation indices [107].

Table 42: Terminal care costs

	Cost (£)	Source
Terminal care	3,759.73	See Appendix L

B.3.6. Summary of base-case analysis inputs and assumptions

B.3.6.1 Summary of base-case analysis inputs

A summary of the variables applied in the economic model is presented in Appendix M.

B.3.6.2 Assumptions

A summary of the base case assumptions and rationale are detailed in Table 43.

Table 43: Summary of model assumptions and rationale

Category	Base case analysis	Rationale
Time horizon	20-year (lifetime) time horizon	The time horizon was considered sufficient to capture all costs and benefits over the lifetime of the modelled population
Comparator	Primary comparator was docetaxel; nintedanib + docetaxel was included as an alternative	Docetaxel was regarded as the appropriate standard of care for previously treated (2L+) NSCLC by UK clinicians. Nintedanib plus docetaxel was explored as a secondary comparator as it is considered for some patients with adenocarcinoma although its use in clinical practice is highly variable.
Population	Population is adult with advanced NSCLC with *KRAS p.G12C* mutation who have been previously treated for advanced or metastatic NSCLC	Aligned with the anticipated licensed indication. The population is generalisable to UK clinical practice and reflects a population with minimal treatment options.
Efficacy	MAIC is a valid approach to model efficacy for sotorasib vs docetaxel monotherapy	Given that sotorasib is a single arm trial, an unanchored MAIC offered the most robust method of comparison to account for potential differences in prognostic characteristics. Multiple MAIC models and unadjusted analyses are presented to fully explore the uncertainty of results.
Efficacy	The Flatiron dataset in the KRAS mutant population provides a reasonable alternative to using the SELECT-1 data to model docetaxel monotherapy.	The Flatiron dataset provides real world evidence on chemotherapy standard of care in this relevant patient population and demonstrated consistency between the outcomes for patients with *KRAS p.G12C* mutant NSCLC and those with KRAS mutant NSCLC. The basket of chemotherapy regimens is likely to sufficiently reflect outcomes with docetaxel.
Efficacy	Efficacy of nintedanib plus docetaxel versus docetaxel monotherapy in adenocarcinoma patients is consistent in *KRAS p.G12C* mutant NSCLC as in the LUME Lung 1 trial cohort.	Limited evidence exists to inform the comparison of sotorasib versus nintedanib plus docetaxel and the indirect method of establishing relative efficacy from the LUME Lung-1 trial was considered appropriate by UK clinicians and health economic experts during an advisory board. The use of the non-squamous cohort from LUME Lung-1 best aligns with CodeBreaK100 cohort and alternative data indicate outcomes are similar with non-

Category	Base case analysis	Rationale
		targeted therapy irrespective of KRAS status.
Efficacy	A piecewise model is appropriate to extrapolate long-term survival for sotorasib vs nintedanib plus docetaxel and the periods chosen were justified	The proportional hazards assumptions were clearly violated and there was a distinct directional change in the instantaneous hazards. A piecewise model was therefore required for long-term extrapolation of survival data. An additional analysis applying a single HR was explored to assess the sensitivity of the model results to this assumption.
HRQOL	Quality of life is appropriately captured using time to death approach	Time to death utilities have been used in previous NICE appraisals in this disease area and were considered by clinical experts to better reflect the quality-of-life deterioration over time in NSCLC, particularly at the end of life.
HRQOL	Treatment specific utility decrement for IV docetaxel	Direct use of reported utility data from CodeBreak100 may underestimate utility decrement associated with a cytotoxic chemotherapy with IV administration. An additional treatment-specific utility decrement identified from the literature is applied to account for this.
Adverse events	Grade 3+ treatment-related adverse events with an incidence of ≥ 5% in any study were included	Treatment-related adverse events are more specific and relevant to capture in the model than treatment emergent adverse events. Sensitivity analysis exploring other treatment emergent adverse events were been conducted.
Costs	No costs are assumed for KRAS mutation testing	KRAS testing is routinely commissioned by NHS in NSCLC [49] and no additional tests beyond those used in the routine diagnostic work up and management of patients with NSCLC are required
Costs	Disease management costs are generalisable to the UK	Disease management costs are consistent with previous NICE appraisals in NSCLC and were considered by UK clinicians to be reflective of health care resource utilisation in this disease area.

Category	Base case analysis	Rationale
Costs	Treatment duration approach is appropriate	The treatment duration for sotorasib was applied to PFS using patient level data for simplicity and was reasonable. The treatment duration for docetaxel was set equal to PFS as it is not expected to be a major cost driver. The treatment duration for nintedanib was set equal to PFS and validated using data from the LUME-Lung 1 trial
Key: 2L, second-line; AE, adverse event; HRQOL, health-related quality of life; IV, intravenous; MAIC, matching-adjusted indirect comparison; NHS, National Health Service; NSCLC, non-small-cell lung cancer; PFS, progression-free survival; OS, overall survival; SAF, safety analysis set; TTD, time to treatment discontinuation		

B.3.6.3 Base-case results

Sotorasib versus docetaxel (primary comparator)

In the model base case where docetaxel is considered the comparator, discounted results are presented in **Error! Not a valid bookmark self-reference.**. Using a 20-year time horizon, the incremental total life-year gain of sotorasib versus docetaxel was [redacted] years. The discounted incremental costs of £[redacted] and incremental QALYs of [redacted] resulted in an incremental cost-effectiveness ratio (ICER) of £47,146 versus docetaxel. This is below the willingness-to-pay threshold of £50,000 per QALY for end-of-life medicines.

Table 44: Deterministic base-case results: sotorasib versus docetaxel

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Docetaxel	[redacted]	[redacted]	[redacted]				
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**47,146**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALYs, quality-adjusted life years.							

In Appendix J, the clinical outcomes of the analysis are presented and compared with the clinical trial results. Appendix J also includes detailed tabulated disaggregated results of the base case analysis.

B.3.7. Sensitivity analyses

B.3.7.1 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was undertaken to explore the uncertainty of all model parameters and their associated impact on cost-effectiveness results. 1,000 iterations were used to ensure convergence. The total costs and QALYs were recorded for each iteration and averaged. PSA results for the comparison to docetaxel are presented in Table

45. The deterministic ICER for sotorasib compared to docetaxel (£47,146) is in line with the PSA results of £46,707 confirming that the results are robust to parameter uncertainty.

Table 45: Probabilistic base-case results: sotorasib versus docetaxel

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Docetaxel	[redacted]	[redacted]	[redacted]				
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**46,707**

Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALY, quality-adjusted life year.

Figure 41 represents the scatter plot of the incremental costs and QALYs from the PSA results based on 1,000 iterations. As shown in the cost-effectiveness acceptability curve (Figure 42), sotorasib has a [redacted] probability of being cost-effective versus docetaxel considering the £50,000 WTP threshold.

Figure 41: Cost-effectiveness plot for sotorasib versus docetaxel

Key: QALY, quality-adjusted life year; WTP, willingness-to-pay.

Figure 42: Cost-effectiveness acceptability plot for sotorasib versus docetaxel

Key: QALY, quality-adjusted life year; WTP, willingness-to-pay.

B.3.7.2 Deterministic sensitivity analysis

A series of one-way sensitivity analyses (OWSA) were performed to evaluate the sensitivity of the model ICER to individual inputs, holding all else constant. In the deterministic sensitivity analysis, the upper and lower bounds of a parameter were taken from their 95% confidence intervals if these were available from the data source. When such information was not available, the upper and lower bounds were assumed to be within ± 20% for all base case values.

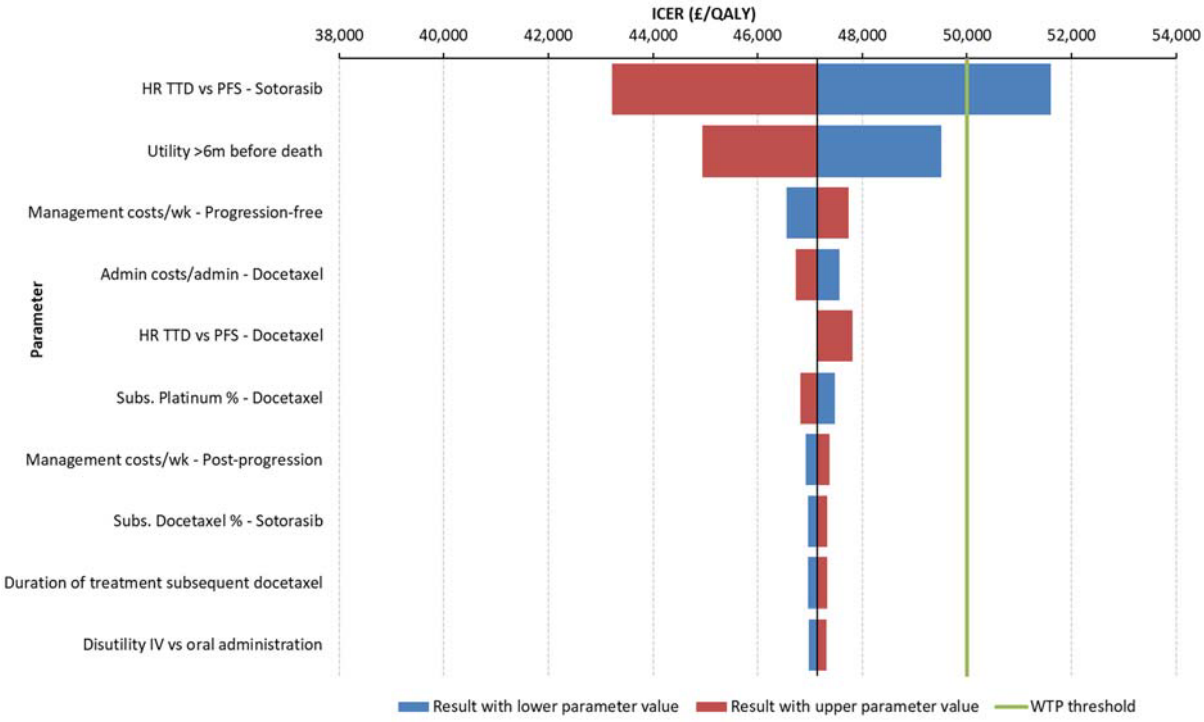
A summary of the variables and distributions applied in the economic model is presented in Appendix M, including references to the corresponding sections in the submission where each is explained in more detail.

Each value was varied based on its uncertainty parameters.

Figure 43 presents the tornado diagram for the top ten parameters in terms of ICER impact which were varied in the OWSA. Parameters are shown in descending order of ICER sensitivity.

These results demonstrate that the ICER is most sensitive to varying the hazard ratio applied to PFS to model sotorasib treatment duration, followed by varying the time to death utility value for > 6 months prior to death. All other analyses had little impact on the results.

Figure 43: One-way sensitivity analysis for sotorasib versus docetaxel



Key: HR, hazard ratio; PFS, progression-free survival; TTD, time to treatment discontinuation.

B.3.7.3 Scenario analysis

B.3.7.3.1 Key scenarios

A comparison of sotorasib to docetaxel based on the Flatiron real world dataset are presented in **Error! Not a valid bookmark self-reference.**. Using a 20-year time horizon, the incremental total life-year gain of sotorasib versus docetaxel was [redacted] years. The discounted incremental costs of £[redacted] and incremental QALYs of [redacted] resulted in an ICER of £39,773 versus docetaxel. This alternative approach using an alternative data source, which included patients closely aligned with the CodeBreaK100 population in terms of prior treatment history, suggests that the primary base case analysis is plausibly conservative.

Table 46: Deterministic results: sotorasib versus docetaxel using Flatiron data

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Docetaxel	[redacted]	[redacted]	[redacted]				
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**39,773**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALYs, quality-adjusted life years							

The secondary comparison of sotorasib vs nintedanib plus docetaxel, assuming the [redacted] for sotorasib and the list price for nintedanib are presented in Table

47. Using a 20-year time horizon, the incremental total life years gain was [redacted] years. The discounted incremental costs of £[redacted] and incremental QALYs of [redacted] resulted in an ICER of £35,779 versus nintedanib + docetaxel. Although the incremental gain in QALYs with sotorasib in this analysis was less than in the comparison against docetaxel monotherapy, the costs of nintedanib increased the total costs of the comparator, leading to a lower ICER compared with docetaxel monotherapy.

Table 47: Deterministic results: sotorasib versus nintedanib + docetaxel

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Nintedanib + docetaxel	[redacted]	[redacted]	[redacted]				
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**35,779**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALY, quality-adjusted life year.							

Additional sensitivity analyses to fully explore the uncertainty around the secondary comparison of nintedanib plus docetaxel are presented in Appendix O.

B.3.7.3.2 Other scenarios

An extensive range of scenario analyses exploring alternative comparative effectiveness estimates, as well as cost and resource inputs, HRQoL and model settings were conducted. A description of the various scenario analyses and the impact on the cost-effectiveness results are presented in Table 48.

Scenario analysis was used to test the selection of parametric fitting to PFS and OS. The results of the analysis remained similar to the base case for the second-best fitting OS and PFS distribution (generalised gamma) resulting in an ICER of £45,612 per QALY. Using the log-logistic distribution (3rd best-fitting) resulted in an increase in incremental costs to £[redacted] and a reduction in incremental QALYs to [redacted], with a resulting ICER of £53,606.

The scenario with the largest increase on the ICER was the use of the unadjusted (naïve) comparison of sotorasib with docetaxel from the SELECT-1 clinical trial whereby the incremental costs increased to £[redacted] and the incremental QALYs reduced to [redacted] and resulted in an ICER of £53,794 per QALY. Alternatively, when utilising the MAIC model which includes all available covariates, the ICER reduces to £39,645

Excluding relative dose intensity resulted in increased incremental costs, with the ICER reported at £52,757, a 11.9% increase compared to the base case. Including drug wastage resulted in an ICER increase of 6.5% to £50,216.

Capturing quality of life using health state utilities by progression status resulted in a reduction in QALY benefit compared to time to death utilities used in the base case ([redacted] versus [redacted], respectively) and the ICER increased to £51,079.

Adjusting the time horizon of the analysis to 15 years produced a small reduction in incremental costs and QALYs, with a reported ICER 2.2% higher than the base case..

Table 48 Scenario analysis results

Scenario	Rationale/Justification	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
15-year time horizon	To investigate the impact on model results of reducing the model timeframe.	[redacted]	[redacted]	48,197
Generalised gamma distribution selected to estimate long-term OS and PFS projections	The generalised gamma was the 2nd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more favourable estimation of survival.	[redacted]	[redacted]	45,612
Log logistic distribution selected to estimate long-term OS and PFS projections	The log-logistic was the 3rd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more conservative estimation of survival.	[redacted]	[redacted]	53,606
Joint (unrestricted) lognormal distribution selected to estimate long-term PFS	Although BIC consistently favoured the restricted versus unrestricted joint fit this scenario tests an alternative parametric distribution using the unrestricted lognormal for PFS which was the best fitting unrestricted model based on BIC criteria.	[redacted]	[redacted]	52,495
Adjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1 using all available covariates	To test the robustness of the MAIC using alternative MAIC model where all available covariates are considered. A joint (restricted) lognormal distribution was used per the base case analysis and based on the survival analysis selection criteria outlined in Appendix N.	[redacted]	[redacted]	39,645
Unadjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1	To investigate the impact of a naïve comparison to the SELECT-1 clinical trial. A joint (restricted) lognormal distribution was used per the base case analysis.	[redacted]	[redacted]	53,794
Unadjusted sotorasib from CodeBreaK100 vs. ATT-adjusted docetaxel from Flatiron	Alternative data source which included patients closely aligned with the CodeBreaK100 population in terms of prior treatment from the real-world	[redacted]	[redacted]	39,773

	Flatiron dataset to test the robustness of the results in the base case analysis			
MAIC-adjusted TTD curve from CodeBreaK100	To test the impact of an alternative approach to estimate long-term treatment duration.	[redacted]	[redacted]	50,810
HR of sotorasib vs. docetaxel = 1 after 5 years	In the base-case PFS and OS were modelled based on parametric survival distributions fit to survival data from CodeBreaK100 and SELECT-1, combined with age- and sex-matched general population mortality. This scenario explicitly limits the duration of benefit to 60 months.	[redacted]	[redacted]	49,956
Apply health state utilities by progression status	To test the impact of an alternative method for measuring health state utilities as described in Section B.3.4.5	[redacted]	[redacted]	51,079
Treatment-emergent AEs	To test the impact of utilising treatment-emergent adverse events for sotorasib and docetaxel as described in Section B.3.3.7	[redacted]	[redacted]	47,495
15-year time horizon	Model timeframe set to 15-years to test the impact of shorter time horizon	[redacted]	[redacted]	48,197
Include drug wastage	To test the impact of potential drug wastage in clinical practice by estimating drug acquisition costs based on total packs as opposed to treatments received	[redacted]	[redacted]	50,216
Exclude RDI	To test the impact of not capturing RDI on drug utilisation calculations	[redacted]	[redacted]	52,757
1.5% discount rate for costs and efficacy	To investigate the alternative discount rate suggested by the NICE Guide to Technology Appraisal. A reduced discount rate of 1.5% is consistent with the Treasury Green Book and is being considered in the ongoing NICE Methods Review consultation.	[redacted]	[redacted]	44,505

Key: ATT, average treatment effect of the treated; HR, hazard ratio; OS, overall survival, ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALY, quality-adjusted life year; RDI, relative dose intensity.

B.3.7.4 Summary of sensitivity analyses results

An extensive range of sensitivity and scenario analyses were conducted to test the robustness of the model inputs and structural assumptions of the economic analyses. Overall, the base case results were robust to most parameters and structural assumptions.

The OWSA found that the results were sensitive to variation in the hazard ratio applied to model sotorasib treatment duration and varying time to death utilities for up to 6 months prior to death. However, all other analyses had minimal impact.

The scenario analyses highlight that the model is robust to changes in key modelling assumptions, with the majority of scenarios remaining below the £50,000 willingness-to pay-threshold for end-of-life medicines. Importantly, key sensitivity and scenario analyses exploring the impact of assumptions on the relative efficacy suggest that the base case ICER is robust and potentially conservative.

B.3.8. Subgroup analysis

Subgroup analyses were not included in this analysis.

B.3.9. Validation

B.3.9.1 Validation of cost-effectiveness analysis

Consistency with trial and literature

As summarised in Appendix J, modelled median PFS and OS are similar to the reported medians in the MAIC adjusted CodeBreak100 clinical trial and the docetaxel arm of SELECT-1.

The clinical plausibility of the parametric models used in the economic analysis were also evaluated by considering the predicted OS landmark results at timepoints of 1-year, 5-years and 10-years based on clinical expert opinion. The base case jointly fitted (restricted) log-normal distribution was considered to be clinically valid and reflective of the expected survival for the population under consideration.

Alternative data sources – supplementary Flatiron Health real-world study

A supplemental, alternative approach was also explored using real-world data from the Flatiron Health database and which may be used as confirmatory validation and to support the robustness of the results generated in the MAIC. As reported in Appendix J, the clinical outcomes generated using this alternative data source were consistent with the conclusions of the MAIC analysis and underline the robustness of the analyses presented.

B.3.9.2 Quality control

The model was subjected to systematic examination of calculations and Visual Basic for Applications coding accuracy by modelling experts in and outside of Amgen who did not build the model. A cost-effectiveness model verification checklist guided the quality control process, which, among others checks, included extreme value analysis and tracing of calculations.

B.3.10. Interpretation and conclusions of economic evidence

B.3.10.1 B.3.11.1 Comparison with published economic literature

To our knowledge this is the first economic evaluating sotorasib; therefore, a comparison of cost-effectiveness results with published literature is not possible.

B.3.10.2 B.3.11.2 Relevance of the economic evaluation to all patients who could potentially use the technology as identified in the decision problem

Sotorasib is a highly innovative, first-in-class targeted therapy, anticipated to receive a conditional marketing authorisation for the treatment of patients with *KRAS p.G12C* -mutated NSCLC following failure of prior therapy. Although data are currently limited to a phase 2 single-arm trial and indirect treatment comparisons, this collective early evidence indicates that sotorasib is highly effective and well tolerated, and provides superior PFS and OS compared with the current primary non-targeted standard of care therapy. Sotorasib should therefore be made available as early as possible to address the high and urgent unmet needs of these patients.

B.3.10.3 B.3.11.3 Generalisability of the analysis

The efficacy and safety of sotorasib observed in the CodeBreak 100 trial, and its comparative effectiveness in the indirect comparisons, are generalisable to UK clinical practice. The economic analyses are reflective of the decision problem and well-align with the anticipated treatment pathway in the UK.

Furthermore, the model was developed using cost sources most relevant to the NHS in England.

B.3.10.4 Limitations of the economic evaluation

Given the context of a disease with no other targeted therapies, high and urgent unmet needs and with limited available data with which to make comparisons, there are limitations to the current evidence base. By necessity, the indirect treatment comparison of sotorasib and the relevant primary comparator is based on an unanchored MAIC, which is also associated with uncertainty; however, given the consistency of the results across different MAIC models and an alternative approach to derive an indirect estimate. There is limited data for sotorasib or the comparators in patients with squamous histology, and the trial

excluded patients with ECOG PS 2, who could particularly benefit from its superior safety profile compared with existing non-targeted standard of care therapies

B.3.10.5 Conclusions

- Sotorasib is a novel and clinically effective treatment option for *KRAS p.G12C* NSCLC which significantly improves life-years and QALYs compared with docetaxel and nintedanib plus docetaxel
- The primary analysis is well-aligned to the decision problem and reflective of UK clinical practice. Results are generated based on MAIC adjusted comparison via the docetaxel arm of SELECT-1 and an indirect comparison of nintedanib plus docetaxel via LUME Lung 1
- The most clinically plausible extrapolations of PFS and OS data were selected for the base case analyses and extensive scenario analyses were presented with only a small impact to the results of the analyses
- The base case modelling approach, including structure, costs included, and utility values applied, is consistent with those accepted in previous TAs for treatments of NSCLC
- In the primary comparison base case analysis the ICER for sotorasib versus docetaxel was £47,146 per QALY gained; results of the alternative analysis using real-world data versus a basket of chemotherapy comparators reduced the ICER to £39,773 per QALY gained, suggesting the base case ICER may be conservative.
- In the secondary comparison, using the list price of nintedanib, the ICER for sotorasib versus nintedanib plus docetaxel reduced to £35,779 per QALY gained
- Sotorasib meets the end-of-life criteria and likely represents a cost-effective use of NHS resources.

References

1. Cancer Research UK. *Lung cancer statistics*. 2021 [cited 2021 26/06/2021]; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer.
2. Cancer Research UK. *Types of lung cancer*. 2021 [cited 2021 26/06/2021]; Available from: https://www.cancerresearchuk.org/about-cancer/lung-cancer/stages-types-grades/types.
3. Planchard, D., et al., *Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up (updated 15 September 2020). .* Ann Oncol, 2020. **29**(Suppl 4): p. iv192-iv237.
4. Román, M., et al., *KRAS oncogene in non-small cell lung cancer: clinical perspectives on the treatment of an old target. .* Mol Cancer, 2018. **17**(1): p. 33.
5. Amgen Inc, *Data on File: A Retrospective Cohort Study of Patient Characteristics and Treatment Outcomes Among Patients With KRAS p.G12C Mutation-positive Advanced Non-small Cell Lung Cancer in the AACR Project GENIE Database (Study 20180277)*. 2020.
6. Cox, A., et al., *Drugging the undruggable Ras: mission possible?* Nat Rev Drug Discov. , 2014. **13**(11): p. 828-851.
7. Al-Batran, S., et al., *The impact of docetaxel-related toxicities on health-related quality of life in patients with metastatic cancer (QoliTax).* Annals of Oncology, 2015. **26**: p. 1244-48.
8. Yared, J. and K. Tkaczuk, *Update on taxane development: new analogs and new formulations.* Drug Design, Development and Therapy, 2012. **6**: p. 371-84.
9. MHRA. *Innovative Licensing and Access Pathway*. 2021 [cited 2021 26/06/2021]; Available from: https://www.gov.uk/guidance/innovative-licensing-and-access-pathway.
10. MHRA. *Promising innovative medicine (PIM) designation.* 2021; Available from: https://www.gov.uk/guidance/apply-for-the-early-access-to-medicines-scheme-eams#promising-innovative-medicine-pim-designation.
11. ClinicalTrials.gov. *Study to Compare Sotorasib With Docetaxel in Non Small Cell Lung Cancer (NSCLC) (CodeBreak 200) [NCT04303780]*. 2020 27 August 2020]; Available from: https://clinicaltrials.gov/ct2/show/NCT04303780.
12. Amgen Inc, *Draft Summary of Product Characteristics: LUMYKRAS (Sotorasib).* 2021.
13. NICE. *NICE Pathways: Lung cancer overview*. 2021 [cited 2021 26/06/2021]; Available from: https://pathways.nice.org.uk/pathways/lung-cancer.
14. Amgen Inc, *Data on File: Amgen NSCLC UK Advisory Board - 11th February 2021*. 2021.
15. Aggarwal, S., et al., *Clinicopathological Characteristics, Treatment Patterns, and Outcomes in Patients with KRAS p.G12C Mutant Advanced Non-Small Cell Lung Cancer in the Flatiron Health-Foundation Medicine Clinico-Genomic Database. Poster Number: 1339P*, in *ESMO Virtual Congress 2020*.
16. NICE, *Lung cancer: diagnosis and management. NICE Guideline NG122.* 2019.
17. Amgen Inc, *Data on File: NICE Scientific Advice: EUnetHTA Report*. 2020.
18. Amgen inc, *Lumakras US FDA Label*. 2021.
19. ACCR Project GENIE Consortium, *AACR Project GENIE: Powering Precision Medicine through an International Consortium.* Cancer Discov, 2017. **7**(8): p. 818-31.
20. Karnoub, A. and R. Weinberg, *Ras oncogenes: split personalities.* Nat Rev Mol Cell Biol., 2008. **9**(7): p. 517-31.
21. Amgen Inc, *Press Release: FDA Approves LUMAKRAS™ (Sotorasib), The First And Only Targeted Treatment For Patients With KRAS G12C-Mutated Locally Advanced Or Metastatic Non-Small Cell Lung Cancer*. 2021.
22. Royal College of Physicians, *National Lung Cancer Audit: Spotlight report on molecular testing in advanced lung cancer*. 2020.
23. Simanshu, D., D. Nissley, and F. McCormick, *RAS Proteins and Their Regulators in Human Disease.* Cell, 2017. **170**(1): p. 17-33.
24. Jones, R., et al., *Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer.* Br J Cancer, 2017. **116**(7): p. 923-9.
25. Goulding, R., et al., *KRAS mutation as a prognostic factor and predictive factor in advances/metastatic non-small cell lung cancer: a systematic literature and meta-analysis.* Cancer Treat Res Commun., 2020. **24:100200**.
26. Biernacka, A., et al., *The potential utility of re-mining results of somatic mutation testing: KRAS status in lung adenocarcinoma.* Cancer Genet, 2016. **209**(5): p. 195-98.

27. MHRA. *Orphan medicinal products.* 2021 26/06/2021]; Available from: https://www.gov.uk/guidance/orphan-medicinal-products-in-great-britain.
28. Gainor, J., et al., *ALK Rearrangements Are Mutually Exclusive with Mutations in EGFR or KRAS: An Analysis of 1,683 Patients with Non–Small Cell Lung Cancer.* Clin Cancer Res, 2013. **19**(15).
29. Martorell, P., et al., *Co-existence of EGFR, KRAS, BRAF, PIK3CA mutations and ALK re-arrangement in a comprehensive cohort of 326 consecutive Spanish non-squamous NSCLC patients.* Clinical Lung Cancer, 2017. **18**(6): p. e395-e402.
30. Scheffler, M., et al., *K-ras Mutation Subtypes in NSCLC and Associated Co-occuring Mutations in Other Oncogenic Pathways.* Journal of Thoracic Oncology, 2019. **14**(4): p. 606-16.
31. Jänne, P., et al., *Selumetinib Plus Docetaxel Compared With Docetaxel Alone and Progression-Free Survival in Patients With KRAS-Mutant Advanced Non–Small Cell Lung Cancer: The SELECT-1 Randomized Clinical Trial.* JAMA, 2017. **317**(18): p. 1844-53.
32. Reck, M., et al., *Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial.* Lancet Oncol, 2014. **15**: p. 143-55.
33. Arbour, K., et al., *Clinical characteristics and anti-PD-(L)1 tretament outcomes of KRAS-G12C mutant lung cancer compared to other molecular subtypes of KRAS-mutant lung cancer.* Journal of Clinical Oncology, 2020. **38**(15).
34. Cui, W., et al., *Real world outcomes in KRAS G12C mutation positive non-small cell lung cancer.* Lung Cancer, 2020. **146**: p. 310-17.
35. Griesinger, F., et al., *Treatment and outcome of a real-world cohort of patients with advanced, non-squamous NSCLC and KRAS mutations with a special focus on KRAS G12C.* Ann Oncol. , 2020. **31**: p. S872.
36. Amgen Inc, *Data on File: A Retrospective cohort Study of Patient Characteristics and Treatment Outcomes Among Patients With KRAS p.G12C Mutation-positive Advanced Non-small Cell Lung Cancer in the Flatiron Health-Foundation medicines Clinico-Genomics Database (Study 20200097).* 2020.
37. Amgen Inc, *Data on File: A Retrospective Cohort Study of Patient Characteristics and Treatment Outcomes Among Patients With Advanced Non-small Cell Lung Cancer in the Flatiron Health-Foundation Medicine Clinico-Genomics Database (Study 20200132).* 2020.
38. Cella, D. and J. Patel, *Improving Health-Related Quality of Life in Non-Small Cell Lung Cancer with Current Treatment Options.* Clinical Lung Cancer, 2008. **9**(4): p. 206-12.
39. Whisenant, M., et al., *What do patients with non-small-cell lung cancer experience? Content domain for the MD Anderson Symptom Inventory for Lung Cancer.* JCO Oncol Pract, 2020. **16**: p. e1151-e1160.
40. Gebbia, V., *Does an optimal therapeutic sequence exist in advanced non-small cell lung cancer?* Expert Opin. Pharmacother., 2008. **9**(8): p. 1321-37.
41. Crawford, J., D. Dale, and G. Lyman, *Chemotherapy-Induced Neutropenia: Risks, Consequences, and New Directions for Its Management.* Cancer, 2004. **100**(2): p. 228-37.
42. Talbot, T., et al., *The burden of neutropenic sepsis in patients with advcanced non-small cell lung cancer tretaed with single agent docetaxel: a retrospective study.* Lung Cancer, 2017. **113**: p. 115-20.
43. Wailoo, A., A. Sutton, and A. Morgan, *The risk of febrile neutropenia in patients with non-small-cell lung cancer treated with docetaxel: a systematic review and meta-analysis.* British Journal of Cancer, 2009. **100**: p. 436-41.
44. Ettinger, S., et al., *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®): Non-Small Cell Lung Cancer. Version 6.2020 — June 15, 2020.* 2020.
45. NHS England, *National Optimal Lung Cancer Pathway For Suspected and Confirmed Lung Cancer: Referral to Treatment*. 2020.
46. NICE, *Entrectinib for treating NTRK fusion-positive solid tumours. NICE TA644.* 2020.
47. HM Government, *GENOME UK: The future of healthcare*. 2020.
48. NHS, *The NHS Long Term Plan.* 2019.
49. NHS England. *2020/21 National Genomic Test Directory for Cancer*. 2020; Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/.
50. Hanna, N. and S.M.G. Temin, *Therapy for Stage IV Non–Small-Cell Lung Cancer Without Driver Alterations: ASCO and OH (CCO) Joint Guideline Update Summary.* JCO Oncol Pract, 2020. **18**(8): p. e844-e848.
51. NICE, *Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer. NICE TA347.* 2015.

52. Amgen Inc, *Data on File: Clinical Study Report: CodeBreaK100 Primary Analysis (September 2020 data cut).* 2020.
53. Hong, D., et al., *KRASG12C Inhibition with Sotorasib in Advanced Solid Tumors.* N Engl J Med, 2020. **383**(13): p. 1207-17.
54. Amgen Inc, *Data on File: Clinical Study Report: CodeBreaK100 Efficacy data (01 December Data cut)* 2021.
55. Amgen Inc, *Data on File: CodeBreaK100 Safety data (01 December 2020 data cut).* 2021.
56. Skoulidis, F., et al., *Sotorasib for Lung Cancers with KRAS p.G12C Mutation.* N Engl J Med, 2021. **384**(25): p. 2371-81.
57. Amgen Inc, *Data on file: Clinical Study Report: CodeBreaK100 Supplementary PROs Primary analysis (September 2020 data cut).* 2020.
58. Amgen Inc, *A Phase 1/2, Open-label Study Evaluating the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics, and Efficacy of AMG 510 Monotherapy in Subjects With Advanced Solid Tumors With KRAS p.G12C Mutation and AMG 510 Combination Therapy in Subjects With Advanced NSCLC With KRAS p.G12C Mutation (CodeBreak 100). Amgen Study 20170543 Protocol.* 2018.
59. NICE. *Osimertinib for treating locally advanced or metastatic EGFR T790M mutation-positive non-small-cell lung cancer. NICE TA614 (replaced by NICE TA 653).* 2016 [cited 2021 26/06/2021]; Available from: Osimertinib for treating locally advanced or metastatic EGFR T790M mutation-positive non-small-cell lung cancer.
60. NICE, *Crizotinib for treating ROS1-positive advanced non- small-cell lung cancer. NICE TA 529.* 2018.
61. NICE, *Lorlatinib for previously treated ALK-positive advanced non-small-cell lung cancer. NICE TA 628.* 2020.
62. NICE, *Entrectinib for treating ROS1-positive advanced non- small-cell lung cancer. NICE TA643.* 2020.
63. EMA, *Guideline on the evaluation of anticancer medicinal products in man (EMA/CHMP/205/95 Rev 4).* 2017.
64. Blumenthal, G., et al., *Overall Response Rate, Progression-Free Survival, and Overall Survival With Targeted and Standard Therapies in Advanced Non–Small-Cell Lung Cancer: US Food and Drug Administration Trial-Level and Patient-Level Analyses.* J Clin Oncol, 2015. **33**(1008-14).
65. Clarke, J., X. Wang, and N. Ready, *Surrogate clinical endpoints to predict overall survival in non-small cell lung cancer trials—are we in a new era?* Transl Lung Cancer Res 2015. **4**(6): p. 804-8.
66. Eisenhauer, E., et al., *New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1).* European Journal of Cancer, 2009. **45**: p. 228-47.
67. NICE, *Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer. NICE TA403.* 2016.
68. Eli Lilly & Co., *Cyramza US Prescribing Information.* 2015.
69. Claret, L., et al., *Prediction of Overall Survival or Progression Free Survival by Disease Control Rate at Week 8 Is Independent of Ethnicity: Western Versus Chinese Patients With First-Line Non-Small Cell Lung Cancer Treated With Chemotherapy With or Without Bevacizumab.* The Journal of Clinical Pharmacology, 2013. **54**(3): p. 253-57.
70. Lara, P., et al., *Disease Control Rate at 8 Weeks Predicts Clinical Benefit in Advanced Non–Small-Cell Lung Cancer: Results From Southwest Oncology Group Randomized Trials.* J Clin Oncol 2008. **26**: p. 463-67.
71. Sterne, J., et al., *ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions.* BMJ 2016, 2016. **355**: p. i4919.
72. Phillippo, D., et al., *NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE.* 2016.
73. Chu, P., et al., *Matching-adjusted indirect comparison: entrectinib versus crizotinib in ROS1 fusion-positive non-small cell lung cancer.* J. Comp. Eff. Res. , 2020. **9**(12): p. 861-76.
74. Claxton, l., et al., *Ceritinib for Untreated Anaplastic Lymphoma Kinase-Positive Advanced Non-Small-Cell Lung Cancer: An Evidence Review Group Evaluation of a NICE Single Technology Appraisal.* PharmacoEconomics, 2019. **37**: p. 645-654.
75. Hamlos, B., et al., *A Matching-Adjusted Indirect Comparison of Pembrolizumab + Chemotherapy vs. Nivolumab + Ipilimumab as First-Line Therapies in Patients with PD-L1 TPS ≥1% Metastatic NSCLC.* Cancers, 2020: p. 3648.

76. Reckamp, K., et al., *Comparative efficacy of brigatinib versus ceritinib and alectinib in patients with crizotinib-refractory anaplastic lymphoma kinase-positive non-small cell lung cancer.* Curr Med Res Opin, 2019. **35**(4): p. 569-576.
77. Boehringer Ingelheim, *Summary of Product Characteristics: Vargatef 100 mg soft capsules.* 2021.
78. Rohatgi, A. *WebPlotDigitizer Web based tool to extract data from plots, images, and maps*. 2020 26/06/2021]; Available from: https://automeris.io/WebPlotDigitizer/.
79. Guyot, P., et al., *Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves.* BMC Medical Research Methodology, 2012. **12**: p. 9.
80. GOV.UK. *Drugs and pharmaceutical electronic market information tool (eMIT)*. 2021 28 April 2021]; Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
81. MIMS. *Monthly Index of Medical Specialities*. 2021 28 April 2021]; Available from: https://www.mims.co.uk/.
82. Royal Pharmaceutical Society. *British National Formulary [BNF Online]*. 2021 28 April 2021]; Available from: https://about.medicinescomplete.com/publication/british-national-formulary/.
83. NICE. *Guide to the methods of technology appraisal 2013 [Internet]*. 2013 19 November 2020]; Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781.
84. NICE, *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer: Final scope.* 2021.
85. NICE. *Pemetrexed maintenance treatment for non-squamous non-small-cell lung cancer after pemetrexed and cisplatin [TA402]*. 2016 10 March 2021]; Available from: https://www.nice.org.uk/guidance/ta402.
86. NICE. *Nivolumab for previously treated non-squamous non-small-cell lung cancer [TA484]*. 2017 10 March 2021]; Available from: https://www.nice.org.uk/guidance/ta484.
87. NICE. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy [TA428]*. 2017 10 March 2021]; Available from: https://www.nice.org.uk/guidance/ta428.
88. NICE. *Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy [TA520]*. 2018 10 March 2021]; Available from: https://www.nice.org.uk/guidance/ta520.
89. Van den Hout, W., et al., *Cost – Utility Analysis of Short- Versus Long-Course Palliative Radiotherapy in Patients With Non – Small-Cell Lung Cancer.* Journal of the National Cancer Institute, 2006. **98**(24).
90. NICE. *Pembrolizumab with pemetrexed and platinum chemotherapy for untreated, metastatic, non-squamous non-small-cell lung cancer [TA557]*. 2019 21 May 2021]; Available from: https://www.nice.org.uk/guidance/ta557.
91. Van Hout, B., et al., *Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets.* Value in health, 2012. **15**(5): p. 708-715.
92. NICE. *Reviewing our methods for health technology evaluation: consultation [Consultation closed - 18 December 2020]*. 2020 17 May 2021]; Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation.
93. Mosteller, R., *Simplified calculation of body-surface area.* The New England journal of medicine, 1987. **317**(17): p. 1098-1098.
94. Latimer, N., *NICE Decision Support Unit - TSD 14: Aurvival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data.* Report by the Decision Support Unit, 2011.
95. NICE. *Nintedanib Committee Papers [NICE TA 347]*. 2015 26/06/2021]; Available from: https://www.nice.org.uk/guidance/ta347/history.
96. Clopper, C.J. and E.S. Pearson, *The use of confidence or fiducial limits illustrated in the case of the binomial.* Biometrika, 1934. **26**(4): p. 404-413.
97. Office for National Statistics. *National life tables: UK*. 2021 28 April 2021]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables.
98. EuroQoL. *EuroQol Instruments - EQ-5D-3L, EQ-5D-5L, EQ-5D-Y*. 2021 30 April 2021]; Available from: https://euroqol.org/.
99. Yang, Y., *Personalized Therapy With Genetic Susceptibility to Advanced Non-small Cell Lung Cancer.* Chest, 2016. **149**(4): p. A280.

100. Schuette, W., et al., *CHAMP: a phase II study of panitumumab with pemetrexed and cisplatin versus pemetrexed and cisplatin in the treatment of patients with advanced-stage primary nonsquamous non–small-cell lung cancer with particular regard to the KRAS status.* Clinical lung cancer, 2015. **16**(6): p. 447-456.
101. Nafees, B., et al., *Health state utilities for non small cell lung cancer.* Health and quality of life outcomes, 2008. **6**(1): p. 1-15.
102. Doyle, S., A. Lloyd, and M. Walker, *Health state utility scores in advanced non-small cell lung cancer.* Lung Cancer, 2008. **62**(3): p. 374-380.
103. Martí, S.G., et al., *A cost-effectiveness analysis of a 10-valent pneumococcal conjugate vaccine in children in six Latin American countries.* Cost Effectiveness and Resource Allocation, 2013. **11**(1): p. 1-17.
104. Lewis, G., et al., *Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom.* Journal of International Medical Research, 2010. **38**(1): p. 9-21.
105. Northern Cancer Alliance. *Lung Chemotherapy protocols and prescriptions*. 2021 27 April 2021]; Available from: https://northerncanceralliance.nhs.uk/advisory_group/lung-expert-advisory-group/lung-chemotherapy-protocols-and-prescriptions/.
106. NHS. *NHS - Reference Costs 2018/2019*. 2021 15 April 2021]; Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819.
107. Curtis L., *Unit Costs of Health and Social Care 2019*. 2018, University of Kent, 176 pp. ISBN 978-1-911353-10-2.
108. NICE. *Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy [TA374]*. 2015 28 April 2021]; Available from: https://www.nice.org.uk/guidance/ta374.
109. NICE. *Atezolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy [TA525]*. 2018 28 April 2021]; Available from: https://www.nice.org.uk/guidance/ta525.
110. NICE. *Ipilimumab for previously untreated advanced (unresectable or metastatic) melanoma [TA319]*. 2014 28 April 2021]; Available from: https://www.nice.org.uk/guidance/ta319.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Clarification questions

July 2021

File name	Version	Contains confidential information	Date
Sotorasib clarification letter	Final_AIC_BIC	Yes	06/08/2021

Important note regarding updated data cut for CodeBreaK100

Please note: An addendum is provided with this response that updates efficacy analyses and modelling results in line with the March 2021 data cut of CodeBreak100. For completeness, all model scenarios and results presented in response to questions also use this updated data cut. An up-to-date model with updated data inputs is also provided with this response.

Section A: Clarification on effectiveness data

Literature searches

A1. Priority question. The Evidence Review Group (ERG) is concerned that studies might have been missed by using EMTREE terms in MEDLINE for the randomised controlled trial (RCT) facet as these terms are not automatically mapping across.

Please see below for searches in MEDLINE using EMTREE terms and also the correct MeSH (medical subject headings) terms:

# ▲	Searches	Results
1	crossover procedure/	0
2	Cross-Over Studies/	50741
3	double-blind procedure/	0
4	Double-Blind Method/	165568
5	single-blind procedure/	0
6	Single-Blind Method/	30452

a. **Please justify not modifying searches in MEDLINE to account for differences in thesaurus headings between Embase and MEDLINE.**

Amgen response: The broad "randomised trials" EMTREE term does map to MEDLINE, which in our experience covers the majority of the components of most RCT filters. In addition, we included the free text terms below, which ensured that all possible synonyms would be covered:

(random$ or factorial$ or crossover$ or cross over$ or cross-over$ or placebo$ or (doubl$ adj blind$) or (singl$ adj blind$) or assign$ or allocat$ or volunteer$).ti,ab.

In addition, extensive hand searches of conference abstracts, trial registries as well as reference checking of included studies and SLRs were used as an additional measure to ensure no studies were missed. We therefore believe our searches were comprehensive and would be at low risk of missing relevant trials that could inform the evidence base for sotorasib or the comparators.

b. Please re-run all searches using the correct MeSH terms and screen the results for relevant references.

Amgen response: As requested, we have re-run the searches by replacing EMTREE terms with suggested MeSH terms. This resulted in additional 13 studies being identified, none of which satisfy the inclusion criteria for the clinical RCT SLR. The screening of newly identified references was performed in a double-blind manner in line with SLR methods. The embedded word document below includes the details of the search and these 13 non-relevant studies. In conclusion, our searches presented in the CS were sufficiently comprehensive to capture all relevant RCTs.

New KRASm RCT SLR citations to scree

A2. Please provide details on the RCT filter applied in clinical effectiveness searches, e.g. whether this is a validated filter.

Amgen response: As described in Appendix D, this SLR was built upon a recently published SLR by Schulz et al. (2019), which was conducted in the pre-treated NSCLC population. To be consistent with the approach we have not modified the search terms used in that publication.

We acknowledge some possible issues including sensitivity of the search reported in Schulz et al. (2019) and the fact that published or validated search filters had not been used by authors. However, the search is logically laid out and uses syntax including MeSH, EMTREE and free text terms that should provide a sufficient degree of sensitivity to identify all relevant studies. As noted in response to Question A1 and A2, the approach we adopted appears to have been comprehensive.

A3. Please explain the use of "[Full texts only]" in line 15 of the June 2020 Embase clinical effectiveness searches (page 4).

Amgen response: The rationale for this was to distinguish between two groups of publications: Full texts and conference abstracts. Line 16 applies "2017 – current" time horizon to conference abstracts, whereas Line 17 applies "2015 – current" time horizon to full papers.

A4. Regarding the searches of the Cochrane Library:

a. Please justify the use of an RCT filter in CENTRAL and CDSR (Cochrane Database of Systematic Reviews) databases in the Cochrane Library as both of these are study design specific resources and the use of a filter may be overly restrictive.

b. Please justify the use of an animal filter as CENTRAL has been pre-screened to remove all animal studies and the use of an animals filter may have removed relevant records.

c. Please justify the use of an animal filter as CDSR does not include systematic reviews in animal research so the use of this filter is redundant.

d. Please re-run all of these searches and screen the results for relevant references.

Amgen response: While these databases are study design specific, the additional use of relevant study design filters to identify either RCTs or systematic reviews of RCTs in the population of interest does not (in our experience) significantly increase the risk of relevant studies being excluded. Similarly, in our experience, retaining the animal filter in the search string applied to these databases is highly unlikely to result in the removal of relevant RCTs. It should be noted that the databases we searched included, but were not restricted to, CENTRAL and CDSR; there is a high degree of overlap between these data bases and the other electronic databases (MEDLINE and EMBASE) we searched. As explained in response to Questions A1 and A2, our searches of these other databases were comprehensive, and we also conducted additional hand searching as part of the review. In conclusion, the use of the RCT and animal filter in the searches of the Cochrane library are highly unlikely to have

resulted in our collective searches missing relevant RCT data that could inform the evidence base for sotorasib or the comparators.

A5. Please provide search terms and hits per resource for conference proceedings and trial registries undertaken for clinical effectiveness.

Amgen response: Please see the embedded word document below for this information:

KRASm NSCLC_RCT SLR-HandSearchingTr

A6. Please justify the use of an English language limit for clinical effectiveness searches in Embase, MEDLINE and the Cochrane Library for single-arm trials conducted in patients with KRAS (Kirsten RAt Sarcoma virus) mutated advanced or metastatic non-small cell lung cancer (NSCLC).

Amgen response: Whilst we acknowledge a small risk of bias, restriction to English language is generally acceptable for pragmatic reasons given that most high-quality studies are generally published in English.

Decision problem

A7. Priority question. According to Tables 1 (The decision problem) and4 (Clinical effectiveness evidence: CodeBreaK100) of the company submission (CS), the population addressed in the CS appears to be narrower than that defined in the National Institute for Health and Care Excellence (NICE)scope.

a. **Please confirm that the population of the decision problem in the CS is narrower than in the NICE scope.**

Amgen response: The population addressed by the submission reflects the anticipated licensed indication, which is *as monotherapy for the treatment of adult patients with KRAS G12C mutated locally advanced or metastatic non-small cell lung cancer (NSCLC), who have progressed on, or are intolerant to, platinum-based*

chemotherapy and/or anti PD-1/PD-L1 immunotherapy (see table 1 and section B.1.2 of the CS – the wording above is the latest working version of the anticipated licensed indication as per the latest draft SmPC provided with this response). This is aligned with and reflects the population of NSCLC patients enrolled in the CodebreaK100 trial (see section B.2.3.2 of the CS and Appendix D, section D.1.2).

The scope for this appraisal was developed by NICE before the anticipated licensed indication had been agreed with the MHRA. NICE appraises medicines within their licensed indication, and as we have not proposed any subgroups of the licensed population for consideration in our submission, Amgen believes that the population of patients meeting the full anticipated licensed indication for sotorasib is the relevant population for the decision problem.

b. **Please confirm that the population included in CodeBreaK100 is narrower than in the NICE scope.**

Amgen response: As above, the full anticipated licensed indication reflects the relevant population for the decision problem. The population enrolled in CodeBreak100 reflects the full anticipated licensed indication and is, therefore, aligned with the relevant population for the decision problem.

c. **Please discuss whether it would be expected that patients with Eastern Cooperative Oncology Group (ECOG) statuses above 1 would be offered sotorasib.**

Amgen response: As noted in our discussion of the generalisability and relevance of the clinical evidence base, section B.2.13.4.1 of the CS, five UK clinical experts at an Amgen Advisory board considered that the population of patients enrolled in the CodeBreaK100 trial was reflective of patients in UK clinical practice who would meet the anticipated licensed indication. Although patients with ECOG PS 2 were not enrolled in the phase 2 part of the trial (as is usual in NSCLC trials), the UK clinical experts considered that these patients could be good candidates for treatment with sotorasib based on its efficacy and favourable adverse event and tolerability profile and its easier administration relative to the current non-targeted, intravenous cytotoxic standard of care therapy (docetaxel monotherapy or nintedanib plus docetaxel). Patients with ECOG PS 2 meeting the licensed indication may therefore have specific unmet needs that could be met by sotorasib. Provided the licensed

indication does not preclude its use, we believe that the exclusion of patients with ECOG PS 2 from the CodeBreaK100 trial should not preclude the use of sotorasib within its licensed indication in such patients in clinical practice. Sotorasib should be an option available to clinicians for use in patients with ECOG PS 2 when clinically relevant. See also the response to Question A20b.

d. **Please note question A12 (regarding the conditional licensing approval).**

Amgen response: we believe this may relate to question A13, rather than A12. Please see the response to question A13.

A8. Priority question. Regarding the comparators listed in the Table 1:

a. **Please provide further justification for the choice of comparators, e.g. while the ERG recognises that the primary comparator listed in theCS (docetaxel monotherapy) is in line with the NICE scope, it noted that the CS quoted the NICE lung cancer pathway and international guidelines regarding the increasing role of combination immunotherapy and chemotherapy.**

Amgen response: As detailed in section B.1.3.2 and B.1.3.3 of the submission, the role of combination immunotherapy and (platinum based) chemotherapy has increased in recent years, driven by the increasing evidence of the benefits of their early use in the treatment pathway. In line with current NICE and international treatment guidelines, five UK clinical experts at a recent advisory board confirmed that an increasing majority of patients in the UK with NSCLC without current actionable mutations can receive anti-PD-1/PD-L1 immunotherapy alone or in combination with platinum-based chemotherapy (but the majority will have been pre-treated with both) as first-line therapy; few patients receive immunotherapy as second-line therapy and re-challenge with immunotherapy is not routine practice in the UK. Given this early use of immunotherapy, usually as part of combination chemo-immunotherapy, immunotherapy is not a relevant comparator for sotorasib based on its anticipated licensed indication (see response to Question A7a).

Given the evidence of benefits and increasing use of combination immunotherapy and (platinum-based) chemotherapy, immunotherapy alone is increasingly unlikely to be a first-line option in patients who could tolerate platinum chemotherapy in

combination with immunotherapy. The use of platinum-based chemotherapy as a second line therapy following first-line immunotherapy is therefore decreasing. Given this, the alternatives available in the current treatment pathway following initial therapy with immunotherapy and/or platinum-based chemotherapy are docetaxel monotherapy or (for some patients with adenocarcinoma and a significantly smaller population) nintedanib in combination with docetaxel. The five UK clinical experts at the recent advisory board confirmed that these would be the relevant comparators for sotorasib in current clinical practice, and these are therefore the comparators included in our submission.

b. **Are results for other comparators, other than those listed as primary and secondary, available? If so, please provide these and include them as options in the model. In particular, for patients who have had first-line immunotherapy as monotherapy please include platinum doublet chemotherapy as a comparator (pemetrexed-based for nonsquamous disease and carboplatin plus gemcitabine or carboplatin plus paclitaxel for squamous disease).**

Amgen response: As noted in the CS and above in response to Question A7a, clinical experts have confirmed that docetaxel monotherapy or nintedanib plus docetaxel are the relevant comparators for sotorasib; other therapies are, therefore, not considered to be relevant comparators for sotorasib.

We are not aware of specific clinical trial evidence to support the use of platinum doublet chemotherapy as a second-line treatment following first line immunotherapy (i.e. in the relevant mutation group). In our submission we provided a supplementary propensity score weighting indirect comparison of sotorasib against a basket of chemotherapy regimens included in the Flatiron database (see section B.2.9.3.1 of the CS and Appendix D, section D.1.6). This was provided as a pragmatic means to validate the primary, trial-based MAIC analysis of sotorasib versus docetaxel monotherapy using another data source to represent docetaxel. However, the most common regimen amongst the basket of chemotherapy regimens in the Flatiron dataset was platinum-based chemotherapy (see Appendix Table 12 in Appendix D, section D.1.6.3). This analysis therefore also provides a pragmatic reflection of the likely relative treatment effects of sotorasib versus platinum-based chemotherapy.

The results of this supplementary comparative analysis are provided in section B.2.9.4.1, Table 15 of the CS. A scenario analysis using these data in the economic model are provided in section B.3.7.3.1, Table 46 of the CS.

Systematic literature review

A9. Priority question. Please provide the missing details on various aspects of the systematic literature review, including the number of reviewers involved in each step of the literature screening and how consensus was reached.

Amgen response: Appendix D and G detail the process of the SLRs, including the number of reviewers and the manner in which discrepancies were resolved. To clarify:

RCTs: Two reviewers independently screened titles and abstracts (Page 15, Appendices, Appendix D). Full papers were examined by two researchers (second pass), and final inclusion and exclusion of citations was verified by the project lead. Disputes as to eligibility were referred to a third party (strategic advisor).

SATs: Two reviewers independently screened the titles and abstracts (see page 17, Appendices, Appendix D). Discrepancies regarding study inclusion were resolved by a third reviewer, and relevant study information was noted in a screening sheet. These studies were then reviewed as full text in a double-blind manner by 2 researchers, and discrepancies were resolved by a third person. (Please note the heading of this section of Appendix D, page 17 should read **D1.1.2.2 Study eligibility criteria and selection for the SLR of single-arm trials)**

Economic burden SLR (covering costs/healthcare resource use, HSUV, and economic evaluation evidence): Titles & abstracts screened in a double-blind manner by two researchers (see page 111, Appendices, Appendix G). Independent dual-review of the full-text papers was conducted. Any uncertainties regarding the inclusion of studies were resolved by a senior reviewer.

A10. Priority question. According to Appendix Table 1, a number of restrictions were applied, e.g. to treatments licensed in the United States or the European Union, by date, by language, and regarding study types.

a. **Please justify the use of these restrictions as potentially relevant studies might have been missed.**

Amgen response: The Clinical RCT SLR was built upon and updated the recently published SLR by Schulz et al. (2019); therefore, eligibility criteria of the original review were mirrored in the update we conducted. These included restriction to 2L+ NSCLC treatments licensed in the United States or European Union at the time the search was conducted. Although there were no therapies specifically licensed for the treatment of KRAS G12C mutated NSCLC (because sotorasib is the first such therapy to progress to regulatory submission), NSCLC patients harbouring KRAS mutations are still treated with therapies that are licensed for the treatment of NSCLC (e.g. in the UK patients anticipated to be eligible for sotorasib are currently treated with docetaxel monotherapy, or nintedanib plus docetaxel, and both regimens are licensed in Europe for use in NSCLC following prior chemotherapy). The restriction to 2L+ therapies that are licensed therapies in Europe or the US is therefore entirely appropriate to the decision problem in this appraisal of sotorasib. As unlicensed therapies, or therapies that are not licensed for use in Europe, would not be relevant to the decision problem in this appraisal of sotorasib, restricting the searches to 2L+ therapies that are licensed in Europe or the US would not have missed any relevant studies.

The restriction to phase II to IV trials of 2L+ NSCLC treatments excludes phase I studies, which are general small, dose finding studies that provide very limited efficacy and safety data. Phase II trials are generally the minimum requirement for initial licensing of oncology therapies and so excluding phase I trials would not be anticipated to result in relevant studies being missed from our searches.

The restriction to English language publications is addressed in our response to Question A6. With respect to the time frame of the searches, studies published prior to 2017 were identified in the previous systematic literature review by Schulz et al. (2019) and were automatically included in the current review.

Collectively, there is little, if any, risk that the restrictions employed in our searches would have missed any relevant evidence that could change the evidence base and conclusions on the effects of sotorasib or the comparators.

b. **Please elaborate on the exclusion of numerous phase II trials listed in the Excel sheet linked in section D1.1.3.1 of the CS.**

Amgen response: please see the attachment below that provides further details on the citations excluded at full-text review as per "Study design" criteria: the exclusion code was updated for 30 records; no missed relevant records were identified as part of the exercise.

Citations excluded at FTR as wrong Study d

c. **Please discuss the implications of excluding non-randomised trials as CodeBreaK100 was a non-randomised study.**

Amgen response: We provided additional searches of single-arm trials, which are by definition non-randomised studies in our original submission (see Appendix .D, section D.1.1.3.2). We did not include searches of comparative trials that were not randomised, but given that few comparative studies are likely to be non-randomised, and given the context of this disease and the comparisons that need to be made with the single-arm codebreak100 trial, this is unlikely to have resulted in significant missed data – the relevant comparators of docetaxel monotherapy and nintedanib plus docetaxel are represented using the most robust data from their key randomised controlled trials.

d. **Please provide an overview of studies excluded based on these criteria.**

Amgen response: The rationale for exclusion of studies from the SLR of RCTs and Single arm trials is provided in the spreadsheets embedded in Appendix D (section D.1.1.3.1 for RCTs and D.1.1.3.2 for SATs). Given that these studies were excluded as irrelevant based on the clear inclusion and exclusion criteria employed in the SLRs, we are unclear what further information is required, or the merit of conducting a detailed overview of these irrelevant studies.

A11. Section D1.4.2 of the CS presents the quality assessment of SELECT-1 and LUME Lung 1. According to the text, the *"quality assessment of these RCTs was conducted using the seven-criteria checklist provided in section 2.5 of the NICE single technology appraisal (STA) user guide"*. According to the footer of Appendix Table 8, the items were "*adapted from Systematic reviews: CRD's* [Centre for Reviews and Dissemination] *guidance*".

a. Please amend the Table to include the last criterion listed as *"minimum criteria for assessment of risk of bias and generalisability in parallel group RCTs"* in the NICE STA user guide, namely *"consider whether the authors of the study publication declared any conflicts of interest"*.

b. Please elaborate how the Table was adapted from the CRD guidance.

Amgen response: Quality assessment of the included RCTs was conducted using both the seven-criteria checklist provided in section 2.5 of the NICE single technology appraisal (STA) user guide and version 2 of the Cochrane Collaboration's tool for assessing risk of bias in RCTs (in line with approach undertaken by Schulz et al. 2019). Appendix Table 8. contains the minimum criteria for assessment of risk of bias in RCTs, which is referenced in "Systematic Reviews: CRD's guidance for undertaking reviews in health care" (Box 1.5, page 37) and is in line with NICE STA user guide, apart from missing question about conflict of interest. We have added this question and the response to the Table below.

	SELECT- 1 [3]	**LUME-Lung 1 [25]**
Was randomisation carried out appropriately?	Y	Y
Was the concealment of treatment allocation adequate?	Y	Y
Were the groups similar at the outset of the study in terms of prognostic factors?	Y	Y
Were the care providers, participants and outcome assessors blind to treatment allocation?	Y	Y
Were there any unexpected imbalances in drop-outs between groups?	N	N
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N	N
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Y	Y

Consider whether the authors of the study publication declared any conflicts of interest.	Y	Y

A12. Please comment why this reference (Blumenschein et al. 2015) was not identified during the systematic literature review: https://www.annalsofoncology.org/article/S0923-7534(19)31514-5/fulltext.

Amgen response: This study was identified in our SLR – see Appendix D1.1.3.1 which states: *Regarding the relevant comparators for sotorasib in this submission, 3 of these 5 studies included docetaxel monotherapy as an intervention or comparator arm (**NCT01362296 (Blumenschein et al 2015) (7);** SELECT-1 (Janne et al 2017) (10); and, TAILOR (Rulli et al 2015 (12)). Of these, SELECT-1 was by far the largest and provided the most robust PFS and OS data in the subgroup of NSCLC patients harbouring the KRAS p.G12C mutation, and sufficient data on the baseline characteristics of enrolled patients to allow its consideration as a data source in indirect comparative analyses with sotorasib; the other 2 studies were much smaller and were more limited in their data.*

Intervention

A13. According to section B.1.1 CS, *"a conditional licensing approval via the Project Orbis regulatory route in the UK is anticipated* [redacted] *based on data from a phase 2 single-arm study (CodeBreaK100), with confirmatory results from a comparative phase 3 trial (CodeBreaK200) anticipated within the next 2 years"*.

a. Please provide the conditional licensing approval or give an update on the status of the approval.

Amgen response: Amgen estimate that a conditional GB marketing authorisation could be issued by MHRA by mid-August 2021, but this has not yet been received at the time of writing of this response. Amgen will keep NICE updated on any regulatory developments.

b. Please confirm whether the company intend to submit interim data from CodebreaK200 if it is available in January 2022?

Amgen response: At this stage we are not clear what interim data may or may not become available (and for which outcomes) in January 2022 and so are not in a position to firmly commit to providing that interim data.

Sotorasib is anticipated to receive a conditional license imminently based on the phase 2 data from the CodeBreaK100 trial. This single-arm trial was accepted in principle by the regulatory authority as sufficient for conditional licensing, balancing the need for sufficiently robust evidence generation against the need to ensure timely availability of sotorasib for patients with urgent, profound unmet clinical needs. As confirmatory data from the phase 3 CodeBreaK200 trial are anticipated to become available within the next 2 years, sotorasib is a potential candidate for the Cancer Drugs Fund (CDF), irrespective of the potential availability of interim data from that trial in January 2022. Given the urgent, profound unmet clinical needs of patients eligible for sotorasib, the appraisal of sotorasib should be concluded as soon as possible.

A14. Regarding the oral administration of sotorasib.

a. Please provide additional information regarding how the drug is meant to be taken, i.e. with food, with water, in the morning, 8 tablets all at once etc.

Amgen response: Table 2 of the CS provides a summary description of sotorasib, and states: *Sotorasib is administered orally at a dose of 960mg (given as 8 x 120mg tablets) once daily until disease progression or unacceptable toxicity.*

The draft SmPC, provided in the reference pack and as an updated version alongside this response document, states these should be taken at the same time each day, with or without food. The tablets should normally be swallowed whole, unless the patient has difficulty swallowing solids, in which case they can be dispersed in non-carbonated, room-temperature water without crushing. Further information is contained in the draft SmPC.

b. Please discuss any implications of the high proportion of participants experiencing diarrhoea (Appendix Table 22) as a treatment-related adverse

event. If available, please provide results separately, i.e. for participants experiencing/ not experiencing that side effect, respectively.

Amgen response: Analyses by the incidence of diarrhoea were not pre-specified in the CodeBreaK100 trial protocol and therefore are not available. We do not believe there are any meaningful implications of the incidence of diarrhoea observed in the trial for the following important reasons:

1) Appendix Table 22 refers to treatment-*emergent* adverse events of any severity, and not treatment-*related* adverse events as suggested by the ERG. Treatment-*related* adverse events are presented in section B.2.10.3, Table 18 of the CS. This shows that the incidence of diarrhoea of any severity was less common as a treatment-*related* adverse event (31.0%) than when reported as a treatment-emergent adverse event (49.2%), and treatment-*related* diarrhoea of severity Grade 3+ occurred in only 4% of NSCLC patients taking sotorasib in the trial. The vast majority of episodes of diarrhoea observed were therefore not severe. Of note, discontinuations of sotorasib due to treatment-*related* adverse events of any cause was low at 7.1% (see Table 17 of the CS).

2) Moreover, the efficacy analyses in the CodeBreaK100 trial were conducted in the full analysis set (for response-based endpoints) and the safety analysis set (for overall survival and safety analyses), both of which include participants who received at least 1 dose of sotorasib and so include participants who experienced diarrhoea as a treatment-related adverse event. The efficacy analyses for sotorasib therefore already fully capture any *theoretical* impact of diarrhoea on the efficacy of sotorasib.

Nevertheless, Amgen has consulted statisticians concerning the feasibility and meaningfulness of performing a subgroup analysis on a post-baseline event (i.e. that can occur at any time around events of interest) and was strongly advised against this. A number of potential biases would reduce any causal interpretation of such an analysis:

- AEs are subjective and subject reported and hence would introduce other forms of bias with a subgroup analysis

- Immortal time bias, meaning that subjects who experienced diarrhoea will have had to survived long enough to experience the AE (i.e. those experiencing the event will tend to have had better outcomes)
- Many of these AEs will have occurred after any events related to efficacy
    - Subjects experience diarrhoea at different times during the study - for example, median time to first diarrhoea was 5.5 months.

Given the multiple reasons above, we believe it would be inappropriate to conduct such an analysis.

A15. Given the close relationship between smoking, alcohol consumption, and the number of patients with liver metastases, please discuss the impact of treatment-related side effects involving the liver.

Amgen response: After receipt of this question Amgen consulted two clinicians for clarity concerning this question. Whilst they could not be certain of the clinical background to the question, they stated that patients who have elevated LFTs at baseline would likely have been excluded from the trial by the "adequate hepatic laboratory assessments" criteria. They did not believe that the stated factors would particularly confound the overall results.

Clinical evidence

Missing information and documents

A16. Priority question. In several places (relating to both, clinical as well as cost effectiveness), the CS refers to clinical expert opinion provided at a meeting of "UK advisory board February 2021".

a. **Please report on the methods used to gather the opinions of clinical experts, as well as the results of this process, and refer to this throughout the provided documentation.**

Amgen response: the UK advisory board was held 11th February 2021. The aims of the meeting were to gain the views and advice of clinical experts and health

economics experts on the appropriate demonstration of clinical and cost effectiveness of sotorasib. The meeting was conducted online.

Clinical experts:

Five UK consultant medical oncologists with expertise in the management of patients with NSCLC were invited and attended the meeting. They were drawn from five different hospitals representing 4 geographical regions in England. Following a presentation of the available clinical data from CodeBreaK100, the views of the clinical experts were sought on the efficacy and safety data available from the CodeBreaK100 trial of sotorasib, its likely position in the clinical pathway, the relevant clinical comparators for sotorasib based on current clinical practice, and the suitability of data sources and methods to provide indirect comparative evidence for sotorasib versus those comparators. The discussion was chaired by one of the clinicians, and all actively participated and provided their individual views based on their experience in current practice. Following the advisory board, all clinicians were provided with a copy of the meeting notes and the opportunity to review and correct the recorded information as appropriate. No requests were received to correct the meeting notes, indicating that the recorded information was a true and fair reflection of the discussions.

As noted in the CS, the clinical experts confirmed:

- The population of CodeBreaK100 was reflective of patients anticipated to meet the licensed indication of sotorasib in UK clinical practice (section B.2.13.4.1)
- The current standard of care for patients anticipated to meet the licensed indication for sotorasib are docetaxel monotherapy or nintedanib plus docetaxel. Use of nintedanib plus docetaxel is highly variable across the country. These are therefore the relevant clinical and economic comparators for sotorasib (section B.1.3.3)
- The ORR and adverse event profile of sotorasib is far better than they would expect to see in patients treated with current standard of care (docetaxel monotherapy or nintedanib plus docetaxel) (section B.2.6.2.1 and B.2.10.3)

- The use of SELECT-1 and LUME Lung 1 trials as reasonable data sources for the comparators (section B.2.9.2.1)
- It would not be possible to match LUME Long 1 to CodeBreak100 via MAIC in a meaningful way (section B.2.9.3.2 and B.3.3.5)

Follow up questions were discussed separately with two of the five clinicians to validate outcomes of survival analyses, cost and resource utilisation and the approach to utility values for the economic model (see further details in response to Question B16b).

Health economics experts:

Two UK-based health economics experts were invited and participated in the meeting 11th February 2021. They provided views and advice on the methods of estimating relative treatment effects from the clinician-agreed data sources, and the economic modelling methods.

b. Please provide the reference to support this, e.g. a document with details of this meeting.

Amgen response: Outputs from the advisory board relevant to the current submission are reflected in the CS as noted above. Whilst meeting notes were recorded these are for internal use only in order to preserve the identity of participants and the commercial in confidence information that was discussed. We are therefore unable to provide the ERG with information beyond that which is already provided in the CS.

A17. Section B.2.9.2.1 refers to "large real-world evidence studies undertaken by Amgen".

Please provide further details on these studies, including methods and full references.

Amgen response: The CS cites reference 37 for the large real-world evidence studies undertaken by Amgen. Reference 37 is the report for the observational study undertaken using the Flatiron Health - Foundation Medicine Clinico-Genomic Database. This was provided in the reference pack as "Amgen Inc_Observational Research Study Report 20200132", which details the full methods and results.

For the avoidance of doubt, this study is the same as the study 20200132 referred to in section B.1.3.1.2 Table 3 in the CS. The references for the companion Flatiron study 20200097 and the further real-world evidence AACR project GENIE study 20180277, also referred to in Table 3, were provided in the reference pack as "Amgen Inc_ Observational Research Study Report 20200097_KRAS G12C" and "Amgen Inc. ACCR Project Genie_Observational Research Study Report_2020-02-11", respectively.

CodeBreaK100 study

A18. Priority question. Appendix E of the CS presents subgroup results for CodeBreaK100.

Please provide results separately for participants with and without adenocarcinoma, respectively.

Amgen response: The majority of patient enrolled in the CodeBreaK100 trial had adenocarcinoma (120/126 [95%]). We have conducted an analysis to meet the ERG's request; however, data in the 6 patients who did not have adenocarcinoma (including 1 patient with squamous histology) are limited and efficacy results should be interpreted with an appropriate degree of caution. Nonetheless, these subgroup data suggest a similar magnitude of effects in ORR, PFS and OS (see table below). It should be noted that, in contrast to nintedanib which is only licensed for use in patients with adenocarcinoma, the anticipated licensed indication for sotorasib does not restrict its use to this specific histology (see the draft SmPC for sotorasib provided in the reference pack, and the updated version accompanying this response document). Given the similar magnitude of effects observed in patients with or without adenocarcinoma, there is no compelling reason to restrict the use of sotorasib based on NSCLC histology. Sotorasib should be available for all patients meeting its anticipated licensed indication.

Efficacy in CodeBreaK100 by adenocarcinoma histology (15 March 2021 data cut, post hoc analysis)

	ORR	PFS				OS			
Adeno-carcinoma	Events/ Subjects (%) (95% CI)	Events/ Subjects	Median (95% CI) (Months)	6 Months KM Estimate (95% CI) (%)	12 Months KM Estimate (95% CI) (%)	Events/ Subjects	Median (95% CI) (Months)	6 Months KM Estimate (95% CI) (%)	12 Months KM Estimate (95% CI) (%)
Yes	44/118 (37.3) (28.6, 46.7)	82/118	6.8 (5.1, 8.2)	52.2 (42.3, 61.2)	26.9 (18.6, 36.0)	62/120	12.0 (10.0, NE)	74.2 (65.2, 81.2)	50.5 (40.9, 59.3)
No	2/6 (33.3) (4.3, 77.7)	5/6	6.2 (1.2, NE)	50.0 (11.1, 80.4)	33.3 (4.6, 67.6)	2/6	NE (6.6, NE)	100.0 (NE, NE)	66.7 (19.5, 90.4)

A19. Priority question. Please provide more details on the CodeBreaK100 study.

a. **Please provide definitions for terms such as "disease progression", "intolerance" etc. (see Table 4 of the CS).**

Amgen response:

Progressive disease is defined in the protocol as:

Target lesions - At least a relative 20% increase and an absolute increase of 5 mm in the sum of the diameters of target lesions, taking as reference the smallest sum on study, or the appearance of 1 or more new lesions.

Non-target lesions - Unequivocal progression of existing non-target lesions and/or appearance of one or more new lesions. To achieve "unequivocal progression" on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease such that, even in presence of stable disease or partial response in target disease, the overall tumour burden has increased sufficiently to merit discontinuation of therapy. A modest "increase" in the size of 1 or more non-target lesions is usually not sufficient to qualify for unequivocal progression status.

Intolerance here refers to intolerance to sotorasib that results in the cessation of sotorasib therapy.

Objective response rate is defined as the proportion of patients with a Best Overall Response of confirmed complete response (CR) or confirmed partial response (PR).

PFS is defined as interval from the start of treatment to disease progression or death due to any cause (whichever comes first).

Duration of response (DOR) is defined as time from first evidence of PR or CR to disease progression or death due to any cause. DOR will be calculated only for subjects who achieve a confirmed best overall response of PR or CR.

Disease control rate (DCR) is defined as the proportion of patients in whom the best overall response is determined as CR, PR or stable disease (SD) >= 5 weeks.

Time to response (TTR) is defined as time from the start of treatment until the first evidence of PR or CR. TTR will be calculated only for subjects who achieve a best overall response of confirmed PR or better.

CR, PR and SD are as defined by RECIST 1.1 criteria.

Please see the protocol for CodeBreaK100, provided in the reference pack with the CS, for any other definitions.

b. **Please provide details on the "blinded independent central review", mentioned in Table 4 of the CS, e.g. who was blinded independent of whom.**

Amgen response: BICR here refers to the assessment of response per RECIST 1.1 criteria by central review rather than by investigators. Whilst treatment assignment in this single-arm trial is known, the central reviewers were blind to patient identity and investigators.

c. **Please provide details on the "deepest average observed response" observed in phase 1 of CodeBreaK100, cf. section B.2.3.1.**

Amgen response: this refers to the criteria used to determine the dose of sotorasib to be used in the phase 2 portion of the trial (see the CSR for the September 2020 data cut, section 8.5.2, which notes that 960 mg orally once per day achieved the highest ORR and the deepest average observed response compared with the other doses examined during phase 1 of the study).This is the same dose that is anticipated to be licensed by the regulator, and as the licensed dose this is not

subject to adjustment outside of the criteria specified in the draft SmPC. See the draft SmPC (provided in the reference pack submitted with the CS) for permitted dose adjustment.

d. **According to section B.2.3.2, participants of CodeBreaK100, had *"adequate haematological, renal, hepatic, and coagulation laboratory assessments"*. Please provide further details.**

Amgen response: Please refer to Appendix D, section D1.2 for these details.

A20. Priority question. Regarding the generalisability of CodeBreaK100:

a. **Please discuss the generalisability of the results of CodeBreaK100 to United Kingdom (UK) clinical practice, given that there were no test centres in the UK.**

Amgen response: The generalisability and relevance of CodeBreaK100 and the indirect comparisons of sotorasib versus the relevant comparators to UK clinical practice are specifically discussed in section B.2.13.4 of the CS. This section discusses the generalisability of the trial and results in a PICOS format, including UK clinical expert opinion that supports our conclusion that there is no reason to doubt that the results of CodeBreaK100 are generalisable to the anticipated use of sotorasib in *KRAS p.G12C*-mutated NSCLC patients in UK clinical practice.

b. **Please also discuss whether the ECOG status of patients in CodeBreaK100 are reflective of patients in UK clinical practice.**

Amgen response: The ECOG PS of patients in CodeBreaK100 was 0-1, as dictated by the trial protocol. As noted in our response to Question A7c, CodeBreaK100 is not unusual amongst oncology trials in enrolling patients with PS 0-1 (validated by clinicians consulted on receipt of these questions). Most oncology trials in relapsed NSCLC exclude patients with PS 2 and above for pragmatic reasons – it allows the measurement of efficacy data over a meaningful period of time. Indeed, both the SELECT-1 trial (providing data for docetaxel monotherapy in our submission [Janne 2017 in the reference pack]) and the LUME Lung 1 trial (providing data for nintedanib plus docetaxel in our submission [Reck 2014 in our reference pack]) enrolled only patients with PS 0-1, and yet docetaxel monotherapy and nintedanib plus docetaxel have been confirmed by clinical experts as the current standard of care therapies

used in routine clinical practice in the UK in the patient population anticipated to receive sotorasib (see the response to Question A16 and section B.1.3.3 in the CS). Therefore, with respect to PS, the patients enrolled in the CodeBreaK100 trial are no less representative of patients in practice than patients enrolled in comparable NSCLC trials.

UK clinical experts agreed that the population of CodeBreaK100 was reflective of patients anticipated to meet the licensed indication of sotorasib in UK clinical practice (as discussed in section B.2.13.4.1 of the CS).

c. **According to sections B.2.5 and B.2.13.4.5 of the CS, *"notwithstanding the single-arm, non-randomised design, the CodeBreaK100 study is considered to be at a low risk of bias, with good external validity"*. Please justify this assertion regarding the good external validity.**

Amgen response: external validity here refers to the extent to which the results of the trial are valid and generalisable to the population and setting in which sotorasib will ultimately be used in clinical practice. Please see our response to Question A20a and the text in CS section B.2.13.4.1.

A21. Please clarify whether in Table 6 of document B of the CS all metastases are listed, i.e. whether patients only had metastases in the brain and liver, and no other locations.

Amgen response: Table 6 refers to key baseline characteristics, of which brain and liver metastases were included as key sites for metastases; however, the trial included patients with metastases at other sites (e.g., bone). Appendix E, Appendix Figures 17, 18 and 19 includes subgroup analyses by metastases of the bone and at any site.

A22. Please conduct a reanalysis of the safety analysis using all patients from CodeBreaK100, not just those with NSCLC.

Amgen response: These data were provided in our original submission in Appendix F, Section F1.2, Appendix Tables 23 and 24. These tables provide the safety and adverse events data across all 234 patients (NSCLC, colorectal cancer, other

tumours, and overall combined population) enrolled in the phase 2 CodeBreaK100 trial.

A23. Table 4 of the CS, presenting the clinical effectiveness evidence for CodeBreaK100, does not include time to treatment discontinuation as a reported outcome.

Please clarify and provide missing data, if needed.

Amgen response: Time to treatment discontinuation (TTD) was not a prespecified efficacy or safety endpoint in the CodeBreaK100 trial (see the CodeBreaK100 trial protocol provided in the reference pack submitted with the CS) but was estimated and used to inform the economic model. On this basis it was not included in Table 4 but was still an endpoint aligned with the outcome measures stipulated in the scope for the appraisal (Table 1 of the CS). See section B.3.3.6 of the CS. Exposure data, which are related to TTD are presented in B.2.10.1 of the CS.

Indirect treatment comparison

A24. Priority question: Section B.2.9.2.3 states that *"given thatPFS* [progression-free survival] *and OS* [overall survival] *outcomes are similar in the absence of targeted therapies, irrespective of KRAS status (…), the inability to match by specific KRAS status is unlikely to lead to biased estimates"*.

Section B.2.9.2.1 refers to Table 3 for showing that outcomes were similar irrespective of KRAS-mutant status. However, Table 3 does not distinguish between patients on targeted or non-targeted therapies, just line of treatment, and then only between non-p.G12C KRAS-mutated and all NSCLC patients (not KRAS mutated and non-KRAS mutated patients).

Please provide evidence to support the assertion that, *"in the absence of targeted therapies"*, PFS and OS outcomes are similar irrespective of KRAS status.

Amgen response: Table 3 of our submission includes data from patients with KRAS p.G12C mutated NSCLC, and KRAS mutated (but non-pG12C) and All NSCLC patients (Table 3 is replicated below for convenience). Given the OS and PFS for All

NSCLC patients are highly consistent to those for patients with KRAS mutations (and the KRAS mutation datasets are included in the All NSCLC dataset), it is reasonable to conclude that the results are consistent for those with and without KRAS mutations.

Regarding the issue of targeted vs non-targeted therapy, it should be noted that sotorasib is the first targeted therapy to have progressed through clinical development and to have been submitted for regulatory approval. The outcomes data in Table 3 therefore relate to outcomes with non-targeted therapies only. Our statement that "*in the absence of targeted therapies, PFS and OS outcomes are similar irrespective of KRAS status*" is therefore supported and justified.

Table 3 (Replicated from the CS)

	***KRAS p.G12C*-mutated NSCLC**		**KRAS-mutated (non-*p.G12C*)**	**All NSCLC**
	Study 20180277 AACR project GENIE	Study 20200097 Flatiron Health Foundation	Study 20200132 Flatiron Health Foundation	
Median (95% CI) OS (months)				
First line	14.9 (12.2, 24.3)	12.0 (9.6, 15.3)	12.2 (10.5, 14.4)	12.9 (11.9, 14.2)
Second line	10.1 (7.1, 16.9)	9.5 (8.1, 13.1)	9.6 (7.7, 12.4)	10.2 (9.5, 11.3)
Third line	6.5 (5.0, NE)	6.7 (5.9, 10.7)	6.6 (5.0, 9.0)	7.9 (6.8, 8.8)
Fourth line	3.0 (2.2, NE)	5.9 (4.3, 12.9)	5.5 (3.9, 8.6)	7.4 (6.4, 8.6)
Median (95% CI) rwPFS (months)				
First line	6.1 (4.4, 9.3)	5.0 (4.4, 5.8)	5.6 (5.4, 6.0)	5.6 (5.3, 5.8)
Second line	3.2 (2.1, 5.3)	4.0 (2.8, 5.3)	4.1 (3.7, 4.4)	4.0 (3.7, 4.4)
Third line	2.3 (1.4, 4.1)	3.1 (2.4, 4.3)	3.5 (3.2, 4.0)	3.5 (3.1, 3.9)
Fourth line	1.8 (1.4, 15.0)	2.6 (2.1, 4.7)	3.1 (2.7, 3.5)	3.0 (2.7, 3.4)

KRAS p.G12C = *KRAS* gene with a mutation resulting in a G12C amino acid substitution; NE = not evaluable; NSCLC = non-small cell lung cancer; OS = overall survival; rwPFS = real-world progression-free survival
Retrospective Studies 20200097 and 20200132 were conducted using the United States Flatiron Health - Foundation Medicine Clinico-Genomic Database in 743 patients with *KRAS p.G12C*-mutated advanced NSCLC and 7069 patients with advanced NSCLC (ie, regardless of *KRAS p.G12C* mutation), respectively. Retrospective Study 20180277 was conducted using the American Association for Cancer Research Project Genomics Evidence Neoplasia Information Exchange database in 416 patients with *KRAS p.G12C*-mutated advanced NSCLC.[5, 36, 37]

A25. Priority question. Table 10 of document B of the CS shows selected baseline characteristics present in at least one of the included studies. Country

of origin, socio-economic status, comorbidities, year of recruitment and number/severity of metastases are not included in this table.

a. **Please clarify whether these variables are not included as they were not considered potential confounders (in which case, please justify why they would not be potential confounders), or whether they were not measured in any included study.**

Amgen response: Based on physician's assessment (see response to A27) of the importance of each of sociodemographic characteristics (age, gender, smoking status, geographic region/ethnicity/race, body mass index/weight or history of alcohol abuse) there was no factor that a majority (at least 4 of the 6) of physicians found very important to consider. Therefore, these covariates were not included in the analyses.

b. **If the variables listed above were not included in the indirect comparisons because they were not measured in the studies (or not able to be included due to no overlap between studies, such as year of recruitment), please describe whether you would expect there to be differences between the studies, and how any differences could impact on the results of the cost effectiveness analyses.**

Amgen response: Given these were not considered important, the impact of excluding these covariates is likely to have no or minimal impact on the cost-effectiveness analyses results.

c. **Please describe and provide evidence whether the standard of care for NSCLC is likely to be equivalent between the studies, and therefore whether differences in care other than the indicated treatment may have contributed to any differences in the cost effectiveness estimates. In particular, please describe what effects different first line treatments might have on subsequent outcomes after starting the second/third/fourth line treatments.**

Amgen response: CodeBreaK100 subjects were more heavily pre-treated than SELECT-1 subjects, with SELECT-1 and LUME-Lung 1 patients receiving only 1 previous line of systemic anticancer therapy. Most patients in CodeBreaK100 had

received an anti-PD-(L)1 regimen in prior lines while no patient received it in SELECT-1. Prior therapies before enrolling patients across different studies are likely to be different because of the timings of these studies (i.e. enrolment in LUME-Lung 1 and SELECT-1 completed in Feb 2011 and January 2016 when immunotherapies were not licensed in the first line setting). This is a conservative limitation for the comparative effectiveness as a more heavily pre-treated population is generally associated with poorer clinical outcomes.

A26. Table 10 of document B of the CS shows differences in the smoking rate in LUME-Lung 1 study compared to CodeBreaK100 and SELECT-1 (64% versus 93% and 92%).

a. Please explain the differences in smoking rates.

Amgen response: The reported smoking rates for CodeBreaK100 and SELECT-1 (the primary analysis we provided in our submission) are highly consistent at 92-93%. For the LUME-Lung-1 study (used in the secondary comparison of sotorasib versus nintedanib plus docetaxel) the figure of 64% smokers presented in Table 10 of the CS appropriately refers to smoking rates in patients with adenocarcinoma; in the entire trial the rate was 75% (see Reck 2014 provided in the reference pack to the CS). We do not know the reasons for why the smoking history of patients enrolled in LUME Lung 1 was different to that in CodeBreaK100 and SELECT-1; however, it can be seen in the LUME Lung 1 trial results that progression-free survival (PFS) and overall survival (OS) were not significantly different between patients with or without a history of smoking (see Figure 4 in Reck 2014, which refers to the adenocarcinoma population). Given that smoking history had no significant impact on outcomes, it is therefore reasonable to adopt the overall PFS and OS data for the adenocarcinoma subgroup as reflective of PFS and OS with nintedanib plus docetaxel, and to use this in the indirect comparison with sotorasib. We note that, whist not significantly different, both PFS and OS were marginally numerically improved in patients who were never smokers. Therefore, given that the LUME Lung 1 trial had a higher proportion of never smokers compared with CodeBreaK100, the indirect comparison is likely to favour nintedanib plus docetaxel, and our estimates of relative treatment effects of sotorasib versus nintedanib plus docetaxel are likely to be conservative.

b. Kindly discuss whether it is expected that the nature of the NSCLCs in LUME-Lung 1 are likely to be equivalent to either CodeBreaK100 or SELECT-1, and what impact this might have on the cost effectiveness analyses.

Amgen response: In the absence of data for nintedanib plus docetaxel specifically in patients with KRAS p.G12C mutated NSCLC, the most appropriate source of data for nintedanib plus docetaxel was the LUME Lung 1 trial. This was confirmed as appropriate by UK clinical experts, as discussed in the CS and in our response to Question A16a.

In terms of KRAS mutant status, as targeted therapy for KRAS mutated NSCLC did not exist at the time of the LUME Lung 1 trial, and screening for KRAS mutant NSCLC was not routine practice, we have no means of knowing the KRAS mutant status of patients enrolled in the LUME Lung 1 trial. It was for this reason that it was not possible to match the LUME Lung 1 trial participants and CodeBreaK100 trial participants in an MAIC, as confirmed by clinical experts (see response to Question A16a and section B.2.9.3.2 and B.3.3.5 of the CS).

In terms of histology, the data for nintedanib+ docetaxel used on the indirect comparison is from patients with adenocarcinoma (in line with the licensed indication for nintedanib plus docetaxel, and as per its recommendation in NICE TA347). In CodeBreaK100, 120/126 (95%) of patients that had adenocarcinoma NSCLC, and based on the very limited data in the 6 patients with non-adenocarcinoma, results appear to be consistent with those in adenocarcinoma (see response to question A18). Therefore, there do not appear to be major differences in the histology of the NSCLCs in CodeBreaK100 and LUME Lung 1 that would significantly impact on the relative treatment effects and resulting cost effectiveness estimates.

In terms of smoking status, given the modelled cost effectiveness is driven by PFS and OS, and the indirect comparisons of sotorasib versus nintedanib plus docetaxel are likely to favour nintedanib plus docetaxel as described in response to Question A26a of this question, we do not anticipate that the differences in smoking status would impact on the nature of NSCLC in LUME-Lung 1 to bias the cost effectiveness analysis in favour of sotorasib; rather it is likely that the cost effectiveness analyses are conservative.

Collectively, based on the most relevant data sources available for sotorasib and nintedanib plus docetaxel, we do not have any evidence that the nature of the NSCLCs in LUME Lung 1 and CodeBreaK100 would differ to the extent that it would invalidate the indirect comparisons and cost effectiveness analyses.

A27. Please present full information regarding the choice of covariables for the matching-adjusted indirect comparison (MAIC), including:

1. The starting list of candidate prognostic variables
2. The pre-read documents, including the questionnaire, circulated to the physicians
3. Information regarding how the variables were selected as very important, somewhat important, or not important

Amgen response: The background reading document, questionnaire (inclusive of starting list of prognostic variables) will be sent with this response. Please note that this was a core global activity to cater to a variety of country HTA requirements. This includes background material and descriptions (e.g. patient profiles) that are relevant to the UK, and additional information that is relevant to other countries' HTA requirements.

Please note that the report summarising the physician responses cannot be shared because clinicians provided responses in confidence and we have no agreement to share their participation externally. Table 11 in the submission summarises the categorisation of the covariates based on the quantitative results of the physician assessment. In addition, the conclusion of the physicians' assessment report was as follows:

"*Interviews conducted with six medical oncologists covering 5 countries in North America and Europe, and with extensive experience of treating patients with advanced NSCLC and with good knowledge of the NSCLC research literature, have allowed to better understand what the key factors are when assessing/adjusting for the prognosis of advanced NSCLC patients.*

More precisely, there was a unanimous opinion from all six physicians that ECOG performance status was the most important factor to predict prognosis of advanced NSCLC patients. As a result, it was strongly recommended when performing comparative effectiveness analyses to ensure that 1) the populations being compared are defined similarly in terms of ECOG performance status at baseline and 2) adjustment based on ECOG score is performed to ensure the populations being compared are balanced. Other factors that were indicated by the physicians to be very important by a majority of physicians to assess the prognosis or response to treatment of patients included presence of brain metastases (ideally distinguishing between active and controlled brain metastases), disease stage at baseline (stage IIIb/c vs IV or IIIb-IVa vs IVc). A number of other factors were considered as being at least somewhat important for prognosis and should be considered when the information is available. Finally, age and gender, although not consistently considered as prognostic or predictive factors, were mentioned as key covariates to include in an adjusted comparative effectiveness analysis."

A28. Please justify why not matching on brain metastases between CodeBreaK100 and SELECT-1 is "unlikely to introduce significant bias", given brain metastases are a very important prognostic indicator.

Amgen response: The proportion of patients with brain metastases was higher in CodeBreaK100 (21%) than in LUME Lung-1 (8%). The proportion with brain metastases in SELECT-1 was not reported. However, all three trials excluded patients with *active* (or symptomatic) brain metastases. It is our understanding, and confirmed by clinical experts, that active brain metastases are a much more significant negative prognostic factor than "non-active" brain metastases, and therefore not matching on "non-active" brain metastases is unlikely to introduce significant bias.

As CodeBreaK100 enrolled a high proportion of patients with brain metastases, and somewhat higher than in patients recruited to LUME Lung 1, it is a reasonable assumption that SELECT-1 did not include a higher proportion of patients with non-active brain metastases than CodeBreaK100. Any negative influence on survival of the presence of brain metastases would therefore impact on the CodeBreak100 population to a greater extent than on the populations in LUME Lung 1 or SELECT-

1. Therefore, the results of the comparison of sotorasib (from CodeBreaK100) vs nintedanib plus docetaxel (from LUME Lung 1) or docetaxel monotherapy (from SELECT-1) would favour the comparators. On this collective basis, our inability to match for brain metastases between CodeBreaK100 and SELECT-1 is unlikely to introduce bias in favour of sotorasib and is more likely to be conservative.

A29. Please clarify whether there could potentially have been patients with brain metastases in SELECT-1, given the percentage of patients with brain metastases in Table 10 of Document B for SELECT-1 is listed as “NR” rather than “0%”.

Amgen response: As stated in the footnote to Table 10, the SELECT-1 trial publication did not report the percentage of patients with brain metastases. Clinical expert opinion suggests it is possible that patients in SELECT-1 had non-active (non-symptomatic) brain metastases; however, we have no basis on which to estimate that number, and “NR” was used appropriately in Table 10. Given that a high proportion of patients in CodeBreaK100 had non-active brain metastases (21% versus 8% in LUME Lung 1), it is reasonable to speculate that the proportion of patients with non-active brain metastases in SELECT-1 is likely to be lower than that seen in CodeBreaK100.

A30. Please clarify whether the participants listed in Table 10 of document B of the CS comprised of participants with confirmed adenocarcinoma. If not, please provide details for all studies of the type of cancers.

Amgen response: For CodeBreaK100, the data in Table 10 are derived from 120 patients with adenocarcinoma and 6 non-adenocarcinoma. It should be noted that there is no robust evidence of a difference in effects of sotorasib in patients with or without adenocarcinoma NSCLC (see response to Question A18), and the anticipated licensed indication for sotorasib does not specify it should be used only in adenocarcinoma (see the draft SmPC provided in the reference pack).

For SELECT-1, the trial publication does not specify the proportion of patients with adenocarcinoma; however, 95% of patients were reported to have non-squamous NSCLC (see Table 1 of Janne 2017, provided in our reference pack for the CS), and as described in section B.1.3.1.1 of the CS, adenocarcinoma is by far the most common type of non-squamous NSCLC. We can therefore be confident that a high proportion of patients in SELECT-1 were of adenocarcinoma histology. It should be

noted that, docetaxel is recommended as a treatment option in the NICE NSCLC pathway for any histology (see Figure 1, section B1.3.3 of the CS).

For LUME Lung 1, the data in Table 10 are derived from patients with adenocarcinoma only, in line with its licensed indication and recommendation by NICE in TA347).

In summary, the data in Table 10 appropriately reflects the populations currently receiving docetaxel monotherapy and nintedanib plus docetaxel as standard of care therapies in UK clinical practice, and the population anticipated to receive sotorasib in line with its licensed indication.

A31. Please justify, with evidence, this sentence in Section B.2.9.2.3: *"The data are considered adequate to reflect PFS and OS outcomes with sotorasib, docetaxel monotherapy and nintedanib plus docetaxel treatment following prior therapy in this population of patients."*

Amgen response: Section B.2.9.2.3 refers to the Conclusions on the feasibility of undertaking indirect treatment comparisons. The conclusion that *"The data are considered adequate to reflect PFS and OS outcomes with sotorasib, docetaxel monotherapy and nintedanib plus docetaxel treatment following prior therapy in this population of patients"* is based on the detailed discussion and consideration of the data sources, the compatibility of the trial designs and assessment of the similarity in patient eligibility criteria and the enrolled population characteristics, which supports this conclusion. Please refer to section B.2.9.1 and B.2.9.2 of the CS, and the response to Question A16 above, which demonstrates how five UK clinical experts confirmed that our approach to demonstrate the comparative efficacy of sotorasib was reasonable (which implicitly requires that these data are *adequate to reflect PFS and OS outcomes with sotorasib, docetaxel monotherapy and nintedanib plus docetaxel treatment following prior therapy in this population of patients*).

Appendices

A32. Appendix F of the CS reports adverse events (AEs) of CodeBreaK100 while appendix M (Model inputs) also reports AE rates for drugs not studied as part of that study.

Please amend the Tables in appendix F with relevant results for SELECT-1 and LUME Lung 1.

Amgen response: Appendix F appropriately refers to adverse events data for sotorasib. Appendix M appropriately refers to model inputs, including adverse event rates for the appropriate comparators. A table of the adverse events and incidence rates used in the model for sotorasib and the comparators is already provided in section B.3.3.7, Table 32 of the CS. Please refer to that section of the CS for information.

A33. Appendix Table 23 shows high rates of serious AEs.

Please provide a detailed breakdown of these serious AEs.

Amgen response: Appendix Table 23 reports treatment-*emergent* adverse events and treatment-*related* adverse events. Serious treatment-*emergent* adverse events were reported in 50% of NSCLC patients in the CodeBreaK100 trial; however, serious treatment-*related* adverse events occurred in a far lower 7.9% of NSCLC patients.

The safety data from the 01 December 2020 data cut, provided in the reference pack as the file called 'Amgen Inc_ CodeBreaK100 Safety data (01 December 2021 data cut)' (note the reference to 01 December 2021 data cut in this file name is simply a typographical error) includes a table of Treatment-related adverse events by system organ class and grade (Table 14b-6.8.2). Although this does not categorise the adverse events as "serious" or not, we have extracted those judged to be of Grade 3 or greater severity and presented these in the Table below.

Of note, there were few Grade 4 treatment-related adverse events in the trial and no fatal adverse events. Also of note, UK clinical experts considered the adverse event profile of sotorasib to be very manageable and far better than that seen with the

current standard of care docetaxel monotherapy or nintedanib plus docetaxel (see B.2.10.3 of the CS and the response to Question A16).

Table of Grade 3+ treatment-related adverse events in CodeBreaK100

Preferred Term Worst Grade	Phase 2 NSCLC 960 mg QD Fasted (N = 126) n (%)	Phase 2 CRC 960 mg QD Fasted (N = 62) n (%)	Phase 2 Other Tumors 960 mg QD Fasted (N = 46) n (%)	Phase 2 Total (N = 234) n (%)
Any preferred term				
Grade 3	25 (19.8)	5 (8.1)	5 (10.9)	35 (15.0)
Grade 4	1 (0.8)	1 (1.6)	0 (0.0)	2 (0.9)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Lymphopenia				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Neutropenia				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Diarrhoea				
Grade 3	5 (4.0)	2 (3.2)	2 (4.3)	9 (3.8)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Abdominal pain				
Grade 3	(0.0)	0 (0.0)	1 (2.2)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Gastrointestinal haemorrhage				
Grade 3	0 (0.0)	0 (0.0)	1 (2.2)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatigue				
Grade 3	0 (0.0)	1 (1.6)	2 (4.3)	3 (1.3)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Drug-induced liver injury				
Grade 3	2 (1.6)	0 (0.0)	0 (0.0)	2 (0.9)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hepatic function abnormal				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hepatotoxicity				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Drug hypersensitivity				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Cellulitis				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Alanine aminotransferase increased				
Grade 3	8 (6.3)	1 (1.6)	0 (0.0)	9 (3.8)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Aspartate aminotransferase increased				
Grade 3	7 (5.6)	1 (1.6)	0 (0.0)	8 (3.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Blood alkaline phosphatase increased				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Gamma-glutamyltransferase increased				
Grade 3	2 (1.6)	0 (0.0)	0 (0.0)	2 (0.9)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Lymphocyte count decreased				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Lipase increased				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Liver function test increased				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Neutrophil count decreased				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Transaminases abnormal				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Blood creatine phosphokinase increased				
Grade 3	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Grade 4	0 (0.0)	1 (1.6)	0 (0.0)	1 (0.4)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hypokalaemia				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Back pain				
Grade 3	0 (0.0)	1 (1.6)	0 (0.0)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Dyspnoea				
Grade 3	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Grade 4	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Pneumonitis				
Grade 3	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Grade 4	1 (0.8)	0 (0.0)	0 (0.0)	1 (0.4)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Pleural effusion				
Grade 3	0 (0.0)	0 (0.0)	1 (2.2)	1 (0.4)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fatal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Section B: Clarification on cost effectiveness data

Literature searches

B1. Please confirm which filters have been applied for cost effectiveness searches.

Amgen response: Please refer the search strategy described in Appendix G (page 89-110, Appendices)

B2. The company lists Database of Abstracts of Reviews of Effects (DARE), Cochrane Methodology Register (CMR), National Health Service Economic Evaluations Database (NHS EED), Health Technology Assessment Database (HTA) and American College of Physicians (ACP) journal club as having been searched in the Cochrane Library, but these resources were no longer available in February 2020 via the Cochrane Library which is when the earliest Cochrane Library searches for cost effectiveness were undertaken.

a. If these searches were undertaken in alternative resources, please provide the resource and search strategies.

b. Please provide a breakdown of Cochrane Library results per database for cost effectiveness searches.

Amgen response: DARE, CMR, NHS EED, and HTA are still considered core databases (as per NICE guidelines, see: https://www.nice.org.uk/process/pmg6/chapter/identifying-the-evidence-literature-searching-and-evidence-submission), despite the fact that these databases don't

capture all years to the present. Below is the full list of Cochrane Library databases that were searched via OVID:

- Cochrane Database of Systematic Review (CDSR) [covers all entries until present; updated online weekly]
- Database of Abstracts of Reviews of Effects (DARE) [covers 1991–2015]
- Cochrane Central Register of Controlled trials (CENTRAL) [covers 1991–present; updated online monthly]
- Cochrane Methodology Register (CMR) [covers 1995–2012]
- NHS Economic Evaluations Database (NHS EED) [covers 1995–2015]
- Health Technology Assessment Database (HTA) [covers 2001–2016]
- American College of Physicians (ACP) journal club [covers 1991–present]

The initial Cochrane cost-effectiveness searches run on 20 February 2020 identified 16 hits. No new hits were identified in Cochrane Library during SLR update in 2021. To demonstrate the breakdown of hits per database, the search was rerun on 29th July and resulted into 19 hits from the Cochrane Library (17 of them via Cochrane Central Register of Controlled trials (CENTRAL), 1 via DARE and 1 via NHS EED).

For search strategy used, please refer to the Appendix G (page 89-110, Appendices).

B3. Please provide search terms and hits per resource for congress searches undertaken for cost-effectiveness.

Amgen response: The following strategies were used for the original congress searches:

• AACR 2017–2019: Searched ‘Cancer Research’ journal for the for terms "NSCLC and KRAS" in volume "77" and issue "13 Supplement" (61 hits; 2017), in volume "78" and issue "13 Supplement" (77 hits; 2018), and volume "79" and issue "13 Supplement" (56 hits; 2019) and exported these for further review.

- ASCO 2017–2019: Searched ASCO website for “NSCLC and KRAS” and retrieved relevant abstracts from 2017 (6 hits), 2018 (3 hits) and 2019 (9 hits) for further review.
- ESMO 2017–2019: Searched ESMO website for “NSCLC and KRAS” and retrieved relevant abstracts from 2017 (17 hits), 2018 (7 hits) and 2019 (1 hits) for further review.
- WCLC 2017–2019: Searched the WCLC abstract booklets as PDFs for “NSCLC and KRAS” and retrieved relevant abstracts from 2017 (6 hits), 2018 (30 hits) and 2019 (39 hits) for further review.

As such, the congress search strategies were not restricted by any outcomes and any relevant cost-effectiveness study identified from AACR, ASCO, ESMO or WCLC was included.

The only congress that was hand-searched as part of the SLR update 2021 was the WLCL (year 2020; held in 2021). The hand-search strategy was done via the WLCL congress website abstract database. The keyword search term used was “KRAS,” and 29 abstracts matched that search term. Seven relevant abstracts were included and given a Study ID. All other congresses of interest were searched via Embase (see congresses coverage on Embase: https://www.elsevier.com/solutions/embase-biomedical-research/embase-coverage-and-content).

Model structure

B4. Priority question. The use of a partitioned survival analysis model instead of a state transition model (STM) was justified by stating that the same structure and health states were used in previous single technology appraisals (STAs). No further justification was provided. Despite the potential limitations of a STM, a partitioned survival analysis has several limitations related to the extrapolation, as mentioned in NICE Decision Support Unit (DSU) Technical Support Document (TSD) 19.

Please provide a scenario analysis (and the accompanying model), enabling STM as a scenario, as recommended in TSD 19.

Amgen response: TSD 19 does not recommend the replacement of partitioned survival analysis (PartSA) with state-transition models (STM, individual simulation or

Markov model) but recommends that assumptions behind each method should be clearly stated. It makes clear that any "choice of modelling approach may be constrained by the available evidence."

To summarise briefly TSD19 notes the two key downsides of the PartSA approach:

1. Assumes certain structural features about the disease progression and course
    - Transitions to states are modelled in an independent fashion and there can be an implicit assumption of no structural relationship between the transition to states
    - Assumptions are made about timing and order of transitions (e.g. patients cannot move from progressed to progression free)
2. Extrapolated periods are strongly related to the within-trial period: "trends in the hazard of each endpoint and treatment effects on these hazards observed within the trial period are assumed to generalise to the extrapolation period".

We would argue strongly that given the available data for sotorasib and the comparators, there is no strong case that a STM approach would mitigate the two downsides listed above.

Firstly, this is because the data requirements for a STM could not be met and so the differences in approach would be superficial in practice. OS and PFS are modelled independently of each other, but there are "baked in" structural relationships in this therapy area in that PFS also accounts for death events. There is no evidence that patients routinely (or at all) move backwards from progressed to a progression free health state and so this structural feature is consistent with the disease area (and virtually all other attempts at modelling NSCLC, irrespective of modelling approach).

Given that there is no external data outside of CodeBreak100 relating to sotorasib and virtually no data relating to the comparators (in the relevant mutation group) the method of informing transition probabilities in an STM (i.e. for the outside-trial period) would be identical (i.e. based on parametric survival analysis) and so there would be no difference in uncertainty relating to extrapolations. There is also no individual patient level data for the comparators, and this would also limit exploration of more

complex STMs (i.e. that could potentially produce divergent cost-effectiveness results).

Secondly, the submission and the responses to these clarification questions has provided the information/scenarios recommended in TSD19 to explore the two downsides listed above. In particular:

- Curve selection (e.g., fit statistics and plots to measure PH and AFT etc) was consistent with NICE recommended methods
- Plots were used to ascertain changes in hazard and inform curve selection
- An alternative (external) data source was used in scenario analyses to validate the base-case (i.e. propensity score matching with Flatiron RWE)
- The relatively blunt tool of treatment effect waning scenarios is explored in the submission and in these responses to explore long term relative treatment effects (see response to question B6 below)
- A CDF recommendation would also provide the opportunity for additional external data collection to validate long term extrapolations.

To conclude, in-line with virtually all NICE submissions in NSCLC (and particularly those in new therapy areas with single-arm trial evidence), Amgen believes the PartSA approach is the most appropriate to model the cost effectiveness of sotorasib.

Efficacy

B5. Priority question. Various distributions are tested to estimate long-term OS and PFS projections in section B3.3 of the CS, however, only a subset of tested distributions is included as scenario analyses in Table 48.

Please provide scenarios for all tested distributions for the OS and PFS projections, using all available datasets.

Amgen response: As requested, scenarios for all the combinations of OS and PFS restricted distributions for both the MAIC-weighted analysis vs. SELECT-1 and the supplementary propensity score weighting analysis vs. Flatiron have been added to the cost-effectiveness model (see table below).

It should be noted that such combinations of selected curves are arbitrary and were considered but ultimately rejected, based on the standard curve selection procedures outlined in NICE TSD documentation.

PFS and OS function scenarios

OS Distribution (Restricted)	PFS Distribution (Restricted)	Source for Comparison	Incremental costs (£)	Incremental QALYs	ICER (£ per QALY)
Exponential	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	51,099
Exponential	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	51,885
Exponential	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	48,710
Exponential	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	50,449
Exponential	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	53,374
Exponential	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	49,197
Gompertz	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	44,400
Gompertz	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	45,077
Gompertz	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,339
Gompertz	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	43,927
Gompertz	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	46,609
Gompertz	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,714
Weibull (PH)	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	59,356
Weibull (PH)	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	60,276
Weibull (PH)	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	56,562
Weibull (PH)	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	58,747
Weibull (PH)	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	61,926
Weibull (PH)	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	57,599
Gen. Gamma	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,232
Gen. Gamma	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,875
Gen. Gamma	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	40,276
Gen. Gamma	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,992
Gen. Gamma	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	44,977
Gen. Gamma	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	41,568
Log-logistic	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	46,117
Log-logistic	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	46,823

Log-logistic	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	43,970
Log-logistic	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	46,602
Log-logistic	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	49,117
Log-logistic	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	45,207
Log-normal	Exponential	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	44,361
Log-normal	Gompertz	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	45,039
Log-normal	Weibull (PH)	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	42,301
Log-normal	Gen. Gamma	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	45,123
Log-normal	Log-logistic	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	47,249
Log-normal	Log-normal	SELECT-1 MAIC, pre-specified vars	[redacted]	[redacted]	43,660
Exponential	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	50,425
Exponential	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	48,941
Exponential	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	48,389
Exponential	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	58,053
Exponential	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	62,123
Exponential	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	59,081
Gompertz	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	33,404
Gompertz	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	32,443
Gompertz	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	32,085
Gompertz	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	38,304
Gompertz	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	42,065
Gompertz	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	39,011
Weibull (PH)	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	52,153
Weibull (PH)	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	50,616
Weibull (PH)	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	50,045
Weibull (PH)	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	60,094
Weibull (PH)	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	64,051
Weibull (PH)	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	61,130
Gen. Gamma	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	31,344
Gen. Gamma	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	30,446
Gen. Gamma	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	30,112
Gen. Gamma	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	36,401
Gen. Gamma	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	39,603
Gen. Gamma	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	37,100
Log-logistic	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	36,151
Log-logistic	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	35,105

Log-logistic	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	34,716
Log-logistic	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	41,950
Log-logistic	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	45,764
Log-logistic	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	42,736
Log-normal	Exponential	CB100 vs Flatiron adjusted	[redacted]	[redacted]	32,326
Log-normal	Gompertz	CB100 vs Flatiron adjusted	[redacted]	[redacted]	31,398
Log-normal	Weibull (PH)	CB100 vs Flatiron adjusted	[redacted]	[redacted]	31,053
Log-normal	Gen. Gamma	CB100 vs Flatiron adjusted	[redacted]	[redacted]	37,554
Log-normal	Log-logistic	CB100 vs Flatiron adjusted	[redacted]	[redacted]	40,859
Log-normal	Log-normal	CB100 vs Flatiron adjusted	[redacted]	[redacted]	38,279

B6. Priority question. The efficacy of sotorasib compared to docetaxel is assumed to remain stable over the full time horizon of the model. Only in a scenario is it limited to 5 years. The ERG could not find any justification for the assumption of an lifelong treatment effect and considers this to be potentially overly optimistic given the short follow-up time of CodeBreaK100.

Please justify the assumption of lasting treatment effect or in fact any effect of treatment beyond what is observed in CodeBreaK100 and provide supporting evidence, including scenario analyses limiting the benefit duration of sotorasib to 2, 3 and 4 years.

Amgen response: The following treatment effect waning (TEW) scenarios have been added to the model and the results are presented in the order listed (see Table below):

1. Gradual 5 years waning of sotorasib vs. docetaxel from year 2
2. Gradual 5 years waning of sotorasib vs. docetaxel from year 3
3. Gradual 5 years waning of sotorasib vs. docetaxel from year 4
4. Gradual 5 years waning of sotorasib vs. docetaxel from year 5
5. No treatment effect of sotorasib vs. docetaxel after 2 years
6. No treatment effect of sotorasib vs. docetaxel after 3 years
7. No treatment effect of sotorasib vs. docetaxel after 4 years
8. No treatment effect of sotorasib vs. docetaxel after 5 years

The first 4 scenarios assume a gradual TEW over 5 years to explore the sensitivity of the model to further variation in TEW assumptions. There is no strong rationale for the use of 5 years and so it is as arbitrary as the chosen timepoint at which there is no treatment effect in scenarios 5 to 8.

TEW scenario results

	Incremental costs (£)		Incremental QALYs		ICER (cost/QALY)	
	[redacted]		[redacted]		43,660	
	SA Result	**Diff vs BC**	**SA Result**	**Diff vs BC**	**SA Result**	**Diff vs BC**
Gradual TEW scenarios	[redacted]	-0.60%	[redacted]	-10.20%	**48,332**	10.70%
	[redacted]	-0.40%	[redacted]	-6.80%	**46,659**	6.90%
	[redacted]	-0.20%	[redacted]	-4.70%	**45,682**	4.60%
	[redacted]	-0.20%	[redacted]	-3.30%	**45,062**	3.20%
No treatment impact scenarios	[redacted]	-1.40%	[redacted]	-24.30%	**56,904**	30.30%
	[redacted]	-0.80%	[redacted]	-15.10%	**50,997**	16.80%
	[redacted]	-0.50%	[redacted]	-10.00%	**48,228**	10.50%
	[redacted]	-0.40%	[redacted]	-6.80%	**46,684**	6.90%

It should be noted that TEW is a very blunt tool and in an ideal world its use should be limited to cases where there is no (or very little) available external data to compare or adjust long term extrapolations with.

It is important to note that a CDF recommendation would allow the collection of data to reduce uncertainty in extrapolations and long-term relative treatment effects. In particular:

- Codebreak100 will have up to 4 years of follow-up data that can be used to inform long term OS and PFS for sotorasib
- As explained in the submission, CodeBreak200 a phase 3 RCT comparing sotorasib against standard of care docetaxel in patients with NSCLC is ongoing and so this can be used to inform longer term relative treatment effects (OS and PFS) in a robust manner (potentially up to 6 years)

Nevertheless, there are several theoretical and clinical reasons that can be used to argue that TEW should be limited and applied carefully in this case:

- To a large extent, the impact of discontinuation on OS and PFS has already been “baked” into the hazard function (and so projected survival estimates) in that in the trial period in which parametric curves have been fitted a significant number of patients have discontinued over the period (>80%).
- According to Appendix E of the company submission and the related publication, by the time of the March 15, 2021 data cut-off around 80% (81.7%) of patients have discontinued treatment, around 40% remain alive and around 20% have yet to progress. Therefore, half of patients who are alive will have remained on sotorasib treatment at this point.
    - Sotorasib is given in CodeBreak100 until progression or the development of unacceptable AEs and so it is inappropriate to apply TEW early when a significant proportion of those alive are still benefiting from treatment.
    - Applying TEW in too blunt a fashion would bias cost-effectiveness results in that sotorasib arm patients continue to accrue the costs of treatment but not the relative benefits of the treatment.
- A case can be made that the mix of patient at the point of the March data cut (around 15 months of follow-up) in the sotorasib arm is in a better average “health state” than the docetaxel patients and so the hazards of survival will continue to be better for the former for some time.
    - In particular, in the docetaxel arm of the model at around 15 months, which is in the trial period of SELECT-1 and not an extrapolated portion, of the <30% alive only 4 percentage points of patients are progression free (the remaining progressed). In contrast, by this point around half of the patients in Codebreak100 who are alive (i.e. around 40%) have yet to progress (20 percentage points of patients).
- TEW is more intuitive and more easily defendable when the two treatments are comparable and have a similar mechanism of action and so a reasonable assumption can be made that the relationship between being on treatment and benefiting longer term are similar (e.g. two EGFR targeting TKI therapies). However, sotorasib and docetaxel are very different medicines with different actions and therefore such an assumption is more uncertain.

B7. Priority: Time to treatment discontinuation (TTD) for sotorasib was estimated by applying a hazard ratio (HR) to a PFS curve. Please clarify:

a. **why this approach was chosen over the approach in the sensitivity analysis, which is stated to be "more complex and ultimately dependent on the variable selection in the MAIC analysis" and would align with what was done in OS and PFS extrapolations, i.e. would have been more consistent?**

b. **whether this approach of applying HR to PFS to estimate TTD (while observed TTD is also available) has any precedent, e.g. by providing supporting evidence.**

Amgen response: There is a strong statistical case to "tether" TTD to PFS and this is consistent with the clinical use of sotorasib. In CodeBreak100, treatment with sotorasib continued until the occurrence of progressive disease, the development of unacceptable side effects, or withdrawal of consent. Indeed, the majority of discontinuation events were listed as due to progression. There is therefore a causal dependency structure between PFS and TTD (i.e. TTD will tend to be similar but "lower" than PFS) and so fitting the curves independently carries the risk of this relationship being violated in long-term extrapolations.

Fitting parametric curves and going through the standard curve selection procedures make sense when there is a large amount of uncertainty around TTD (i.e. it is immature). However, more than 80% of patients had discontinued sotorasib by the March 2021 data cut snapshot and so the relatively simple HR approach suffices and reflects the low uncertainty in this parameter.

It is important to note that a CDF recommendation would facilitate further reduction in uncertainty in this parameter: a fully mature CodeBreak100 TTD curve would be available, mature TTD for sotorasib vs docetaxel would be available from CodeBreak200, and time on treatment data would be available from real world use of sotorasib in the NHS.

There is a strong precedent based on previous NICE submissions (and acceptance by appraisal committees) for methods that "tether" TTD to PFS for oncology

medicines where this reflects how the treatment will be used – i.e. where TTD tends to be around PFS but not the same. For example:

- In multiple myeloma, a fitted HR applied to PFS to obtain TTD has been used and accepted before (e.g. see NICE TA657)
- In NSCLC it is common to assume TTD=PFS if this is in-line with the licence and SmPC wording or otherwise tether TTD to PFS by either applying a HR or adding a mean number of cycles of treatment at progression in the model (i.e. if TTD is slightly higher than PFS). For example, see NICE TA628, NICE TA670 and NICE TA406.

B8. It is not entirely clear whether the target populations for docetaxel and docetaxel + nintedanib are comparable. If they are, why were the comparators not combined in a full incremental analysis; if they are not, is the approach of applying the relative treatment effect from the LUME-Lung 1 comparison onto the SELECT-1 modelled docetaxel curve valid and justified?

Please explain and elaborate.

Amgen response: There is no easy way to produce a relative treatment effect between sotorasib and the nintedanib+docetaxel arm in LUME-Lung 1. Deriving fitted HRs from the LUME-Lung 1 arms and applying to the docetaxel arm in the model is imperfect, but it is at least applying a treatment effect that is relatively unbiased and has internal validity (i.e. it is derived from a large randomised phase 3 trial). For this reason, it is relatively balanced on treatment effect modifiers (and for selection bias) but an assumption must be made that the treatment effect can be carried over to the KRAS mutation sotorasib population.

The alternative would be an unanchored MAIC with CodeBreak100, but as explained in Section B.2.9.3.1 of the submission, such a MAIC would lack adequate baseline variables to be meaningful. In particular, most patients will not have KRAS mutated tumours in LUME-Lung 1 (and the proportion is not reported) and there are no patients in Codebreak100 that could be weighted to match the LUME-Lung 1 mix in terms of proportion with KRAS mutation. Such an analysis will not reduce the

amount of uncertainty in this secondary comparison and in addition would add a layer of uncertainty related to an unanchored comparison.

In addition, it should be noted that clinical consultation after receipt of these questions reinforced the clinical views presented in the submission. Nintedanib as add-on therapy to docetaxel is a minor comparator in terms of use and a minority who are eligible for docetaxel will have add-on nintedanib.

B9. Figures 14 and 19 show the fitted curves for OS and PFS, respectively, for the approach selected (i.e. restricted joint model) and with a time horizon of 5 years.

Please also provide Figures for the other approaches (i.e. unrestricted model and independent fits) and for the full time horizon of the analysis.

Amgen response: As requested, the independent fit models (i.e. with the same function for each arm) for OS and PFS are shown below (Figure 1 and Figure 2) for the full time horizon of the model. It should be noted that as clarified in response to Question B10, unrestricted and independent fit models are identical in practice.

Figure 1. Unrestricted (independent fit) parametric fits for OS

Figure 2. Unrestricted (independent fit) parametric fits for PFS

B10. Please provide explanation and detail on the differences between the unrestricted and restricted models, e.g. details on how the models were restricted.

Amgen response:

Practical explanation

The joint fit (unrestricted) model in practice is identical to independent fitting curves (i.e. fitting independent survival functions by arm). The unrestricted model fit statistics were presented so that the fit statistics (AIC/BIC) could be meaningfully compared between independent fitted and joint fitted models (please refer to the unrestricted column in fit statistics tables in the CS as the combined AIC/BIC for independent fits).

Statistical explanation

For the unrestricted models, any information relating to treatment arm does not inform the shape of the parametric distribution. In consequence, the curves of both treatment arms do not only differ in terms of a location parameter, but also the parameters that determine the shape are being estimated independently.

In contrast, for the restricted model the treatment difference in both parameters depends solely on a location parameter. The shape determining parameters are estimated jointly. For the generalized gamma, log-logistic and log-normal distribution, the restricted model corresponds to an accelerated failure time model. For the Gompertz and the Weibull distribution the implemented restricted time to event

approach corresponds to a proportional hazards model. For the exponential distribution there is no difference between the restricted and the unrestricted model, as by treatment arm there is only one location parameter (i.e. the time-independent event rate).

B11. Please add real-world data to the predictions in Table 25 of the CS.

Amgen response: Table 25 in the submission presents landmark model predicted OS for the purposes of clinical validation (and curve selection). There is no real-world data that Amgen is aware of that can be used as validation.

The alternative real-world Flatiron data is used to proxy docetaxel by weighting (via propensity score matching) observational data for a basket of chemotherapies to match CodeBreak100.

B12. In Figures 24 and 29 of the CS, please provide fitted curves for all considered models, comparable to Figures 14 and 19, for the full time horizon of the analysis.

Amgen response: This question requests ATT (Flatiron analysis) fitted OS and PFS curves for all survival models – see Figure 3 and

Figure **4** below.

Figure 3. ATT (Flatiron analysis) OS parametric fits for restricted and unrestricted models

Figure 4. ATT (Flatiron analysis) PFS parametric fits for restricted and unrestricted models

B13. Please comment on the appropriateness of choosing joint fit curves for sotorasib vs. docetaxel OS and PFS, given that data are collected from two separate trials. The argument of fewer parameters and a larger data set may not be sufficient to overcome the differences between the two comparators.

Amgen response: The issues that arise from using different trial arms to inform the relative efficacy between sotorasib and docetaxel relate to the biases that arise from unanchored comparisons that may not fully adjust for selection bias and the

imbalance of treatment effect modifiers. These issues are related to the rationale for conducting a MAIC and are discussed in section B.2.9 of the CS and Appendix D, D.1.5 (feasibility assessment, variable selection vs sample size, etc). Therefore, to the extent that the MAIC is robust these differences have been adjusted for.

Much of the argumentation about Joint fits vs independent fits is analogous to the arguments around function selection and have been framed this way in the relevant submission sections – i.e. the choice has been justified based on comparison of fit statistics and diagnostic plots (see Section B.3.3.3).

To illustrate the point, these methods that measure the fit to the observed trial period could, for example, suggest selection of independent curves over joint fits when data is also from the same trial (e.g. proportional hazards is not satisfied and AIC/BIC favour the former).

B14. For the nintedanib + docetaxel analysis, please provide

a. log cumulative hazard plots for both OS and PFS,

b. the survival curves resulting from the extrapolations using the piecewise approach, alongside the Kaplan-Meier (KM) curves for both OS and PFS (also in the accompanying model), for various fitted parametric functions as in Figure 19,

c. predictions of OS and PFS in percentages as in Table 25, and

d. any information available on the clinical plausibility of these curves.

Amgen response: As requested, below are the log cumulative hazard plots for OS and PFS (Figure 5 and Figure 6). In the submission the justification for the piecewise application of HRs (and related time points) was argued based on visual inspection of the KM survival and instantaneous hazard plots from LUME-Lung 1. The log-cumulative hazard plots displayed are also further evidence of the choice of these time points: 6m and 26m for OS, 2m and 6m for PFS.

Figure 5 shows that at 6m the curves cross indicating an inflection point whereby the curves separate until around 26m where they become almost parallel. Figure 6 shows that just before 2m the curves are broadly parallel but separate sharply and

then begin to close in around 6m. Nevertheless, to support the answer to question B16, sensitivity analyses are undertaken for the OS piecewise analysis.

Figure 5: Log-cumulative hazard plot (OS)

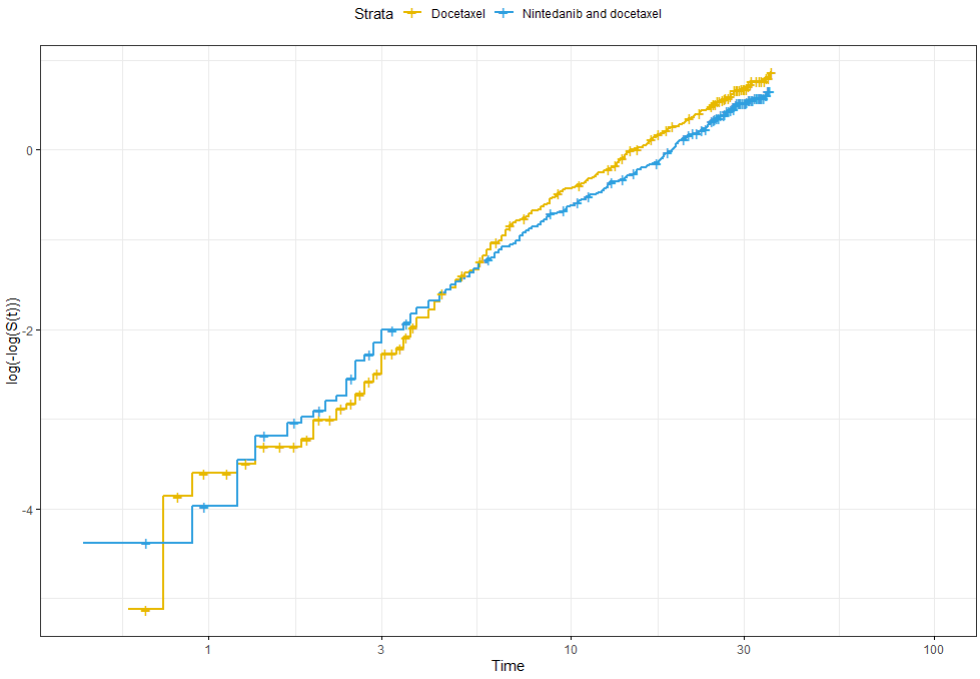
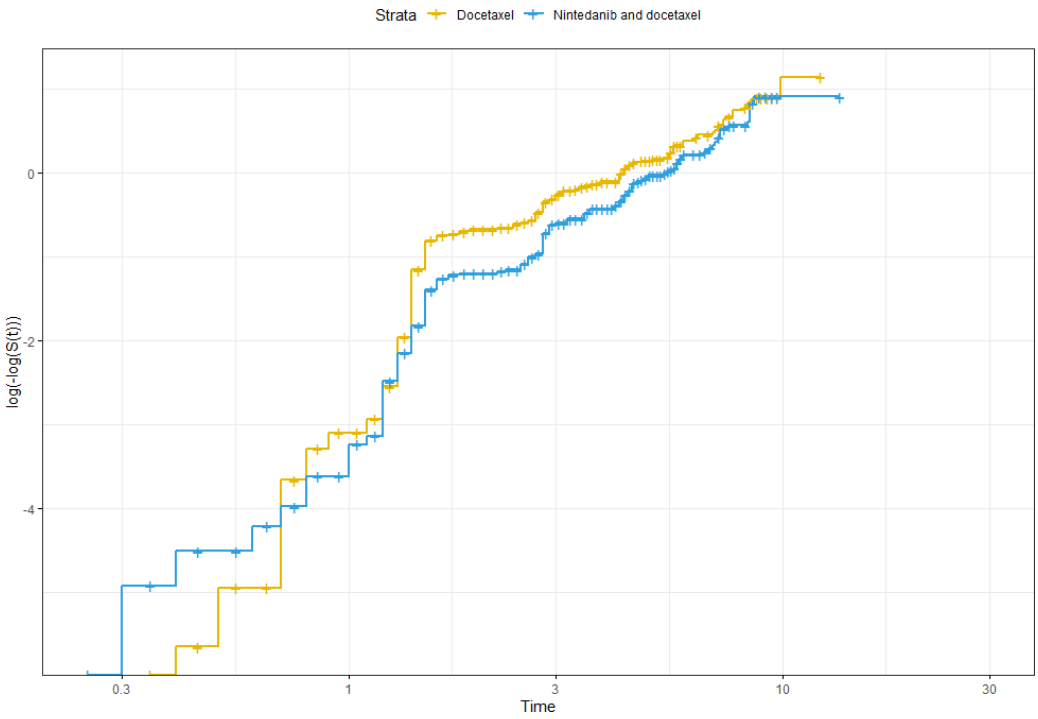


Figure 6: Log-cumulative hazard plot (PFS)



Part b of this question requests the visual representation of curve fittings across the model time horizon for all functions related to the nintedanib+docetaxel analysis. It is important to understand that this analysis involved applying LUME-Lung 1 derived HRs to the SELECT-1 docetaxel arm to produce an estimated survival curve for nintedanib+docetaxel - this latter curve will depend on the chosen survival function for docetaxel selected in the model.

In

Figure 7 OS fitted distributions for docetaxel are shown and for easy viewing underneath the respective docetaxel function after the HRs have been applied (i.e. the estimated nintedanib+docetaxel curves conditional on choice of survival function for the docetaxel arm). Note that only the KM for docetaxel is presented because there is no meaningful KM curve for nintedanib+docetaxel. The nintedanib+docetaxel KM from LUME-Lung 1 is not relevant for comparison because these parametric

curves are derived by applying a relative treatment effect to docetaxel curves from a different study (SELECT-1). An analogous approach is taken for PFS (

Figure 8).

Figure 7: OS for nintedanib and docetaxel vs. docetaxel alone (all distributions)

Figure 8: PFS for nintedanib and docetaxel vs. docetaxel alone (all distributions)

In the table below we have presented the proportions surviving for different survival functions. There has been no explicit validation of the proportions for nintedanib+docetaxel undertaken with clinicians. However, based on discussions with clinicians there is some consensus that adding nintedanib to docetaxel does not change efficacy greatly and this is largely borne out by the proportions presented. These proportions also confirm with the piecewise HR approach: the hazard of death is higher for nintedanib early then lower later and this is consistent with these proportions (i.e. adding nintedanib lowers OS at 1 year relative to docetaxel but increases it later on).

Proportion of alive patients at different landmark timepoints

	Exp	Gompertz	Weibull	GG	Loglogistic	Lognormal
Docetaxel						
1 Year	40.9%	40.7%	40.8%	38.9%	37.8%	38.8%
5 Years	1.1%	1.7%	0.4%	6.4%	5.2%	4.6%
10 Years	0.0%	0.1%	0.0%	2.2%	1.9%	1.1%
Nintedanib and docetaxel						
1 Year	39.2%	38.9%	39.8%	37.0%	36.5%	37.2%
5 Years	2.4%	3.5%	1.2%	9.2%	7.9%	7.2%
10 Years	0.0%	0.2%	0.0%	3.1%	2.9%	1.7%
Key: Exp, exponential; GG, generalized gamma; OS, overall survival						

B15. The ERG has great difficulty identifying the landmark 26-month time point at which the hazards are supposed to change direction in Figure 32 (OS LUME-Lung 1 instantaneous hazards plot). One could also argue they remain rather constant, and sample size at 26 months may not justify changing HR at this point.

Please elaborate on whether the 3-piece approach is justified and why.

Amgen response: As described above in the response to B14, the log-cum plots add some additional weight to the selected timepoints, especially for OS. However, we accept that there is a certain amount of uncertainty about what points in the piecewise HR analysis to select. Therefore we have presented in the table below additional sensitivity analyses around this second chosen time point of 26m for the OS analysis: 24m and 28m are chosen, HR refitted in the same way via Cox PH model, and cost-effectiveness results presented for these scenarios. It can be seen that fitted HRs are relatively stable and therefore so are model results.

Scenario analyses related to 2nd time point for piecewise HR analysis

Time periods	Hazard ratios	Inc. results
Current base-case		
0 -6 months	xxxx	Incremental costs (£): [redacted]
6 – 26 months	xxxx	Incremental QALYs: [redacted]
26+ months	xxxx	ICER (£/QALY gained): 30,899
[redacted]		
0 -6 months	xxxx	Incremental costs (£): [redacted]
6 – 24 months	xxxx	Incremental QALYs: [redacted]
24+ months	xxxx	ICER (£/QALY gained): 31,330
[redacted]		
0 -6 months	xxxx	Incremental costs (£): [redacted]
6 – 28 months	xxxx	Incremental QALYs: [redacted]
28+ months	xxxx	ICER (£/QALY gained): 30,712

Health-related quality of life (HRQoL)

B16. Priority question. The justification for using time to death (TTD) utilities in the base case instead of health state utilities was mainly justified by a) utilities

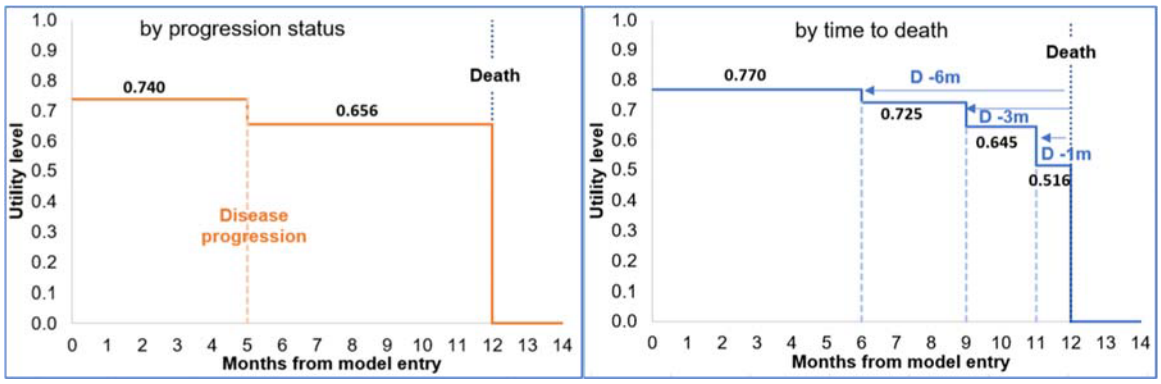
were shown to decrease towards end of life, b) other STAs that used the TTD approach as well, and c) because UK clinical experts considered this to better reflect the experience of patients.

a. **Please elaborate on why the TTD approach is superior to the health state approach and was chosen to be the base case.**

b. **Please provide detailed information on the contents of the clinical expert opinion mentioned.**

Amgen response: The analysis presented in section B.3.4.1 of the submission uses MMRM models (scenario analyses and stepwise selection of variables etc) to ascertain the impact of progression and time to death on EQ-5D measured quality of life in CodeBreak100. This analysis suggests that both progression and time to death are important variables for predicting QoL; in particular, there is evidence of a marked drop in QoL from greater than 6m to death and within a month to death (and relatively stable between these points). There is simulation study evidence that suggests that health state utilities can produce a poor fit to observed trial QoL if patient QoL varies significantly with closeness to death (see van den Hout WB, et al. JNCI 2006;98 (24):1786-94 and Hatswell AJ, et al. Health and Quality of Life Outcomes 2014;12:140).

Following the initial advisory board described in response to Question A16, two independent additional in-depth interviews were conducted with oncologists to validate outcomes of survival analyses, cost and resource utilisation and utility approach. With respect to the utilities, the methodology behind each approach was presented alongside conclusions of the literature cited in our original submission and past precedent from NICE appraisals. The clinicians were shown graphical representations of both utility approaches and asked to comment on which approach best reflects their experience of quality of life in NSCLC:

- 

Clinician responses were clear and consistent with previous literature on the subject – in both cases, the clinicians recognised that patients were likely to maintain a relatively high-quality of life and experience rapid and substantial decline as they approached death. Importantly, it was clear that progression-status was likely to be less of a driver of reduced quality of life than time-to-death, and it was unrealistic to maintain a relatively high-quality of life up until the point of death. Based on this feedback and the additional rationale presented in our submission the TTD approach is a highly defensible approach to modelling lifetime modelled QALYs.
It is important to note that a CDF recommendation would facilitate collection of comparative QALY accrual for sotorasib vs docetaxel because EQ-5D will be collected in the phase 3 CodeBreak200 study. This will reduce the uncertainty on the QoL side of the cost-effectiveness modelling.

B17. No age-related decrement was applied (or correction for general population utilities).

Please provide a scenario analysis including age related utility decrements.

Amgen response: A scenario analysis has been conducted applying an adjustment to utilities based on the sex-matched general population utilities from Ara and Brazier, to ensure that the estimated patient utilities never exceed that of the general population. This scenario had a minor effect on the results with the ICER rising only slightly to £43,715. It should be noted that the time-to-death utility values already account for aging in that the method applies utility decrements as a patient age and grows closer to death.

B18. The source for the utility decrement of docetaxel intravenous (IV) is a study using visual analogue scale (VAS) scores to determine quality of life (QoL) in the PFS state for oral vs. IV therapy. The utility scores reported here for PFS (0.426 and 0.451) are vastly different from the one reported in CodeBreaK100 (0.74). Please clarify:

a. why the decrement from this specific source is considered to be valid for the population in this STA?

b. whether it is reasonable to apply this decrement continuously while on treatment, given that docetaxel is administered once every 3 weeks. NB: Cytotoxicity should be captured in AE decrements.

c. why no utility decrement is applied to taking 8 tablets on a daily basis, as is the case for sotorasib treatment.

Amgen response: It is important to note that the utility analysis described in the submission is highly conservative in general and so the question can be answered in this context. The analysis assumes that utility weights (either for base-case TTD or HS approaches) are equal for a targeted therapy (sotorasib) vs chemotherapy (docetaxel). This is unusual in the modelling of targeted NSCLC medicines (EGFR, ALK etc) where, for example, utilities weights are rarely assumed to be the same between a targeted therapy and chemotherapy when patients are still on treatment (or in pre-progression) but are usually only assumed equal only when off treatment (progressed).

This is supported by previous literature:

- This has been the approach in many NSCLC appraisals to NICE when the comparator is a chemotherapy (NICE TA628, NICE TA416, NICE TA406, NICE TA422)
    - For example, in the most recent ALK appraisal (TA628) this approach was considered acceptable by the ERG and committee and was justified as follows in the submission: "*It was considered appropriate to apply treatment-specific utilities given that patients receiving chemotherapy are likely to have a poorer HRQoL than patients on ALK TKIs. This was found in PROFILE 1007, where utilities for the ALK TKI*

crizotinib (0.82, 95% CI: 0.79−0.85) were significantly greater (p<0.05) than for PDC (0.73, 95% CI: 0.70−0.79). Further to this, within the HRQoL SLR, seven out of the 10 studies identified progression free treatment-specific utilities. For four of these studies, a comparison between ALK TKIs and chemotherapy was available and, in all instances, a utility decrement was applied for patients on chemotherapy compared to those receiving treatment with an ALK TKI (0.02–0.08)."

- It is supported by a real-world study, that is heavily referenced in NICE appraisals, which compares utilities of the most appropriate targeted therapy to chemotherapy for different NSCLCs (Labbe et al. Clin Lung Cancer 2017; 18(4): 388-95.)

As noted in the CS, the SLR did not identify any studies that provided HRQoL estimates for NSCLC with KRAS p.G12C mutation or more broadly with any KRAS mutation and so there are no point estimates available for HRQoL in general (or PFS specific) for patients on chemotherapy that could be applied in scenario analyses. However, it should be noted that the applied 0.025 decrement is on the lower end of the 0.02-0.08 decrement range referenced above (i.e. for ALK mutation NSCLC PFS utility for targeted therapy vs chemotherapy).

Two alternative scenarios are provided to explore the uncertainty:

1) Proportional decrement equivalent to 0.04 (base-case TTD utility): the decrement referenced in the question is taken as a proportional drop instead of absolute (so a 5.5% decrement) and applied to the PFS utility from CodeBreak100 (0.734), which results in an absolute decrement of 0.040

2) Lower 0.687 PFS utility for docetaxel (HS utility scenario only): EQ-5D utilities from SELECT-1 are not available (as evidenced by the results of the SLR), however a PFS utility is available from LUME-Lung 1 (used in NICE TA347 and NICE TA416) and this is applied for PFS in the docetaxel arm only

    - The effective decrement applied would be 0.734 - 0.687 = 0.047 for PFS (sotorasib vs docetaxel)

The table below presents the results for these alternative scenarios. For scenario 1, when a proportional decrement is applied to the TTD base-case scenario it results in higher incremental QALYs and a lower ICER. For scenario 2, as expected the ICER is slightly lower compared to base-case settings with HS utility weights applied.

Results for alternative utility scenarios

Scenario	Treatment	Incremental costs (£)	Incremental QALYs	ICER (£ per QALY)
1	Docetaxel	[redacted]	[redacted]	£43,253
	Sotorasib			
2	Docetaxel	[redacted]	[redacted]	£45,783
	Sotorasib			

It is important to note that a CDF recommendation would facilitate collection of comparative QALY accrual (and HRQoL over time) for sotorasib vs docetaxel because EQ-5D will be collected in the phase 3 CodeBreak200 study.

Costs

B19. Relative dose intensity (RDI) is included in cost calculations for all three agents.

a. Given that TTD for sotorasib is already shorter than PFS, please discuss the reasons be for dose reductions when the AE profile is relatively mild.

b. Would it be reasonable to assume that sotorasib has the lowest RDI of all three agents? If so, please elaborate on the underlying reasons.

Amgen response: The AE profile is relatively mild in general and compared with comparator treatments (see Table 32 of the CS), as confirmed by UK clinical experts (see response to Question A16).The AE profile is relatively mild because sotorasib, as a targeted therapy (in contrast to chemotherapy) has low/no off-target cytotoxicity. However, like all active therapies, sotorasib is not free of the potential for adverse events, and the draft SmPC (provided in the reference pack) details the recommended dose modifications in the event of specific adverse reactions. As an

oral therapy, sotorasib allows the easy implementation of dose reductions to help manage AEs.

It would not be reasonable to assume sotorasib has the lowest RDI of all 3 agents as evidenced by table 37 in the submission – all 3 have very similar RDIs, the differences in which may reflect random sampling error.

B20. Subsequent treatment costs are included in the model. After disease progression, 40 % of patients are expected to switch to docetaxel monotherapy in the sotorasib treatment arm.

a. Please elaborate on how many patients with an adenocarcinoma histology would be expected to switch to docetaxel + nintedanib therapy.

b. What are the costs associated with this switch and what is the effect on the incremental cost effectiveness ratio (ICER)?

Amgen response: Following receipt of the clarification questions, Amgen consulted two clinicians who again reiterated that 50% or more of patients will receive BSC as a subsequent treatment to sotorasib, with most of the remaining patients receiving docetaxel.

From CodeBreak100 (March 2021 data cut) there are 77 records for 44 patients having received active/systemic anti-cancer therapies subsequent to discontinuing sotorasib. Assuming around 20% of patients (≈25 patients) had yet to discontinue from sotorasib (see response to Question B6), of the remaining 101 patients, 57 (i.e. 101-44 = 57) received no active anti-cancer therapy. Therefore, based on CodeBreak100 data, around 56% of patients (i.e. 57/101) are estimated to have received BSC subsequent to sotorasib, which is consistent with the original company submission.

Of the 77 records of subsequent treatment there are no recorded cases of nintedanib or nintedanib+docetaxel. (see table below) Therefore, there is no evidence either based on clinical opinion or trial data, that suggests patients will receive nintedanib+docetaxel subsequent to sotorasib.

Treatment mix for 44 patients receiving subsequent treatments

Treatment	N	Proportion of 77 treatments (%)
Pemetrexed or docetaxel	xxxx	xxxx
Platinum based chemotherapy	xxxx	xxxx
Others* or non-interventional therapy	xxxx	xxxx
Total	xxxx	xxxx

* other includes novel treatments assessd in clinical trial settings and other treatments not relevant to UK clinical practice (n=39) or unknown (n=1)

B21. The base case assumes zero wastage of sotorasib.

Please elaborate on this and provide supporting evidence.

Amgen response: In the economic analysis two scenarios relating to the costing of sotorasib were explored to capture the potential impact of drug wastage in clinical practice. The first scenario, used in the base case analysis, estimated the total treatment cost per treated day in the model and explicitly maintains the dose-efficacy relationship observed in CodeBreaK100 to provide an unbiased estimate of the cost-effectiveness of sotorasib.

The second scenario, captured in the reported scenario analyses, estimated the total treatment cost per opened pack. In other words, this captures the total cost of all opened packs thus reflects an upper bound of potential drug wastage in clinical practice.

However, based on the disease management of NSCLC and discussion with an NHS pharmacist, we maintain that the base case may provide a better reflection of likely drug utilisation in clinical practice. As discussed in Section B.3.5.1 of the CS, the ability to implement dose reductions and the single-strength formulation of sotorasib allows pharmacists to optimise the sotorasib-dose (without wastage) and provide the appropriate supply of drug to patients until disease progression is recorded in clinical practice. Furthermore, given the close relationship between PFS and TTD and the low number of PFS events reported as deaths (13/83, 15.6%; CodeBreak100 Dec 2020 data cut) the scenario analysis presented will significantly overestimate the true drug utilisation of sotorasib and does not reflect likely clinical experience.

Sensitivity analyses

B22. A probabilistic sensitivity analysis (PSA) was run for 1,000 iterations for the base-case analysis (sotorasib versus docetaxel).

In the model when the patients' characteristics are set to zero (for example sex and body surface area (BSA)), there are negative values for the low values column in "Model parameters" sheet. The ICER for base case does not change by setting sex and/or weight to zero. The ICER had a small change when was BSA set to zero (£47,820 vs £47,146).

a. Please explain why there are no or only small changes by setting these characteristics to zero.

b. Please elaborate on how this was controlled for the negative values.

Amgen response: In the cost-effectiveness model, sex is used to estimate the general population mortality at a given age. General population mortality is applied as a floor to ensure that the projected survival for patients in our model never exceeds that of the general population. As survival in this disease area is low, there was never the case that survival projections exceeded that of the general population. Therefore, changing the sex distribution does not affect results in the model as there is never a timepoint that we switch to general population mortality.

None of the treatments modelled in the cost-effectiveness model were weight-based therapies, therefore changing the weight has no impact on model results. We have now removed the weight input from the model as it is not relevant.

Both docetaxel and vinorelbine, which is included in the platinum doublet subsequent therapy, are dosed based on a patient's BSA. When setting BSA to 0, it can be seen in the model that the medication costs of docetaxel and vinorelbine are both 0, as is expected. The reason that this has a small impact on the ICER is that these treatments are relatively inexpensive - £17.95 per 160 mg vial for docetaxel, dose once every three weeks, and £27.73 per 20 mg for vinorelbine, dosed twice every three weeks.

To ensure that the low values of age, sex, and BSA never fall below 0, in the updated model we have now switched from calculating the low and high values using the normal distribution to the gamma distribution.

Validation

B23. The real-world data from the Flatiron study are presented as a way to validate the model results. However, in appendix J, the information provided is very minimal and not sufficient to assess validity of model results.

a. Please provide more information on the Flatiron study, e.g. how representative it is for the population in the current STA.

b. Kindly provide Flatiron observed OS and PFS for the time points reported in Table 25 instead of the 'Flatiron model analysis' now provided in Appendix J.

Amgen response: Appendix D (section D.16) provides a detailed exploration of the propensity score Flatiron analysis, including:

- Study design
- Description of the Flatiron patient mix
- Statistical model implemented
- Covariates before and after weighted/adjustment and diagnostic plots
- Treatment mix
- Unweighted and weighted results

As requested, survival analysis predictions for the alternative Flatiron analysis are provided below for OS and PFS.

OS predictions for models using joint fitted (restricted model), Flatiron analysis

	Exp	Gompertz	Weibull	GG	Loglogistic	Lognormal
Sotorasib						
1 Year	56.5%	56.3%	56.5%	56.6%	55.8%	56.6%
5 Years	5.8%	11.2%	5.2%	14.8%	13.2%	14.0%
10 Years	0.3%	3.9%	0.2%	6.0%	5.7%	5.3%
Docetaxel proxy						

1 Year	38.2%	37.5%	38.2%	34.9%	34.7%	35.0%
5 Years	0.8%	2.4%	0.7%	5.7%	6.0%	5.1%
10 Years	0.0%	0.4%	0.0%	1.9%	2.5%	1.5%
Key: Exp, exponential; GG, generalized gamma; OS, overall survival						

PFS predictions for models using joint fitted (restricted model), Flatiron analysis

	Exp	Gompertz	Weibull	GG	Loglogistic	Lognormal
Sotorasib						
1 Year	29.1%	28.8%	27.9%	29.1%	28.6%	29.2%
5 Years	0.2%	0.0%	0.0%	1.8%	3.1%	2.1%
10 Years	0.0%	0.0%	0.0%	0.3%	1.1%	0.4%
Docetaxel proxy						
1 Year	19.5%	19.4%	18.5%	17.5%	17.3%	17.6%
5 Years	0.0%	0.0%	0.0%	0.6%	1.7%	0.8%
10 Years	0.0%	0.0%	0.0%	0.1%	0.6%	0.1%
Key: Exp, exponential; GG, generalized gamma; PFS, progression-free survival						

B24. As reported in Appendix J, *"the clinical outcomes generated using this alternative data source were consistent with the conclusions of the MAIC analysis and underline the robustness of the analyses presented"*. The ERG agrees that OS estimates seem consistent. However, the PFS estimates from the Flatiron approach are substantially higher compared to the base case though for both arms (xxxx for sotorasib, and xxxx for docetaxel).

Please explain this difference.

Amgen response: We agree that the OS results from the Flatiron conform more with the MAIC analysis than do the PFS results. However, considering the following contextual points both analyses seem to produce remarkably similar results:

- There will be differences that derive from a real-world setting (Flatiron) vs a controlled trial setting (SELECT-1)
    - In the real-world disease progression is derived from physician notes in a clinical practice setting and may be informed by RECIST criteria in conjunction with other signs of progression

- As discussed in the submission, the Flatiron data is used as a proxy for docetaxel and is based on a basket of standard of care chemotherapies
- There will necessarily be other differences in the patient mix related to the data sources for the comparator in each analysis even after the impact of statistical weighting (i.e. SELECT-1 vs Flatiron)

- The data points compared in the question are 1-year proportions taken from the model that have been informed by treatment effects derived from two statistical analyses that use the CodeBreak100 sotorasib and comparator source (SELECT-1 or Flatiron) in different ways
  - In-line with recommended MAIC methodology (and data availability) the MAIC analysis weighted CodeBreak100 patients to match SELECT-1 to derive a meaningful relative treatment effect. In contrast, for the Flatiron propensity score matching analysis the RWE source is weighted to match CodeBreak100.

Section C: Textual clarification and additional points

C1. According to section B.3.3.2, *"the relative treatment effect of nintedanib plus docetaxel vs docetaxel was applied to the SELECT-1 modelled docetaxel curve"*.

Please confirm that by 'modelled docetaxel curve' both the OS and PFS curves are meant.

Amgen response: Yes – please see section B.3.3.5 of the CS for full details.

C2. Please provide legends for Figures 30 and 33 of the CS.

Amgen response: Please see figure 30 and figure 33 with legends inserted below

Figure 30 from the CS: OS Kaplan–Meier plot from LUME Lung-1 for nintedanib versus placebo

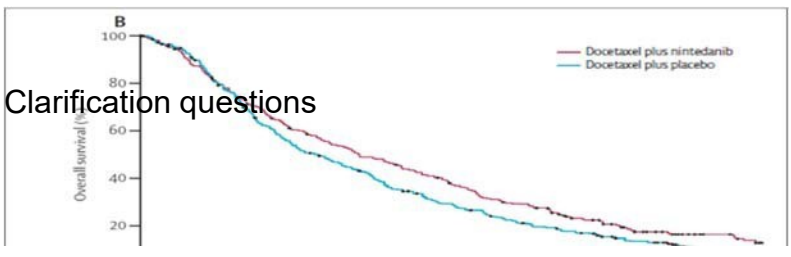
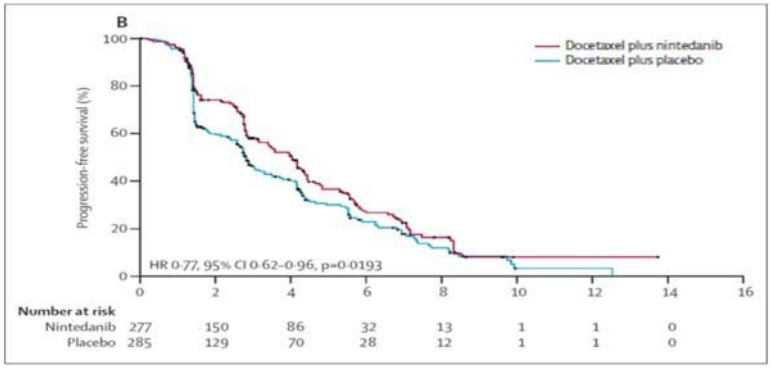


Figure 33 from the CS: PFS Kaplan–Meier plot from LUME Lung-1 for nintedanib versus placebo



C3. Please correct the footnotes in Table 37 of the CS as “a” is mentioned twice.

Amgen response: Please see table 37 corrected for the footnote.

Table 37 from the CS: Unit drug costs

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)[f]
Sotorasib	240 x 120 mg tablets	[redacted]	[redacted]	960mg per day	89.2%[b]	[redacted]
Docetaxel	160 mg per vial	17.95	eMIT[80]	75 mg/m^2 on day of treatment	90.3%[c]	19.93[e]
Nintedanib	120 x 100 mg tablets	2,151.10	BNF[82]	400 mg per day (21-day cycle)[a]	92.1%[d]	1,926.28

Key: BNF, British National Formulary; eMIT, electronic market access tool.
Note:
[a] Nintedanib administered on days when docetaxel is not taken, i.e., 20 days per 21 day cycle,
[b] CodeBreaK100 CSR (01DEC2020), Table 14b-5.1, Exposure to sotorasib (AMG510)
[c] Jänne, 2017[31]
[d] Reck 2014[32]
[e] Docetaxel cycle cost is based on cost per mg x dose per administration (75 mg/m^2) x body surface area (1.81 m2)
[f] calculated from CEM

C4. In Table 48, reporting scenario analysis results, the 15-year time horizon has been duplicated, please correct.

Amgen response: please see Table 48 below corrected by removal of the duplicate row for the scenario analysis using a 15 year time horizon.

Table 48 from the CS: Scenario analysis results

Scenario	Rationale/Justification	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
15-year time horizon	To investigate the impact on model results of reducing the model timeframe.	[redacted]	[redacted]	48,197
Generalised gamma distribution selected to estimate long-term OS and PFS projections	The generalised gamma was the 2nd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more favourable estimation of survival.	[redacted]	[redacted]	45,612
Log logistic distribution selected to estimate long-term OS and PFS projections	The log-logistic was the 3rd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more conservative estimation of survival.	[redacted]	[redacted]	53,606
Joint (unrestricted) lognormal distribution selected to estimate long-term PFS	Although BIC consistently favoured the restricted versus unrestricted joint fit this scenario tests an alternative parametric distribution using the unrestricted lognormal for PFS which was the best fitting unrestricted model based on BIC criteria.	[redacted]	[redacted]	52,495
Adjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1 using all available covariates	To test the robustness of the MAIC using alternative MAIC model where all available covariates are considered. A joint (restricted) lognormal distribution was used per the base case analysis and based on the survival analysis selection criteria outlined in Appendix N.	[redacted]	[redacted]	39,645
Unadjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1	To investigate the impact of a naïve comparison to the SELECT-1 clinical trial. A joint (restricted) lognormal distribution was used per the base case analysis.	[redacted]	[redacted]	53,794

Unadjusted sotorasib from CodeBreaK100 vs. ATT-adjusted docetaxel from Flatiron	Alternative data source which included patients closely aligned with the CodeBreaK100 population in terms of prior treatment from the real-world Flatiron dataset to test the robustness of the results in the base case analysis	[redacted]	[redacted]	39,773
MAIC-adjusted TTD curve from CodeBreaK100	To test the impact of an alternative approach to estimate long-term treatment duration.	[redacted]	[redacted]	50,810
HR of sotorasib vs. docetaxel = 1 after 5 years	In the base-case PFS and OS were modelled based on parametric survival distributions fit to survival data from CodeBreaK100 and SELECT-1, combined with age- and sex-matched general population mortality. This scenario explicitly limits the duration of benefit to 60 months.	[redacted]	[redacted]	49,956
Apply health state utilities by progression status	To test the impact of an alternative method for measuring health state utilities as described in Section B.3.4.5	[redacted]	[redacted]	51,079
Treatment-emergent AEs	To test the impact of utilising treatment-emergent adverse events for sotorasib and docetaxel as described in Section B.3.3.7	[redacted]	[redacted]	47,495
Include drug wastage	To test the impact of potential drug wastage in clinical practice by estimating drug acquisition costs based on total packs as opposed to treatments received	[redacted]	[redacted]	50,216
Exclude RDI	To test the impact of not capturing RDI on drug utilisation calculations	[redacted]	[redacted]	52,757
1.5% discount rate for costs and efficacy	To investigate the alternative discount rate suggested by the NICE Guide to Technology Appraisal. A reduced discount rate of 1.5% is consistent with the Treasury Green Book and is being considered in the ongoing NICE Methods Review consultation.	[redacted]	[redacted]	44,505

Key: ATT, average treatment effect of the treated; HR, hazard ratio; OS, overall survival, ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALY, quality-adjusted life year; RDI, relative dose intensity.

C5. Please correct the legend of the PFS figures in the model ("Sheet PFS").

Amgen response: the legend of the figure has been corrected in the revised model that has been provided based on the March 2021 data cut.

C6. In a number of places throughout the submission (e.g. sections B.3.3.3 and B.3.9.1), only clinical expert opinion is cited in support of text provided in the CS.

Please confirm that this was indeed the only evidence available.

Amgen response: Clinical expert opinion was generally used to validate assumptions and/or data in the CS, rather than being the source of evidence. E.g. in section B.3.3.3, it is noted that clinical expert opinion was sought to advise on which variables were important prognostic factors for the MAIC analysis. The starting list of variables was informed by the literature as detailed in section B.2.9.3.1, so whilst the text in B.3.3.3 refers only to clinical expert opinion, there was other evidence considered upon which the clinical experts expressed their views. Also in B.3.3.3, clinical expert opinion was used to validate survival curves, but of course the survival curves were informed by evidence from the relevant trials.

C7. There is a mismatch between section B.2.6 (Overall response rate (ORR) defined as *"complete response + partial response"*) and Table 5 (ORR defined as *"Proportion of subjects with best overall response of complete response or partial response as assessed by RECIST 1.1"*).

Please clarify the definition of ORR used in the trial.

Amgen response: An objective response in an individual patient is defined as either a complete response (CR) or a partial response (PR). The best response achieved in a patient is used to define their objective response. The Overall response rate (ORR) is the proportion of patients who achieve an objective response - either a CR or a PR (but not both – only a patient's best response counts). ORR is estimated from the sum of patients who achieve a CR and patients who achieve a PR (i.e. CR + PR). The definitions of ORR in B2.6 and in Table 5 are therefore compatible. The correct definition is that presented in Table 5, with ORR= CR+PR used as a shorthand.

C8. Please clarify the date of the data cut for the document "Amgen Inc_CodeBreaK100 Safety data (01 December 2021 data cut).pdf" which wasprovided as a reference for this CS.

Amgen response: The data of the data cut is 01 December 2020. The reference to 01 December 2021 in the file name is a typographical error.

C9. Please amend the incomplete sentence, ending "including", at the end of section 1.1 of document B of the CS.

Amgen response: We are unclear where the incomplete sentence is located in section B.1.1.

C10. Sections B.2.6.2.3 and B.2.13.2.1 refer to "section 0".

Please correct the cross-reference or provide the missing section.

Amgen response: apologies for the loss of this cross reference.

In both section B.2.6.2.3 and B.2.13.2.1 the text should read: *…Given that existing non-targeted standard of care therapies are associated with toxicity, intolerance and quality of life impairment (see section B.1.3.1.2), these data indicate …*

C11. Please provide the footnotes missing from Appendix Table 1.

Amgen response: Apologies for the missing footnote. Appendix Table 1 with the footnote inserted is provided below:

Appendix Table 1 from CS appendix D: Eligibility criteria for the SLR of RCTs

Criteria	Include	Exclude
Population	Adults (18 years) with KRAS mutated locally advanced and unresectable or metastatic (stage IIIB-IV) NSCLC who had received at least 1 prior systemic therapy. Studies with non-adult participants if information specific to adults was reported separately¶. Subgroups of particular interest including but not limited to: • PD-L1 expression • Prior PD-(L)1 therapies • Early vs late progressors	• Paediatric and adolescent (<18 years) patients • Patients with cancers other than NSCLC • Early-stage NSCLC patients (Stage<IIIB) • Trials studying safety and efficacy of treatment administered in adjuvant setting • Treatment naïve patients
Intervention/comparator	Any therapies licensed in the United States or European Union for the second or later line treatment of patients with NSCLC	Treatments specifically targeting EGFR/ALK or ROS 1 mutations Or other targetable mutation

Criteria	Include	Exclude
	Retreatment with Immuno-oncology therapies will be considered as is in scope even if not a specified retreatment post progression on an anti PD-(L)1	
Outcomes	**Efficacy** • Overall survival (OS)§ • Progression-free survival (PFS)§ • Progression after next line of therapy (PFS2) § • Time to progression (TTP)§ • Time to next treatment (TTNT) • Event-free survival§ • Objective response rate (ORR) • Partial response (PR) • Complete response (CR) • Odds ratio for response rates • Duration of response • Disease control rate or clinical benefit rate • Treatment duration and dosing (median duration, mean duration, mean number of doses, cumulative doses, etc.) **Safety and tolerability:** • All-grade treatment-emergent AEs • Treatment related Grade 3 or 4 AEs • Treatment related SAEs • Tolerability: dose reductions and interruptions, discontinuation (any reason), discontinuation (due to AEs)	Non-clinical outcomes
Study design/setting	Phase II – IV randomised controlled trials	• Trials with a phase I component only • Non-randomized clinical trials • Studies with <10 participants
Language of publication	English language publications (English language abstracts of foreign language publications will be considered for inclusion.)	Non-English language publications without an English abstract.
Date of publication	For the replicated searches, full papers published during or after 2015‡ Conference abstracts published during or after 2017‡	• Studies published prior to 2015‡ • Conference abstracts published prior to 2017
Countries	No restriction	-

Abbreviations: AEs, adverse events; CI, confidence interval; HR, hazard ratio; mOS, Median OS; NSCLC, non-small cell lung cancer, ORR, objective response rate; OS, overall survival; PFS, progression-free response; PR, partial response; SAE, serious adverse events; TTP, time to progression.
† Studies were included even if the investigational dose was not the approved treatment
‡ Studies published prior to 2017 were identified in a previous systematic literature review by Schulz et al. 2019 and will be automatically included in the current review.
§ For time to event endpoints, data will be captured as median, hazard ratios (HR) with 95% confidence intervals (CI) and at rates at landmark timepoints (12, 24, 36, 48, 60 months)

1. March 2021 data cut updates

The submission used the 01 December 2020 data cut to inform the clinical sections (B2) and inputs into the model that produced the modelling results (B3) presented in the submission. There is no updated CSR with information beyond what was presented in the 4th June publication and so the efficacy results presented for comparison in Appendix E (and those presented in the Skoulidis et al. publication) remain the most recent. There is also no update to the (September 2020) PRO data presented in the submission.

Individual patient-level data is now available for the 15th March data cut and so updates to the relevant analyses were made so as to inform the following modelling inputs:

a) OS & PFS
    i) Unadjusted CodeBreaK 100 parameters
    ii) Unadjusted HR CodeBreaK 100 vs. SELECT-1
    iii) CodeBreaK 100 parameters MAIC-adjusted to match SELECT-1
        (1) Includes 2 MAIC models:
            (a) Pre-specified important variables
            (b) All available variables
    iv) MAIC-adjusted HR of CodeBreaK 100 vs. SELECT-1 docetaxel
        (1) Includes 2 MAIC models:
            (a) Pre-specified important variables
            (b) All available variables
    v) HR of CodeBreaK 100 vs. Flatiron
    vi) CodeBreaK 100 vs. Flatiron parameters
b) TTD
    i) CodeBreaK 100 HR of TTD vs. PFS for sotorasib
    ii) Unadjusted TTD parameters from CodeBreaK 100
    iii) MAIC-adjusted TTD parameters from CodeBreaK 100
        (1) Includes 2 MAIC models:
            (a) Pre-specified important variables
            (b) All available variables
c) Incidence of AEs for sotorasib from CodeBreaK 100
d) Sotorasib RDI from CodeBreaK 100
e) Utilities
    i) By health state (progression status)
    ii) By time to death category

In the sections below, the updates to each of these analyses is presented in the order presented in the original submission. Updated results of the cost-effectiveness model are also shown.

2. Indirect Comparisons

Results of the indirect treatment comparisons

Primary comparison – sotorasib vs docetaxel monotherapy

Results of primary analysis – MAIC using CodeBreaK100 and SELECT-1

Results for the primary MAIC analysis, the sensitivity analysis using all available covariates, and an unadjusted analysis provided for reference, are reported in terms of hazard ratios (HRs) with 95% confidence intervals (CIs) based on robust standard errors (Table 1). Point estimates related to ESS, median OS, median PF and related HRs remain similar. Median OS for both MAIC models have now been reached and show a substantial difference compared with docetaxel.

Table 1. Results of MAIC for primary comparison of sotorasib vs docetaxel monotherapy

Analyses	**CodeBreaK 100 N (OS / PFS)**	**CodeBreaK 100 ESS (OS / PFS)**	**Median OS Sotorasib vs Docetaxel**	**Median PFS Sotorasib vs Docetaxel**
Unadjusted	126	126	xxxx	xxxx
MAIC Model: "all variables of prognostic importance" **(Primary analysis)**	**123/ 121**	**108.9/ 106.1**	xxxx	xxxx
MAIC Model: "all available covariates" (sensitivity analysis)	98/ 96	53.3/ 53.3	xxxx	xxxx

Key: *Median OS in this data cut has now been reached for both models; CI, confidence interval; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; mths, months; OS, overall survival; PFS, progression-free survival

Figure 1. Kaplan-Meier plot for primary MAIC analysis of OS for sotorasib and docetaxel monotherapy

Figure 2. Kaplan-Meier plot for primary MAIC analysis of PFS for sotorasib and docetaxel monotherapy

Results of the supplementary primary comparison – Propensity score weighting analysis using CodeBreaK100 and Amgen Flatiron Health real-world evidence study

Median and HR Point estimate results remain stable after implementation of the data cut for both the KRAS and subgroup with *KRAS p.G12C*-mutated NSCLC. The point estimates of the OS and PFS hazard ratios clearly favour sotorasib numerically and suggest it is possible that the hazard ratio estimates based on the wider KRAS mutant population may be conservative; however, the effective sample size is very small.

Table 2. Results of supplementary primary comparison using propensity score weighting analysis

Outcome	Flatiron N before adjustment	*KRAS* mutant		*KRAS-p.G12Cmutant* subgroup	
		ESS	Median HR (95% CI)	ESS	Median HR (95% CI)
Overall survival	206	104.8	xxxx	17.8	xxxx
Progression-free survival	206	104.8	xxxx	17.8	xxxx

Key: ATT, average treatment effect of the treated; CI, confidence interval; ESS, effective sample size of Flatiron population following adjustment to CodeBreaK100 population; HR, hazard ratio; mths, months; OS, overall survival

3. Clinical parameters and variables

Sotorasib versus docetaxel

Overall survival

Statistical Goodness of Fit

Goodness of fit statistics remain similar to those presented in the main submission (Table 3). For individually fitted curves, the lognormal distribution was the best statistically fitting curve with the second lowest AIC (but not significantly lower than generalised Gamma) and BIC across both sotorasib and docetaxel.

For the jointly fitted curves, AIC and BIC indicate that the best fitting curve for both the restricted and unrestricted models was the lognormal followed by the generalised gamma and log-logistic models.

Table 3: Goodness of fit statistics for independent and jointly fitted models

Model	Independent fit – sotorasib		Independent fit - docetaxel		Joint fit (unrestricted)		Joint fit (restricted)	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	454.3	457.1	1209.7	1213.2	1663.9	1671.8	1663.9	1671.8
Gompertz	456.2	461.9	1211.4	1218.5	1667.7	1683.4	1665.8	1677.6
Weibull	454.1	459.8	1209.6	1216.7	1663.8	1679.5	1662.2	1674.0
Generalized Gamma	446.9	455.3	1194.6	1205.2	1641.5	1665.1	1639.3	1655.0
Loglogistic	450.6	456.2	1196.3	1203.4	1646.9	1662.6	1645.0	1656.8
Lognormal	447.4	453.0	1192.8	1199.9	1640.2	1656.0	1638.2	1650.0

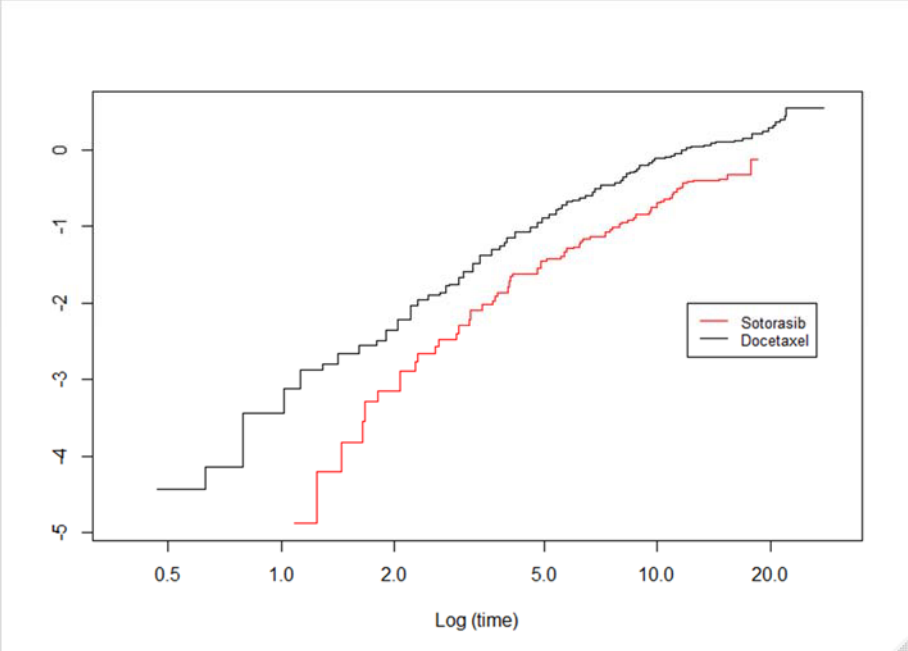
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion
Note: Underlined values indicate the best statistically fitting parametric distribution

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

To further inform the most appropriate distribution to extrapolate OS, the proportional hazards assumption and the presence of accelerated failure time was assessed using log

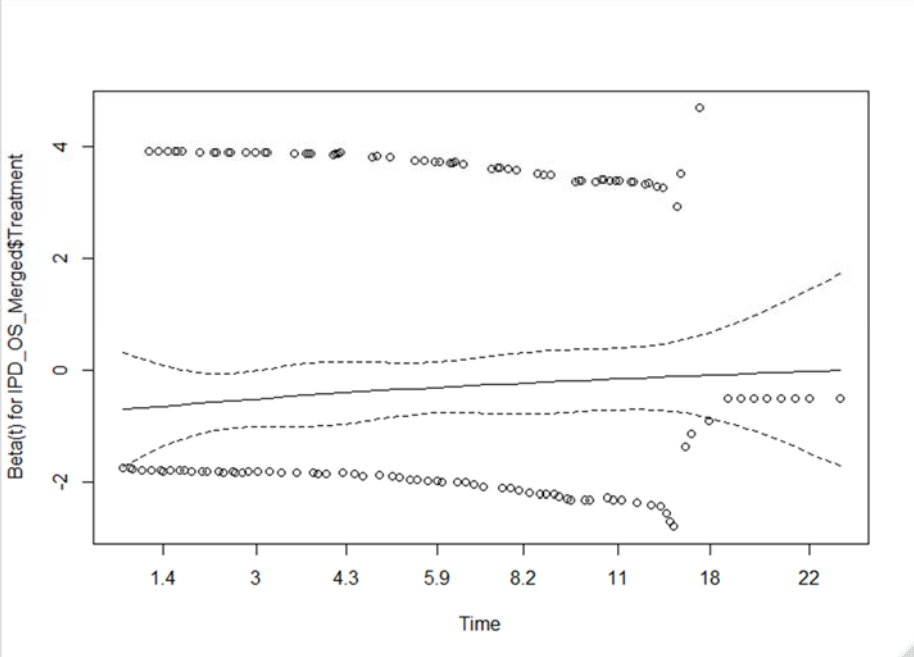
cumulative plots, Schoenfeld residuals and QQ plots. Results remain consistent with the descriptions provided in the submission.

Figure 3: Log-cumulative hazards plot for OS using base case MAIC



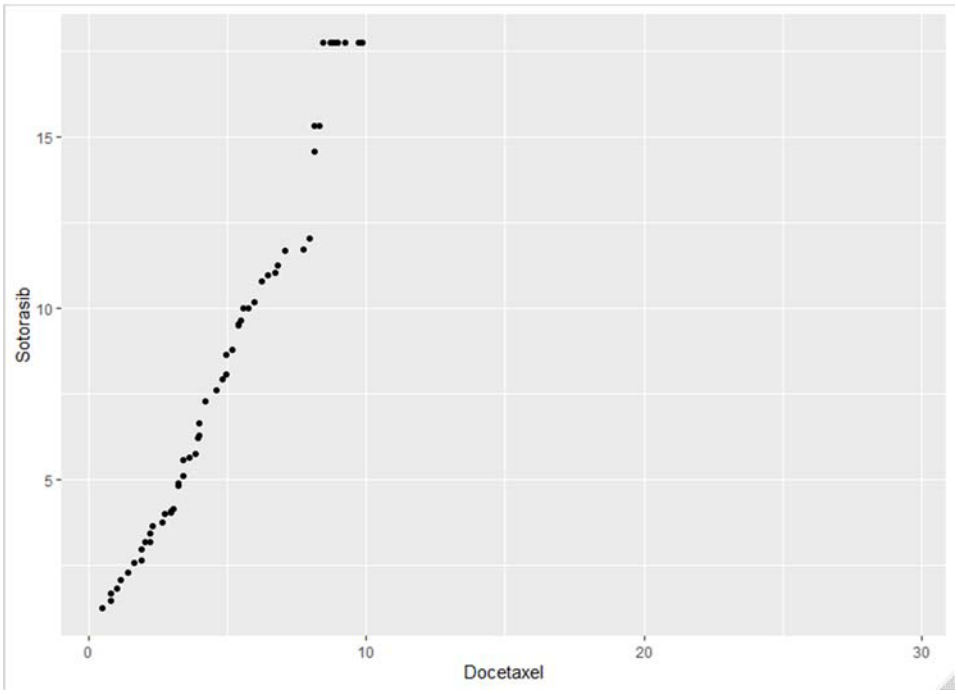
Key: MAIC, matching-adjusted indirect comparison; OS, overall survival.

Figure 4: Schoenfeld residuals plot for OS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; OS, overall survival.

Figure 5: Q-Q plot for OS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; OS, overall survival; Q-Q, quartile-quartile.

Visual Inspection to Observed Data

A plot of jointly fitted parametric distributions fitted to sotorasib and docetaxel using the base case MAIC adjusted Kaplan–Meier curves is shown below (Figure 6). The restricted joint models are presented given the superior statistical fits observed previously.

Results remain almost identical with the docetaxel plot indicating that the Weibull, Gompertz and exponential distributions overestimated OS in the early periods (up to 14 months) with consistent underestimation of OS after this timepoint (Figure 6). Similarly, the sotorasib plot indicated that the Weibull and Gompertz plots underestimated OS up to 2 months and were the most conservative (with exponential) OS estimates for the long-term projections (Figure 6). These findings are consistent with the AIC and BIC results previously presented.

Visual inspection of the best statistically fitting distributions generally indicated that the extrapolated data matched the Kaplan–Meier plots well and captured the longer term shape of the survival function (Figure 6). In all cases sotorasib was shown to improve long-term OS versus docetaxel.

Figure 6: OS KM for sotorasib and docetaxel from base case MAIC with parametric functions fitted

Key: KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; OS, overall survival.

Progression-free survival

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 4). The ordering as presented in the main submission holds with the updated data cut. For individually fitted curves, the AIC and BIC both indicated that the lognormal distribution provided the best statistical fit for sotorasib, whereas the generalised gamma performed the best for docetaxel.

For the jointly fitted models, the AIC indicates that the generalised gamma distribution is the best performing, whereas the BIC indicates that the lognormal provides the best statistical fit to the observed data, although again differences are minor.

Table 4: Goodness of fit statistics for independent and jointly fitted models

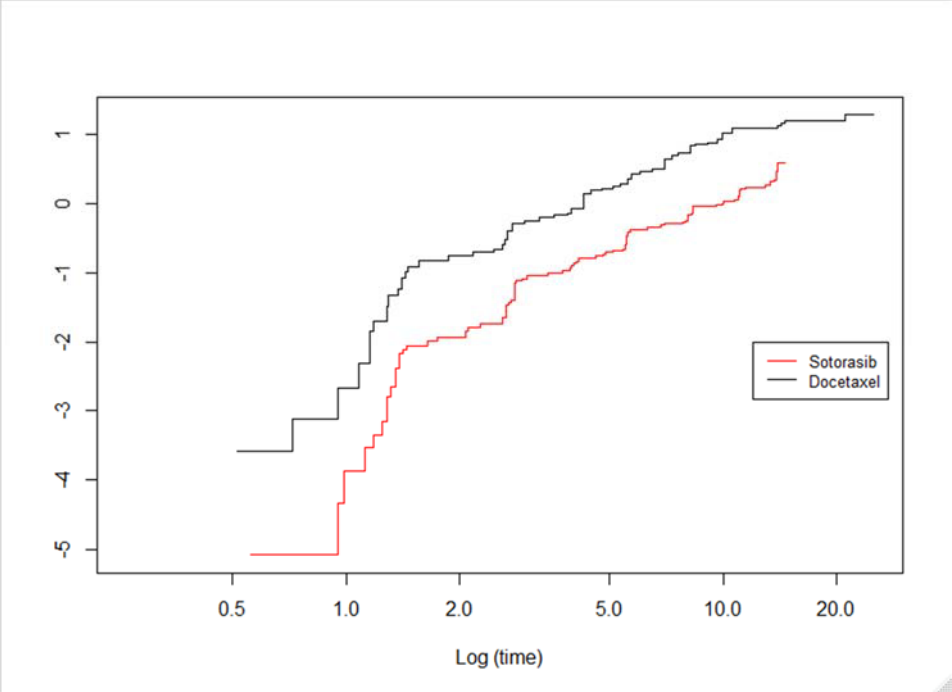
Model	**Independent fit – sotorasib**		**Independent fit - docetaxel**		**Joint fit (unrestricted)**		**Joint fit (restricted)**	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	562.5	565.3	1166.5	1170.0	1729.0	1736.9	1729.0	1736.9
Gompertz	561.8	567.4	1166.9	1174.0	1728.6	1744.4	1730.9	1742.7
Weibull	558.4	564.0	1160.6	1167.7	1718.9	1734.7	1717.8	1729.6
Generalized Gamma	554.3	562.7	1099.5	1110.1	1653.8	1677.4	1655.3	1671.1
Loglogistic	556.5	562.1	1113.5	1120.6	1670.0	1685.7	1670.1	1682.0
Lognormal	552.4	558.0	1105.7	1112.8	1660.2	1675.9	1658.2	1670.0

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion
Note: Underlined values indicate the best statistically fitting parametric distribution

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

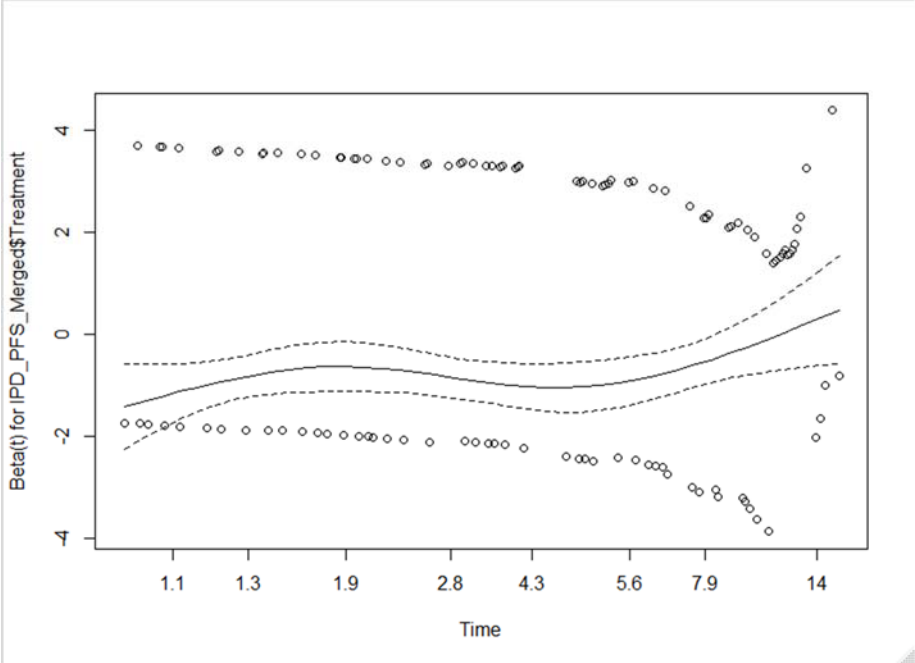
Similar to OS, the assumption of proportional hazards between the two datasets was assessed using the log-cumulative hazards plot (Figure 7) and the Schoenfeld residuals plot (Figure 8). As in the company submission, the log-cumulative hazards and the Schoenfeld residuals plot for sotorasib and docetaxel indicated that the proportional hazards assumption is unlikely to hold: the log-cumulative hazards plot demonstrated the convergence of both curves in the first 2 months, which diverges before 3 months and then remains parallel beyond 4 months. Likewise, the confidence bands of the scaled Schoenfeld residuals did not include zero for the majority of the time horizon.

Figure 7: Log-cumulative hazards plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival.

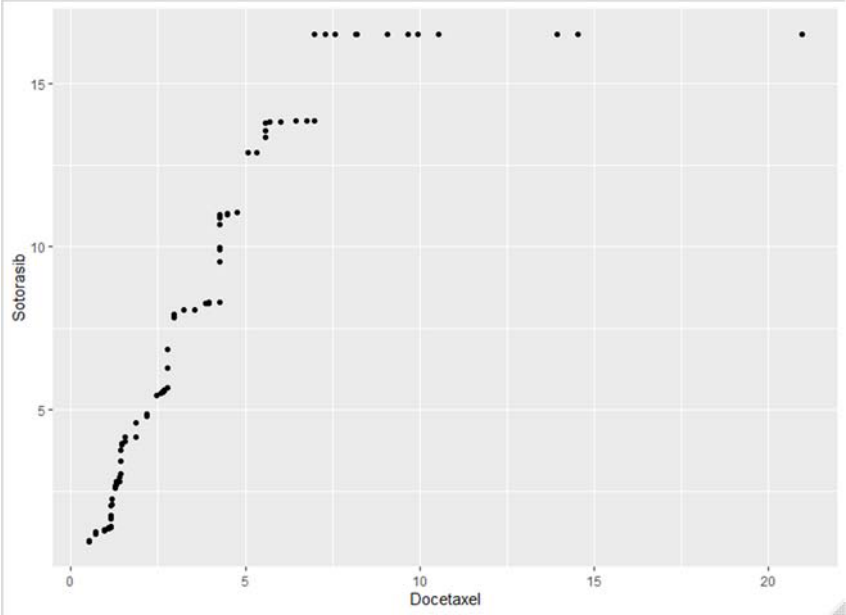
Figure 8: Schoenfeld residuals plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival.

The Q-Q plot, however, indicated that accelerated failure time assumption was sufficiently valid with a straight-line trend clearly discernible in the main (Figure 9).

Figure 9: Q-Q plot for PFS using base case MAIC



Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; Q-Q, quartile-quartile.

Visual Inspection to Observed Data

A plot of jointly fitted parametric distributions fitted to sotorasib and docetaxel using the base case MAIC adjusted Kaplan–Meier curves is shown below (Figure 10). Visual inspection of the docetaxel plot indicates similar conclusions to the main submission: the Weibull, Gompertz and Exponential distributions overestimated PFS in the early periods (up to 12 months) with underestimation of PFS after this timepoint (Figure 10). Both the lognormal and log-logistic model fit the data well and the generalised gamma, although performing well on statistical tests, shows a slight underestimation for docetaxel between 6 and 12 months.

Figure 10: PFS KM for sotorasib and docetaxel from base case MAIC with parametric functions fitted

Key: KM, Kaplan-Meier; MAIC, matching-adjusted indirect comparison; PFS, Progression Free Survival

Visual inspection of the best statistically fitting distributions indicated that the extrapolated data matched the Kaplan–Meier plots well and captured the longer-term shape of the survival function (Figure 10). In all cases sotorasib was shown to improve long-term PFS versus docetaxel.

Alternative comparative efficacy based on Flatiron real-world data

Overall survival

The weighted KM plot for OS from the propensity score weighting analysis is presented in Figure 11. Results are visually consistent with those in the main submission.

Figure 11. Kaplan-Meier plot of OS from the propensity score weighting analysis

Statistical Goodness of Fit

Updated goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 5). In line with the analysis for with the December data cut: for individually fitted curves, the lognormal distribution was the best statistically fitting curve with the lowest AIC and BIC across both sotorasib and chemotherapy, with the exception of the BIC for chemotherapy which marginally favoured the exponential distribution. Consistent with the approach taken previously, jointly fitted survival models (either restricted or unrestricted) were considered the most appropriate for this analysis.

Table 5: Goodness of fit statistics for jointly fitted OS models for KRAS mutant ATT

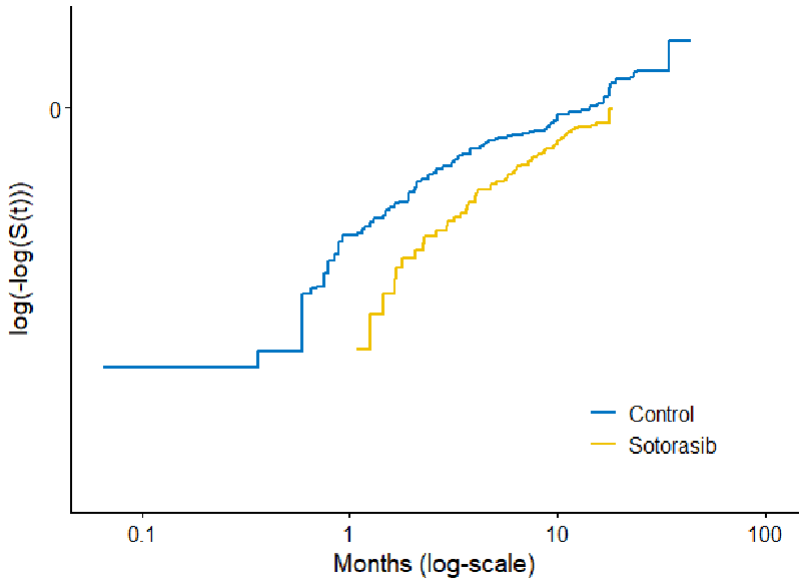
Model	**Independent fit – sotorasib**		**Independent fit - chemo**		**Joint fit (unrestricted)**		**Joint fit (restricted)**	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	479.1	481.9	605.1	608.4	1084.2	1091.7	1084.2	1091.7
Gompertz	481.0	486.5	605.7	612.4	1086.7	1101.8	1085.5	1096.8
Weibull	478.5	484.0	606.6	613.2	1085.0	1100.2	1086.1	1097.4
Generalized Gamma	472.2	480.1	604.1	614.1	1076.4	1099.1	1075.8	1090.9
Loglogistic	474.9	480.5	604.6	611.3	1079.5	1094.7	1078.9	1090.3
Lognormal	472.0	477.54	602.4	609.0	1074.4	1089.5	1073.9	1085.2

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion
Note: Underlined values indicate the best statistically fitting parametric distribution

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

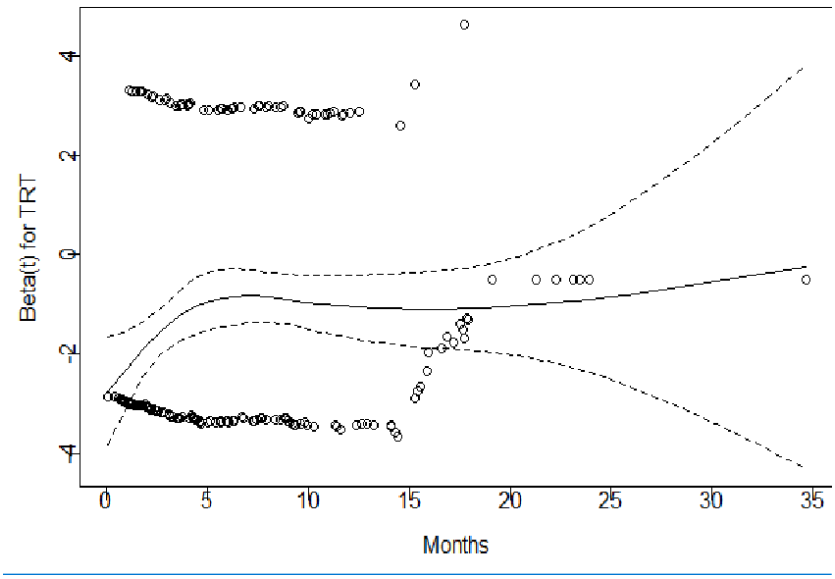
The proportional hazards assumption was evaluated for OS using the log cumulative hazards plot (Figure 12) and the Schoenfeld residuals plot (Figure 13). In line with the plots in the company submission, these updated plots indicated the proportional hazards assumption was unlikely to be valid. Accelerated time failure for OS was assessed using a Q-Q plot (Figure 14). The plot indicated that despite some deviation either side of the from the fitted line the assumption of accelerated failure time appears reasonable.

Figure 12: OS log cumulative hazards plot for sotorasib and control



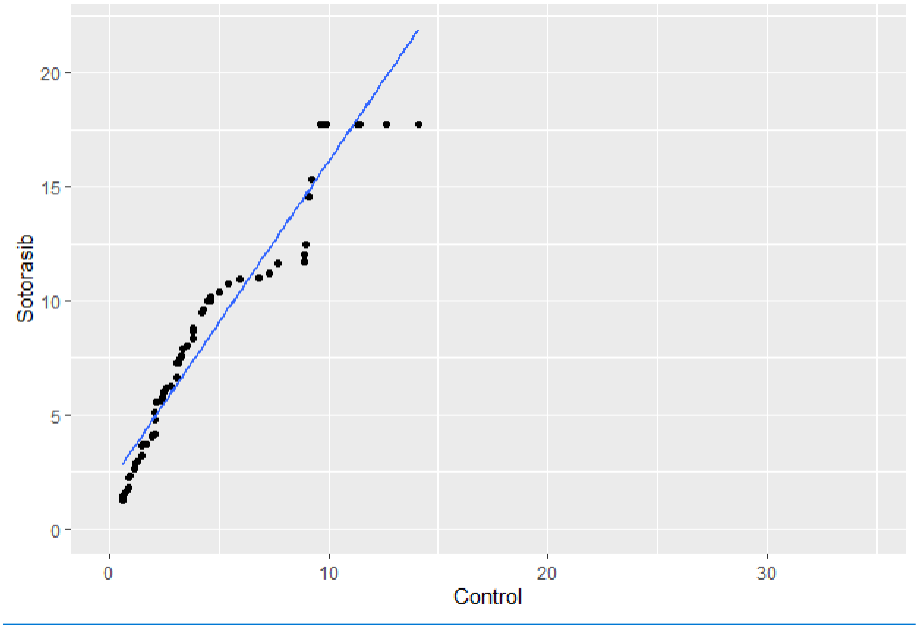
Key: OS, overall survival.

Figure 13: OS Schoenfeld residuals plot for sotorasib and control



Key: OS, overall survival.

Figure 14: OS Q-Q plot for sotorasib and control



Key: OS, overall survival; Q-Q, quartile-quartile.

Distribution Selection

Given that the joint fit (restricted) lognormal provides the best statistical fit to the observed ATT propensity adjusted data and the assumption of accelerated failure time appears to hold, this curve was selected to inform this sensitivity analysis. The (March data cut) updated visual fit of the ATT propensity KM data to the lognormal distribution is presented in Figure 15.

Figure 15: ATT OS KM versus fitted lognormal model using restricted model

Key: ATT, average treatment effect of the treated ; KM, Kaplan–Meier, OS, overall survival.

Progression-free survival

The weighted KM plot for PFS from the propensity score analysis, with updated March data cut IPD, is presented in Figure 16.

Figure 16. Kaplan-Meier plot of PFS from the propensity score weighting analysis

Statistical Goodness of Fit

Goodness of fit statistics using AIC and BIC are presented for jointly fitted (unrestricted and restricted) and independent models (Table 6). As before, the lognormal distribution was the consistently best statistically fitting curve with the lowest AIC and BIC across both sotorasib and chemotherapy. As a result, jointly fitted survival models (either restricted or unrestricted) were considered the most appropriate since they can reduce uncertainty due to the estimation of fewer parameters and the use of a larger data set.

Table 6: Goodness of fit statistics for jointly fitted PFS models for KRAS-mutant ATT

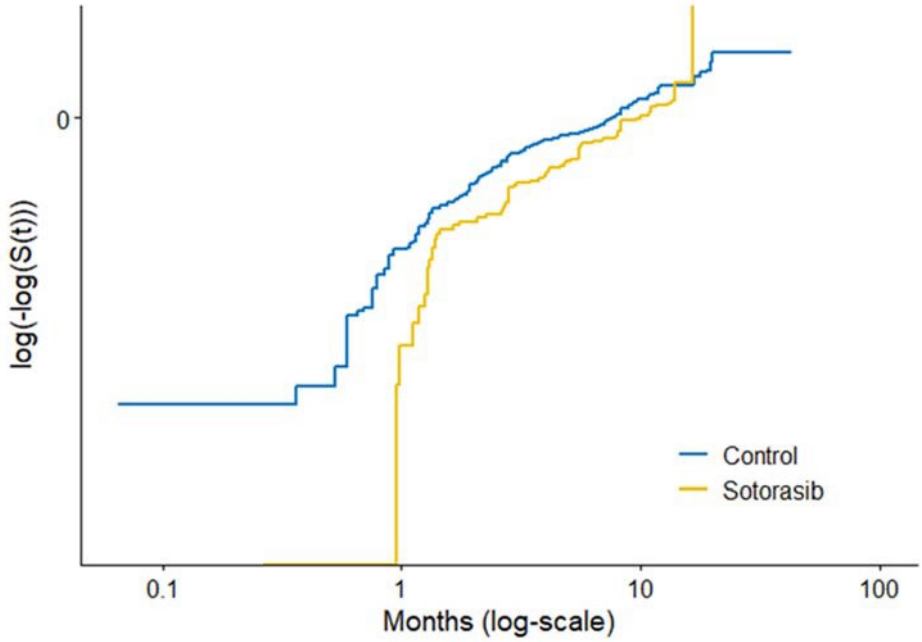
Distribution	**Independent fit – sotorasib**		**Independent fit - chemotherapy**		**Joint fit (unrestricted)**		**Joint fit (restricted)**	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	564.5	567.3	601.5	604.8	1166.0	1173.6	1166.0	1173.6
Gompertz	562.8	568.4	603.0	609.7	1165.8	1180.9	1167.7	1179.0
Weibull	559.5	565.1	603.4	610.1	1162.9	1178.0	1164.5	1175.8
Generalized Gamma	556.7	565.1	597.7	607.6	1154.4	1177.1	1152.3	1167.5
Loglogistic	558.8	564.4	598.8	605.4	1157.6	1172.7	1156.8	1168.1
Lognormal	554.7	560.3	595.9	602.5	1150.6	1165.7	1150.4	1161.8

Key: ATT, average treatment effect of the treated; AIC, Akaike information criterion; BIC, Bayesian information criterion; PFS, progression-free survival.
Note: Underlined values indicate the best statistically-fitting parametric distribution.

Diagnostics (Proportional Hazards, Shoenfeld residuals and QQ Plots)

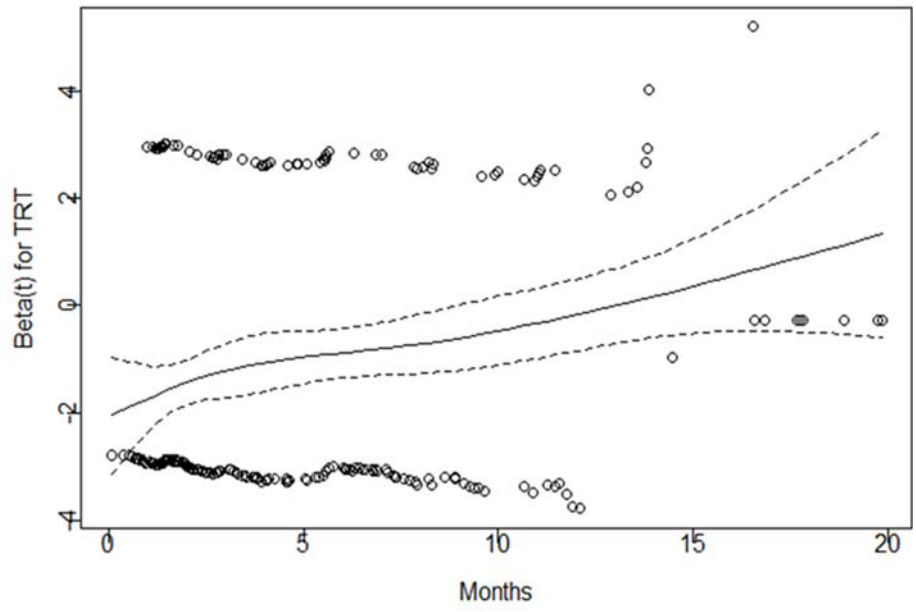
The proportional hazards assumption was evaluated for PFS using the log cumulative hazards plot (Figure 17) and the Schoenfeld residuals plot (Figure 18). These plots, as before, indicated the proportional hazards assumption was not valid. Accelerated time failure for PFS was assessed using a Q-Q plot (Figure 19). The plot again indicated some deviation either side of the from the fitted line. However, overall the assumption of an accelerated failure time model appeared acceptable.

Figure 17: PFS log cumulative hazards plot for sotorasib and control



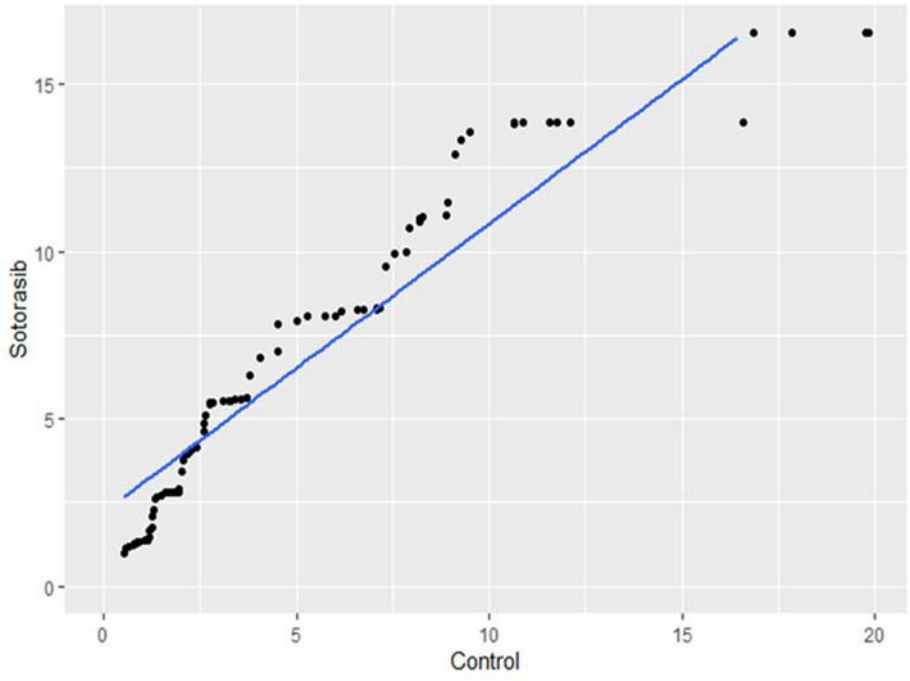
Key: PFS, progression-free survival.

Figure 18: PFS Schoenfeld residuals plot for sotorasib and control



Key: PFS, progression-free survival.

Figure 19: PFS Q-Q plot for sotorasib and control



Key: PFS, progression-free survival; Q-Q, quartile-quartile.

Distribution Selection

Given that the joint fit (restricted) lognormal provides the best statistical fit to the observed ATT propensity adjusted data and the assumption of accelerated failure time appears to

hold, this curve can again be used to inform this sensitivity analysis. The visual fit of the ATT propensity KM data to the lognormal distribution is presented in Figure 20 and as before shows a close visual fit of the extrapolation to the Kaplan–Meier data.

Figure 20: ATT PFS KM versus fitted lognormal model using restricted model

Key: ATT, average treatment effect of the treated; KM, Kaplan–Meier, PFS, progression-free survival.

Nintedanib + docetaxel

The analysis related to and inputs (i.e. piecewise hazard ratios) for this comparison have not been updated because they are not based on CodeBreaK100 trial data.

Treatment duration

Sotorasib

Sotorasib treatment duration was modelled using a hazard ratio applied to PFS from CodeBreaK100. The hazard ratio was estimated using a Cox model with the effect estimated between time to treatment discontinuation and progression-free survival (updated march data cut xxxx, 95% CI: xxxx).

Adverse events

Point estimates for adverse event incidence for the March data cut are the same as for the December data cut and so Table 32 in the main submission remains valid.

4. RDI and utilities

Relative dose intensity (RDI)

The updated RDI for Sotorasib is 89.0% and so is slightly lower than that reported in the company submission (89.2%).

Utilities

The health state utilities analysis gives similar results based on the March data cut, with a slightly higher PF and PP utility (Table 7). The time to death utilities analysis also gives similar point estimate results to the company submission (Table 8).

Table 7: Summary of health state utility values

Health state	Mean (95% CI)	Reference
Progression-free	0.734 (0.700, 0.769)	CodeBreaK100 EQ-5D-5L analyses [a] [57] using UK crosswalk tariffs [91]
Disutility in progressed disease	0.064 (0.097, 0.031)	
Post-progression	0.670	Calculation
Key: CI, confidence interval. **Note:** [a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-4.7.702 and subsequent analyses		

Table 8: Time to death utilities

Health state	Mean (95% CI)	Reference
Utility more than 6 months to death	0.762 (0.698, 0.767)	CodeBreaK100 EQ-5D-5L analyses using UK crosswalk tariffs (1 September 2020 data cut-off)[91]
Disutility between 3 and 6 months to death (versus. more than 6 months)	0.047 (0.090, 0.004)	
Disutility between 1 and 3 months to death (versus. more than 6 months)	0.125 (0.176, 0.074)	
Disutility less than 1 month to death (versus. more than 6 months)	0.233 (0.312, 0.153)	
Utility between 3 and 6 months to death	0.715	Calculated
Utility between 1 and 3 months to death	0.637	Calculated
Utility in last month of life	0.529	Calculated
Key: CI, confidence interval.		

5. Updated deterministic results and scenario analyses

Base-case results

Sotorasib versus docetaxel (primary comparator)

In the model base case where docetaxel is considered the comparator, discounted results are presented in Table 9. Using a 20-year time horizon, results remain similar but with an increase in incremental costs and a disproportionately higher increase in incremental QALYs which reduces the ICER by around 6%.

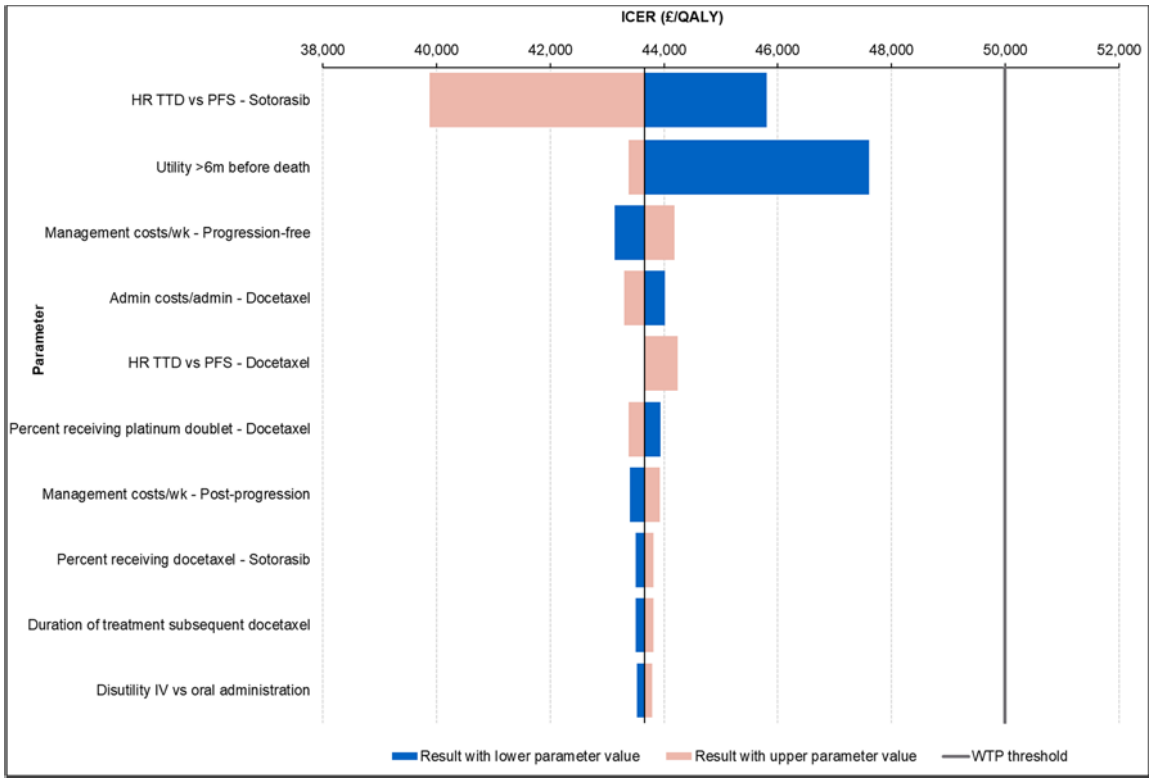
Table 9: Deterministic base-case results: sotorasib versus docetaxel

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Docetaxel	[redacted]	[redacted]	[redacted]				
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**43,660**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALYs, quality-adjusted life years.							

Deterministic sensitivity analysis

Figure 21 presents the updated tornado diagram for the top ten parameters in terms of ICER impact which were varied in the OWSA. Parameters are shown in descending order of ICER sensitivity. The ordering of sensitivity to input variables remains the same.

Figure 21: One-way sensitivity analysis for sotorasib versus docetaxel



Key: HR, hazard ratio; PFS, progression-free survival; TTD, time to treatment discontinuation.

Scenario analysis

Key scenarios

A comparison of sotorasib to docetaxel based on the Flatiron real world dataset are presented in Table 10. Results are again more favourable for Sotorasib compared with the base-case but are broadly comparable to the results with the December data cut.

Table 10: Deterministic results: sotorasib versus docetaxel using Flatiron data

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Docetaxel	xxxx	xxxx	xxxx				
Sotorasib	xxxx	xxxx	xxxx	[redacted]	[redacted]	[redacted]	**38,279**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALYs, quality-adjusted life years							

The secondary comparison of sotorasib vs nintedanib plus docetaxel, assuming the [redacted] xxxx for sotorasib and the list price for nintedanib are presented in Table 11. Results are again similar with a disproportionally large increase in incremental QALYs relative to costs leading to a slight decrease in the ICER.

Table 11: Deterministic results: sotorasib versus nintedanib + docetaxel

Technology	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Nintedanib + docetaxel	xxxx	xxxx	xxxx				
Sotorasib	xxxx	xxxx	xxxx	[redacted]	[redacted]	[redacted]	**33,628**
Key: ICER, incremental cost-effectiveness ratio; Incr., incremental; LYG, life years gained; QALY, quality-adjusted life year.							

Other scenarios

The same extensive range of scenario analyses exploring alternative comparative effectiveness estimates, as well as cost and resource inputs, HRQoL and model settings were conducted with the update to the new March data cut. As seen in Table 12, the results remain consistent with the company submission in terms of direction and magnitude of impact on the ICER.

Table 12 Scenario analysis results

Scenario	Rationale/Justification	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
15-year time horizon	To investigate the impact on model results of reducing the model timeframe.	[redacted]	[redacted]	44,779
Generalised gamma distribution selected to estimate long-term OS and PFS projections	The generalised gamma was the 2nd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more favourable estimation of survival.	[redacted]	[redacted]	42,992
Log logistic distribution selected to estimate long-term OS and PFS projections	The log-logistic was the 3rd best-fitting distribution based on the survival analysis selection criteria outlined in Section B.3.3. This scenario provides a more conservative estimation of survival.	[redacted]	[redacted]	49,117
Joint (unrestricted) lognormal distribution selected to estimate long-term PFS	Although BIC consistently favoured the restricted versus unrestricted joint fit this scenario tests an alternative parametric distribution using the unrestricted lognormal for PFS which was the best fitting unrestricted model based on BIC criteria.	[redacted]	[redacted]	47,917
Adjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1 using all available covariates	To test the robustness of the MAIC using alternative MAIC model where all available covariates are considered. A joint (restricted) lognormal distribution was used per the base case analysis and based on the survival analysis selection criteria outlined in Appendix N.	[redacted]	[redacted]	37,082
Unadjusted sotorasib from CodeBreaK100 vs. unadjusted docetaxel from SELECT-1	To investigate the impact of a naïve comparison to the SELECT-1 clinical trial. A joint (restricted) lognormal distribution was used per the base case analysis.	[redacted]	[redacted]	50,981
Unadjusted sotorasib from CodeBreaK100 vs. ATT-adjusted docetaxel from Flatiron	Alternative data source which included patients closely aligned with the CodeBreaK100 population in terms of prior treatment from the real-world	[redacted]	[redacted]	38,279

	Flatiron dataset to test the robustness of the results in the base case analysis			
MAIC-adjusted TTD curve from CodeBreaK100	To test the impact of an alternative approach to estimate long-term treatment duration.	[redacted]	[redacted]	44,496
HR of sotorasib vs. docetaxel = 1 after 5 years	In the base-case PFS and OS were modelled based on parametric survival distributions fit to survival data from CodeBreaK100 and SELECT-1, combined with age- and sex-matched general population mortality. This scenario explicitly limits the duration of benefit to 60 months.	[redacted]	[redacted]	46,587
Apply health state utilities by progression status	To test the impact of an alternative method for measuring health state utilities as described in Section B.3.4.5	[redacted]	[redacted]	47,208
Treatment-emergent AEs	To test the impact of utilising treatment-emergent adverse events for sotorasib and docetaxel as described in Section B.3.3.7	[redacted]	[redacted]	44,116
Include drug wastage	To test the impact of potential drug wastage in clinical practice by estimating drug acquisition costs based on total packs as opposed to treatments received	[redacted]	[redacted]	46,387
Exclude RDI	To test the impact of not capturing RDI on drug utilisation calculations	[redacted]	[redacted]	48,944
1.5% discount rate for costs and efficacy	To investigate the alternative discount rate suggested by the NICE Guide to Technology Appraisal. A reduced discount rate of 1.5% is consistent with the Treasury Green Book and is being considered in the ongoing NICE Methods Review consultation.	[redacted]	[redacted]	41,120
Key: ATT, average treatment effect of the treated; HR, hazard ratio; OS, overall survival, ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALY, quality-adjusted life year; RDI, relative dose intensity.				

6. Conclusions related to March data cut updates

As shown in previous sections, MAIC results and diagnostics related to curve selection (fit statistics and plots) remain stable across data cuts with the broad conclusions made in the submission holding. Results are also similar, with a slight improvement in the relative cost-effectiveness of sotorasib.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

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.

About you	
1.Your name	[redacted]
2. Name of organisation	Roy Castle Lung Cancer Foundation
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Roy Castle Lung Cancer Foundation is a UK wide lung cancer charity. We fund lung cancer research and work in lung cancer patient care (information, support and advocacy activity). Our funding base is a broad mixture including community, retail, corporate, legacies and charitable trusts. Clearly, our patient group members and contacts are a self-selected group, who have taken the step to seek out information or have accessed specialist support services. As most lung cancer sufferers tend to be older, from lower social class groups and with the five year survival being around 15%, less physically well, we acknowledge that our patients are perhaps not representative of the vast majority of lung cancer patients, who are not so well informed. It is, however, important that the opinions expressed to us, be passed on to NICE, as it considers the place of this product in the management of lung cancer
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?	As a result of the COVID pandemic, our contact with patients and carers has become virtual. The Foundation has contact with patients/carers through its UK wide network of Lung Cancer Patient Support Groups, patient/carer panel, online forums, Keep in Touch' service and its nurse-led Lung Cancer Information Helpline.
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?	According to the National Lung Cancer Audit, the one year survival for lung cancer is 37%. Thus, this group of lung cancer patients have a particularly poor outlook. with an obvious impact on family and carers. Symptoms such as breathlessness, cough and weight loss are difficult to treat, without active anti-cancer therapy. Furthermore, these are symptoms which can be distressing for loved ones to observe. Approximately 13% of patients with non small cell lung cancer (nsclc) have the KRAS G12C mutation. Historically, with no drugs available to target KRAS, these patients have had a poorer prognosis.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	In recent years, we have seen new targeted therapy options for some patients with nsclc. As above, this has, so far, not been the case for those with KRAS G12C mutations. There are currently no NICE recommended treatments, specifically for G12C mutation positive lung cancer patients. Current systemic treatment (first and second line treatment) would be with standard NSCLC treatment – a combination of chemotherapy and immunotherapy.

8. Is there an unmet need for patients with this condition?	Yes
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	As above, this would be the first NICE approved therapy available specifically targeted at G12C mutation positive lung cancer. We refer to the multicentre, single arm study, Code Break 100 clinical trial. The 124 study participants had KRAS G12C mutation positive disease and had received one previous systemic treatment. Data showed that Sotorasib shrank the tumours of 36% of participants and response lasted for a median of 10 months. Sotorasib is a once a day, oral treatment (tablet), with the obvious advantages of home/ease of administration, reduction in patient time at hospital (important in this new COVID world) etc..
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	The side effects associated with the therapy. Most commonly reported were diarrhoea, musculoskeletal pain and nausea. Most side effects were mild, however, 20% experienced more serious side effects. Of note, 28% had treatment delays and / or treatment dose reductions due to side effects and 7% stopped treatment,

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.	
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	

Other issues	
13. Are there any other issues that you would like the committee to consider?	We understand that further studies of Sotorasib are ongoing in several clinical trials. As data matures and as new data emerges, this is perhaps a therapy, at this time, which could be made available through the Cancer Drugs Fund.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- First targeted therapy being assessed specifically for KRAS G12C mutation positive nsclc.
- Oral treatment
- Consider availability through the Cancer Drugs Fund, reassessing after data matures and new data emerges.
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

NICE National Institute for Health and Care Excellence

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

Professional organisation submission

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

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	
1. Your name	[redacted]
2. Name of organisation	**British Thoracic Oncology Group**
3. Job title or position	**Consultant in Medical Oncology; Professor of Experimental Cancer Medicine**

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 this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	BTOG is a not-for-profit association (with charitable status) of clinicians and healthcare professionals interested in the management of thoracic cancers including non-small cell lung cancer (NSCLC)
4b. 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.	Not to my knowledge

5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
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.)	Palliative: To improve symptoms of lung cancer To delay progression and death
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.)	Docetaxel remains a widely used standard of care in patients progressing after first line treatment. The pivotal trial of this agent showed a response rate (RR) of <10% (Shepherd F et al. J Clin Oncol 2000), associated with very significant toxicity including myelosuppression and alopecia. Single agent docetaxel formed the control arm of a more recent study (Reck, M et al. Lancet Oncol 2014), where RR was only 3%. The experimental arm of this trial was combination docetaxel + nintedanib, with a RR of 4%.
8. In your view, is there an unmet need for patients and	*KRAS* mutations are found in approximately 30% of non-small cell lung cancers with adenocarcinoma histology (by for the most common histological subtype), and of these 50% are G12C mutations (Burns, T et al. J Clin Oncol 2020), which are specifically targeted by sotorasib. The *KRAS* oncogene has previously

healthcare professionals in this condition?	been described as undruggable because of its unusual biological characteristics (it is not a protein kinase unlike the target of most small molecule targeted therapies developed in oncology to date). *KRAS* mutations generally occur mutually exclusively with other genetic drivers for which good therapeutic options already exist (such as *EGFR, ALK, ROS1* and *RET*), and therefore there is a clear unmet need in this large molecularly-defined subgroup.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Even when a *KRAS* mutation is identified, these patients receive first line treatment with platinum-based doublet chemotherapy, an immune checkpoint inhibitor (often pembrolizumab), or often a combination of these. Subsequent treatment will often be as part of a clinical trial where available, given the paucity of other good options, or docetaxel in the absence of alternatives. Docetaxel in combination with the antiangiogenic oral therapy nintedanib is also approved and funded in the UK (Reck, M et al. Lancet Oncol 2014), but is more toxic than docetaxel alone, which is the more widely used option.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Clinicians practicing in the UK most commonly refer to ESMO guidelines in this context (Planchard, D et al. Ann Oncol 2018), but ASCO and NCCN guidance are consulted.
• 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.)	Few options are available as “standard care” in patents previously treated with doublet chemotherapy and a checkpoint inhibitor (together or in sequence), and these are described above.

• What impact would the technology have on the current pathway of care?	All fit patients with an identified G12C *KRAS* mutation would have the opportunity of receiving a relatively well-tolerated personalised therapy in the 2nd and subsequent line setting, with a response rate around 33% (Hong D et al. New Engl J Med 2020).
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	There is no KRAS-directed therapy currently available outside clinical trials.
• How does healthcare resource use differ between the technology and current care?	Oral targeted therapy versus chemotherapy with predictable toxicities associated with myelosuppressive drugs.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Following platinum-based doublet chemotherapy, and administration of a checkpoint inhibitor (combination or sequentially). Almost all patients treated in the Hong et al. 'CodeBreaK100' study had received prior therapy.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Not all centres routinely test tumours for *KRAS* mutations. This can be achieved with a simple PCR test, similar to widely-used *EGFR* testing for alternative NICE-approved therapy, and will also be part of the GLH NGS testing panel, due national implementation in England later this year.

11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. PFS in the lung cohort treated in the 'CodeBreaK100' trial of sotorasib was 6.3 months. PFS was 3.4 months in this setting [in molecularly-unselected patients] for docetaxel + nintedanib, compared to 2.7 months for docetaxel + placebo (Reck, M et al. Lancet Oncol 2014).
• Do you expect the technology to increase length of life more than current care?	Yes, this seems likely. Addition to antiangiogenic therapy to docetaxel in this setting was associated with the above PFS improvement, from 2.7 to 3.4 months. This corresponded to a statistically significantly superior overall survival in adenocarcinoma patients (around 1/3 of whom would be expected to harbour a *KRAS* mutation), with HR=0.83 (Reck, M et al. Lancet Oncol 2014).
• Do you expect the technology to increase health-related quality of life more than current care?	Oral targeted therapies are generally better tolerated than conventional cytotoxic therapies, although not without side effects.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Patient selection according to the presence of *KRAS* G12C mutation in the patients' tumour material, or cfDNA, is essential.
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 affecting patient acceptability or ease of use or additional tests or monitoring needed.)	Oral therapies can be administered without outpatient attendance at IV Day Units, which frees up space and staffing resource. Myelosuppression resulted in admission with neutropenic fever in 15% of patients receiving docetaxel + nintedanib, and in 10% of those receiving docetaxel alone (Reck, M et al. Lancet Oncol 2014). These patients require emergency admission for intravenous antibiotics. This toxicity was not reported with sotorasib (Hong D et al. New Engl J Med 2020).
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No
15. Do you consider that the use of the technology will result in any substantial health-	No

related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
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?	Sotorasib provides the first ever approved KRAS-targeted therapy.
• Is the technology a 'step-change' in the management of the condition?	Yes.
• Does the use of the technology address any particular unmet need of the patient population?	Unlike other commonly-identified genomic drivers in lung cancer, no targeted therapy has previously been available for treatment of *KRAS*-mutated lung adenocarcinoma.

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?	Compared with chemotherapy-based alternatives, the side effect profile of sotorasib (even administered long-term) is preferable to the universally-experienced myelosuppression and alopecia, and common nausea and vomiting, associated with comparator chemotherapy.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Randomised comparative data in not available.
• If not, how could the results be extrapolated to the UK setting?	Historical controls in NSCLC adenocarcinoma histology would be fair (eg Reck, M et al. Lancet Oncol 2014).
• What, in your view, are the most important outcomes, and were they measured in the trials?	Response rate and PFS are important in the early assessment of targeted therapies in selected populations. Both were measured for sotorasib (Hong D et al. New Engl J Med 2020), and appear superior to historical controls treated with docetaxel +- nintedanib.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	PFS in this setting seems to correlate with OS benefit (see above). Furthermore, progression of NSCLC is generally associated with the emergence of more prominent symptoms and deteriorating quality of life.

• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not to my knowledge.
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]	No.

21. How do data on real-world experience compare with the trial data?	N/A
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.	N/A
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	N/A

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.

- *KRAS* mutation is very common in NSCLC
- The commonest specific *KRAS* mutation in this disease is G12C, the target of sotorasib
- Docetaxel, or docetaxel + nintedanib, are appropriate comparators in NSCLC previously treated with platinum-doublet chemotherapy and immune checkpoint inhibition
- Use of sotorasib is better tolerated and less resource-intensive than comparator chemotherapy-based treatment
- Efficacy of sotorasib seems superior to comparator chemotherapy-based treatment, although randomised comparisons are not available.

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 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.

in collaboration with:

university of groningen

University Medical Center Groningen

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Produced by Kleijnen Systematic Reviews (KSR) Ltd in collaboration with University Medical Center (UMC) Groningen, the Netherlands

Authors
Robert Wolff, Managing Director, KSR Ltd, United Kingdom (UK)
Thea van Asselt, Health Economist, UMC Groningen, the Netherlands
Sajad Emamipour, Health Economist, UMC Groningen, the Netherlands
Simon van der Pol, Health Economist, UMC Groningen, the Netherlands
Maarten Postma, Health Economist, UMC Groningen, the Netherlands
Annette Chalker, Systematic Reviewer, KSR Ltd, UK
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Sean Harrison, Statistician, KSR Ltd, UK
Shelley de Kock, Information Specialist, KSR Ltd, UK
Jos Kleijnen, Director, KSR Ltd, UK

Correspondence to
Robert Wolff, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York YO19 6FD
United Kingdom

Date completed 03/09/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number 13/51/34.

Declared competing interests of the authors:
None.

Acknowledgements
[redacted]

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:
Wolff R, van Asselt T, Emamipour S, van der Pol S, Postma M, Chalker A, Posadzki P, Ahmadu C, Armstrong N, Harrison S, de Kock S, Kleijnen J. Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors
Robert Wolff 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. Thea van Asselt acted as health economic project lead, critiqued the company's economic evaluation, and contributed to the writing of the report. Sajad Emamipour, Simon van der Pol, Maarten Postma, Charlotte Ahmadu and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Annette Chalker and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence, and contributed to the writing of the report. Sean Harrison acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Shelley de Kock critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen 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 of Cancer Research
ACP	American College of Physicians
AE	Adverse event
AIC	Akaike Information Criterion
AiC	Academic in confidence
ALK	Anaplastic lymphoma kinase
ALT	Alanine aminotransferase
ANZCTR	Australian New Zealand Clinical Trials Registry
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
ATE	Average treatment effect
ATT	Average treatment effect of the treated
BIC	Bayesian information criterion
BICR	Blinded independent central review
BNF	British National Formulary
BRAF	B-Raf Proto-oncogene
BSC	Best supportive care
CDF	Cancer Drugs Fund
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CiC	Commercial in confidence
CMR	Cochrane Methodology Register
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
DARE	Database of Abstracts of Reviews of Effects
DNA	Deoxyribonucleic acid
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EED	Economic Evaluations Database
EGFR	Epidermal growth factor receptor
ELCC	European Lung Cancer Congress
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-5L	European Quality of Life-5 Dimensions-5 Levels
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
ESS	Effective sample size
EU CTR	European Clinical Trials Register
FDA	Food and Drug Administration
FE	Fixing errors
FV	Fixing violations
G12C	G12C amino acid substitution
GID	Guideline in development
HCHS	Hospital and Community Health Services
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment

IASLC	International Association for the Study of Lung Cancer
ICTRP	International Clinical Trials Registry Platform
ICER	Incremental cost effectiveness ratio
IPW	Inverse probability weighting
ISRCTN	International Standard Randomized Controlled Trial Number
ITC	Indirect treatment comparison
IV	Intravenous
KM	Kaplan-Meier
KRAS	Kirsten rat sarcoma viral oncogene homolog
KSR	Kleijnen Systematic Reviews
LYG	Life years gained
MAIC	Matching adjusted indirect comparison
MedDRA	Medical Dictionary for Regulatory Activities
MEK	Mitogen activated protein kinase
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MJ	Matters of judgement
MMRM	Mixed models with repeated measures
mRCT	metaRegister of Controlled
MRI	Magnetic resonance imaging
MTA	Multiple technology appraisal
N	Number of participants in the analysis set
n	Number of participants with observed data
N/A	Not applicable
NCI	National Cancer Institute
NHS	National Health Service
NHSCII	National Health Service Cost Inflation Index inflation indices
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic tyrosine kinase
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD-1	Programmed cell death-1
PD-L1	Programmed death-ligand 1
PF	Progression free
PFS	Progression-free survival
PP	Post progression
PPS	Post progression survival
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PSWA	Propensity score weighted analysis
QALY	Quality adjusted life year
QLQ	Quality of Life Questionnaire
RA	Regression adjustment
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 of Interventions
ROS	Proto-oncogene tyrosine-protein kinase
SAE	Serious adverse event
SD	Standard deviation
SLR	Systematic literature review
STM	State transition model
TA	Technology appraisal
TEAE	Treatment-emergent adverse events
TEW	Treatment effect waning
TRAE	Treatment related adverse events
TK	Tyrosine kinase
TSD	Technical support document
TTD	Time to treatment discontinuation
UK	United Kingdom
UKCCCR	United Kingdom Coordinating Committee on Cancer Research
UMC	University Medical Center
VAS	Visual analogue scale
VEGFR	Vascular endothelial growth factor receptor
WCLC	World Conference on Lung Cancer
WHO	World Health Organization
WTP	Willingness-to-pay

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 Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG'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, Section 1.4 issues relates to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness. A summary is presented in Section 1.6.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1 provides a summary of the key issues identified by the ERG.

Table 1.1: Summary of key issues

Issue	Summary of issue	Report Sections
1	Population narrower than NICE scope	2.1
2	Generalisability / lack of UK participants	2.1.3, 3.2.1
3	High risk of bias of CodeBreaK100	3.2.3
4	High number of serious adverse events observed in CodeBreaK100	3.2.4.5
5	Validity of ITC without a common comparator	3.3, 3.4
6	Partitioned Survival Model structure not validated or justified	4.2.2
7	Exclusion of platinum-based chemotherapy as a comparator in 2nd line	4.2.4
8	Docetaxel plus nintedanib modelling approach leading to worse survival	4.2.6
9	No waning of treatment effect	4.2.6
10	TTD modelling approach inconsistent with OS and PFS modelling	4.2.6
11	Time-to-death utilities do not seem well-informed	4.2.8
12	Disutility for IV administration not well justified	4.2.8
13	Relative dose intensity and wastage assumption not justified	4.2.9

ITC = indirect treatment comparison; IV = intravenous; NICE = National Institute for Health and Care Excellence; TTD = time to treatment discontinuation; UK = United Kingdom

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are a different approach to estimating utility values, a different approach to estimating time to treatment discontinuation (TTD), the incorporation of treatment waning, and, specifically for the secondary comparison, assuming that docetaxel plus nintedanib cannot be worse than docetaxel in terms of overall survival (OS).

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (OS) and quality of life in a quality-adjusted life year (QALY). An incremental cost effectiveness ratio (ICER) is the ratio of the extra cost per QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increasing survival, which accrues in progression-free survival (PFS; [redacted] vs. [redacted] months) as well as in post-progression survival (PPS; [redacted] vs. [redacted] months).
- Increasing health-related quality of life (HRQoL) because of longer survival (via time-to-death utilities) and because of the treatment-related disutility for docetaxel.

Overall, the technology is modelled to affect costs by:

- The higher cost of sotorasib compared to docetaxel [redacted] vs. £17.95).
- Early treatment discontinuation for sotorasib compared to docetaxel.

The modelling assumptions that have the greatest effect on the ICER are:

- The hazard ratio applied to PFS to model sotorasib treatment duration (TTD).
- The time to death utility for >6 months prior to death.
- The OS hazard ratio for docetaxel plus nintedanib versus docetaxel (for the secondary comparison only).

1.3 The decision problem: summary of the ERG's key issues

The decision problem addressed in the company submission (CS) is slightly narrower than that specified in the final scope, see Table 1.2.

Table 1.2: Key issue 1. Population narrower than NICE scope

Report Section	**2.1**
Description of issue and why the ERG has identified it as important	There is a discrepancy of populations 1) defined in the NICE scope, 2) addressed in the CS decision problem, and 3) included in CodeBreaK100, providing the primary clinical trial evidence: 1. Adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC 2. Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated 3. Adult patients with KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of 0 or 1 Of note, the anticipated marketing authorisation is for the "treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemo-therapy and/or anti PD-1/PD-L1 immunotherapy, unless contra-indicated". The ERG would bring this issue to the attention of the committee as it potentially limits the population for which a decision is made.

Report Section	2.1
What alternative approach has the ERG suggested?	Further evidence should be gathered to cover the population defined in the NICE scope.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should be gathered to cover the population defined in the NICE scope.
CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; RECIST = Response Evaluation Criteria in Solid Tumours	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

Generalisability to England and Wales is unclear due to the lack of centres in the United Kingdom, see Table 1.3.

The ERG assessed the risk of bias of the CodeBreaK100 study, the primary clinical trial evidence, using the ROBINS-I (Risk Of Bias in Non-randomised Studies of Interventions) tool and rated it the risk of bias to be "serious", see Table 1.4.

Furthermore, the ERG would like to highlight the high number of treatment-emergent adverse events (TEAEs) observed in the CodeBreaK100 study, see Table 1.5.

Finally, the ERG has concerns regarding the validity of indirect comparisons performed in the CS, see Table 1.6.

Table 1.3: Key issue 2. Generalisability / lack of UK participants

Report Section	2.1.3, 3.2.1
Description of issue and why the ERG has identified it as important	The participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the UK. The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample).
What alternative approach has the ERG suggested?	Further analyses of countries similar to the UK would be informative.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further analyses of countries similar to the UK would be informative.
ERG = Evidence Review Group; UK = United Kingdom	

Table 1.4: Key issue 3. High risk of bias of CodeBreaK100

Report Section	**3.2.3**
Description of issue and why the ERG has identified it as important	Using the ROBINS-I tool, the company rated overall risk of bias of CodeBreaK100 to "low to moderate". However, the ERG re-assessed the study and rated the risk of bias to be "serious". Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.
What alternative approach has the ERG suggested?	Further evidence should aim to minimise the risk of bias
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased
What additional evidence or analyses might help to resolve this key issue?	Further evidence should aim to minimise the risk of bias
CS = company submission; ERG = Evidence Review Group; ROBINS-I = Risk Of Bias in Non-randomised Studies of Interventions	

Table 1.5: Key issue 4. High number of serious adverse events observed in CodeBreaK100

Report Section	**3.2.4.5**
Description of issue and why the ERG has identified it as important	The ERG is concerned with the high number of treatment-emergent adverse events, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.
What alternative approach has the ERG suggested?	None. The ERG wants to highlight the issue for the committee.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Potential guidance should reflect this issue.
AE = adverse event; ERG = ERG = Evidence Review Group; NSCLC = non-small cell lung cancer	

Table 1.6: Key issue 5. Validity of ITC without a common comparator

Report Section	3.3, 3.4
Description of issue and why the ERG has identified it as important	The ITC is unanchored i.e. no common comparator. Therefore, there are potentially relevant differences in prognostic factors between the studies included in the ITCs (CodeBreaK100, SELECT-1, LUME-Lung 1), e.g. regarding G12C KRAS mutation status, prior therapies, presence of brain metastases, and factors like sex and smoking history. It is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC. The company chose a MAIC for their primary analysis of the main comparison with docetaxel, which is particularly prone to bias given lack of identification of all relevant prognostic factors and clinical experts identified factors to be "very important", e.g. brain metastases and disease stage at baseline. However, these, alongside G12C mutation status, were not considered for the MAIC comparing CodeBreaK100 and SELECT 1. Also, because only summary statistics were available from SELECT-1, the CodeBreaK 100 had to be adjusted to match the SELECT-1 population. The company also conducted a supplementary analysis using the Flatiron study, which, using a method of adjustment, referred to as PSWA that appears to involve IPW, allowed the comparator data to match the CodeBreaK 100 population. A richer set of individual patient data also afforded a greater number of potential prognostic factors. In addition to the underlying uncertainty introduced by an indirect comparison of treatments (compared to a direct comparison), the differences between studies, the choice of baseline variables for matching, the choice of underlying data source and adjustment method can be questioned, and the ERG would have liked to see further analyses.
What alternative approach has the ERG suggested?	1. For the MAIC, an analysis with mutation status as covariate could be informative 2. For the PSWA, methods other than IPW, such as RA or doubly robust (RA plus IPW), could have been employed and so scenario analyses using these methods could be informative 3. For the PSWA, limiting to the docetaxel only population could be informative 4. In principle, evidence directly comparing treatments would provide more robust evidence.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	See suggestions above.

ERG = Evidence Review Group; G12C = G12C amino acid substitution; IPW = inverse probability weighting; ITC = indirect treatment comparison; KRAS = Kirsten rat sarcoma viral oncogene homolog; MAIC = matching adjusted indirect comparison; PSWA = propensity score weighted analysis; RA = regression adjustment

1.5 The cost effectiveness evidence: summary of the ERG's key issues

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues in the cost effectiveness evidence are discussed in the Tables below.

Table 1.7: Key issue 6. Partitioned Survival Model structure not validated or justified

Report Section	**4.2.2**
Description of issue and why the ERG has identified it as important	The company used a partitioned survival model without elaborate justification and without an accompanying scenario implementing an STM to validate the results
What alternative approach has the ERG suggested?	The ERG did not suggest an alternative approach other than the STM
What is the expected effect on the cost effectiveness estimates?	The expected effect cannot be predicted
What additional evidence or analyses might help to resolve this key issue?	The ERG recognises that it is difficult and intensive to provide results from a model with an alternative structure.
ERG = Evidence Review Group; STM = state transition model	

Table 1.8: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line

Report Section	**4.2.4**
Description of issue and why the ERG has identified it as important	Compared to the final scope for this appraisal, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority
What alternative approach has the ERG suggested?	The ERG has no alternative approach as adding the comparator to the model would require structural and substantial changes which are outside the scope of work for the ERG.
What is the expected effect on the cost effectiveness estimates?	Could potentially have a substantial impact on the cost effectiveness, direction unknown.
What additional evidence or analyses might help to resolve this key issue?	Implementing platinum-based chemotherapy in the model as an additional comparator would help to resolve the issue and reduce uncertainty.
ERG = Evidence Review Group	

Table 1.9: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The indirect way of estimating OS and PFS for the secondary comparator docetaxel plus nintedanib leads to worse survival for

	docetaxel plus nintedanib compared to docetaxel plus placebo in the first six months of the OS curve.
What alternative approach has the ERG suggested?	The ERG prefers to assume that the HR for docetaxel plus nintedanib versus docetaxel plus placebo cannot go above 1.
What is the expected effect on the cost effectiveness estimates?	Lowering the HR for docetaxel plus nintedanib versus docetaxel plus placebo will increase the ICER for sotorasib versus docetaxel plus nintedanib.
What additional evidence or analyses might help to resolve this key issue?	Direct evidence for this comparison.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression-free survival	

Table 1.10: Key issue 9. No waning of treatment effect

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The company's assumption of continued effect of sotorasib does not seem justified and is difficult to maintain given immature evidence.
What alternative approach has the ERG suggested?	The ERG suggested to start waning of the treatment effect at the 2-year timepoint and have it gradually decreased to an HR of 1 over a period of 5 years (with exploratory scenario analyses for 3 and 7 years).
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on lasting treatment effect.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio	

Table 1.11: Key issue 10. TTD modelling approach inconsistent with OS and PFS modelling

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The TTD was modelled by applying a hazard ratio to PFS from CodeBreaK100. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, fitting a parametric curve on TTD data using weights based on the MAIC.
What alternative approach has the ERG suggested?	The ERG suggested to use the company's alternative approach, based on the MAIC, in the base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on observed treatment duration in sotorasib and comparator arms

Report Section	4.2.6
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation	

Table 1.12: Key issue 11. Time-to-death utilities do not seem well-informed

Report Section	4.2.8
Description of issue and why the ERG has identified it as important	The time to death utilities which the company used in the base-case did not seem well-informed. The data underlying the estimates were sparse, and increasingly so for the closer to death states.
What alternative approach has the ERG suggested?	The ERG suggested to use utilities based on disease progression as base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Fully specified models using also AN02 dataset should be provided to see which approach is most appropriate. But given that even AN02 probably has many missing data this may still not be ideal.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 1.13: Key issue 12. Disutility for IV administration not well justified

Report Section	4.2.8
Description of issue and why the ERG has identified it as important	A disutility for IV administration of docetaxel is applied without sufficient justification for the size of the disutility or the exclusion of the potential disutility for taking eight tablets of sotorasib daily.
What alternative approach has the ERG suggested?	The ERG suggested to exclude the IV disutility in the base-case
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Comparative evidence on (observed) health state utilities in sotorasib and comparator arms could resolve this
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratios; IV = intravenous	

Table 1.14: Key issue 13. Relative dose intensity and wastage assumption not justified

Report Section	4.2.9
Description of issue and why the ERG has identified it as important	In their base-case, the company assumed a lower RDI for sotorasib than for comparators, which was not justified. The company also assumed zero wastage for sotorasib, which the ERG also considered not justified.

Report Section	4.2.9
What alternative approach has the ERG suggested?	The ERG proposed to take the average RDI as base-case, and to include wastage based on opened packs.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	For the wastage, the company would have to make a convincing case that opened packs, when not used, would be returned for usage by other patients, i.e. a specific program would have to be in place.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratios; RDI = relative dose intensity	

1.6 Summary of the ERG's view

In conclusion, cost effectiveness estimates of sotorasib compared with docetaxel and with docetaxel plus nintedanib are subject to considerable uncertainty, mainly because of immaturity of data and lack of comparative evidence in various areas. Even when all the ERG preferred assumptions were implemented in the model, uncertainty remained on a number of issues, such as whether all relevant comparators were included in the analysis, treatment duration and long-term efficacy of sotorasib, and comparative HRQoL values. The comparison for docetaxel plus nintedanib is potentially more heavily biased even because of the indirectness of the two-step approach to model OS and PFS, see Tables 1.15 to 1.18.

Table 1.15: ERG base-case adjustments (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	43,660
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,394
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,496
MJ 4: Including drug wastage (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	46,387
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	47,208
MJ 6: Subsequent treatments based on alternative distribution					
Docetaxel	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	43,825
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,339
MJ 8: gradual waning of treatment effect over 5 years, starting at 2-year timepoint (key issue 9)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	48,332
ERG base-case					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,415
Based on CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violations; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation					

Table 1.16: ERG base-case adjustments (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,628
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,909
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,692
MJ 4: Including drug wastage (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,673
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	35,990
MJ 6: Subsequent treatment based on alternative distribution					
Docetaxel + nintedanib	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,839
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,087
MJ 8: gradual waning of treatment effect over 5 years, starting at 2-year timepoint (key issue 9)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,618
MJ 9: Assuming HR of 1 for OS for nintedanib for the first period (key issue 8)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,969
ERG base-case					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,051

Based on CS updated model

CS = company submission; ERG = Evidence Review Group; FV = fixing violations; HR = hazard ratio; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; OS = overall survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation

Table 1.17: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	57,567
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,444
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,986
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	60,809

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	60,428
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	57,206
Based on CS updated model AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year					

Table 1.18: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	50,249
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	51,874
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,733
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,851
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,179
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,074

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 6: Assuming constant HR of OS and PFS for nintedanib from 2nd period onwards					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	49,664

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC	Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immuno-therapy, unless contra-indicated	Patient population in the CodeBreaK100 trial included KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of 0 or 1.	The population is slightly narrower than population outlined in NICE scope, see Section 2.1 for details.
Intervention	Sotorasib	Sotorasib (LUMYKRAS™) administered orally at a dose of 960 mg (given as 8x 120 mg tablets) once daily until disease progression or unacceptable toxicity	N/A – in line with the NICE final scope.	The intervention is in line with the NICE scope.
Comparator(s)	**Non-squamous NSCLC:** • pemetrexed with carboplatin with or without pemetrexed maintenance • other platinum doublet chemotherapy with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • docetaxel monotherapy	**Primary comparator:** Docetaxel monotherapy **Secondary comparator:** Nintedanib + docetaxel	Docetaxel monotherapy, i.e. the primary comparator– is outside the final scope issued by NICE by not targeting people with KRAS p.G12C mutation	The NICE lung cancer pathway and international clinical guidelines recognise the increasing role of combination immunotherapy and chemotherapy in the first-line setting for NSCLC. It is unclear why results for other comparators are

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	• atezolizumab • nivolumab (subject to ongoing CDF review) • pembrolizumab (PD-L1-expressing tumours) • best supportive care **Squamous NSCLC:** • gemcitabine with carboplatin or cisplatin • vinorelbine with cisplatin or carboplatin • docetaxel monotherapy • pembrolizumab (PD-L1-expressing tumours) • atezolizumab • nivolumab • best supportive care **People with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1):** Established clinical management without sotorasib, including: • atezolizumab combination (after EGFR-TK or ALK-targeted therapies) • lorlatinib (after ALK-targeted therapies			unavailable, see Section 2.3 for details.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	• brigatinib (after ALK-targeted therapies) • ceritinib (after ALK-targeted therapies) • osimertinib (EGFR T790M mutation-positive after EGFR-TK targeted therapies) • pemetrexed with carboplatin • platinum doublet chemotherapy with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • nivolumab (subject to ongoing CDF review)			
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • time to treatment discontinuation • adverse effects of treatment • health-related quality of life	• overall survival • progression-free survival • response rates • duration of response • adverse effects of treatment • health-related quality of life	The outcomes reported are largely in line with the NICE scope	Time to treatment discontinuation is missing in the CS, see Section 2.4 for details.
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.	[Not completed in the CS]	[Not completed in the CS]	The approach taken for the economic analysis is largely in line with the reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	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 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. The use of sotorasib is conditional on the presence of KRAS G12C mutation. The economic modelling should include the costs associated with diagnostic testing for KRAS G12C in people with NSCLC who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test. See Section 5.9 of the Guide to the Methods of Technology Appraisals'.			No full incremental analysis was performed though, see Table 4.3. The costs associated with diagnostic testing for KRAS G12C mutation was not included in the economic modelling because KRAS testing is routinely commissioned by NHS in NSCLC.
Special considerations including issues related to equity or equality	N/A	• In contrast to NSCLC patients with other oncogenic mutations, patients with advanced or metastatic KRAS p.G12C-mutated NSCLC who have failed prior therapy currently have no targeted therapy options, and very few other effective therapy options. Their prognosis is very poor, with OS significantly less than 2 years.		N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
		• Sotorasib is a highly innovative, first in class therapy for KRAS p.G12C-mutated NSCLC. It provides an effective and tolerable targeted treatment option where previously there was none. It has been designated as a Promising Innovative Medicine via the UK Early Access to Medicines Scheme, and was granted an Innovation Passport under the Innovative Licensing and Access Pathway. UK orphan designation is pending. • Subject to approval, sotorasib is anticipated to be granted conditional marketing authorisation by the MHRA via the Project Orbis regulatory route on the basis of the results of the phase 2 CodeBreaK100 single arm trial. • As sotorasib is the first KRASG12Cinhibitor to progress to licensing by any regulatory authority there is a lack of data specifically in patients with KRAS p.G12C mutated NSCLC for the relevant comparators, or any other agents. • Indirect comparative data using the most robust methods possible indicate that sotorasib is highly effective in achieving clinically meaningfully improvements in PFS and OS by >3 months compared with relevant comparators. • Based on these data, sotorasib provides a step change in therapy for patients with KRAS p.G12C mutated NSCLC and is highly likely to be cost effective under the NICE end of life policy. • Phase 3 data from the CodeBreaK200 RCT are anticipated within the next 2 years. • Sotorasib may therefore be a candidate for the CDF.		

Based on Table 1 of the CS[1]

ALK = anaplastic lymphoma kinase; CDF = Cancer Drugs Fund; CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; MHRA = Medicines and Healthcare Products Regulatory Agency; N/A = not applicable; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; OS =

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
overall survival; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; ROS = proto-oncogene tyrosine-protein kinase; TK = tyrosine kinase; UK = United Kingdom				

2.1 Population

The NICE scope defined the population of interest as *"adults with previously treated Kirsten rat sarcoma viral oncogene homolog (KRAS) p.G12C mutated, locally advanced or metastatic non-small cell lung cancer (NSCLC)"*.[2]

The company submission (CS) defined the population of interest as *"adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated"*.[1]

ERG comment: The population addressed in the CS is narrower than the population defined in the NICE scope:

1. The CS only considered patients *"previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated"*.[1]
2. The population in CodeBreaK100, providing the primary clinical trial evidence for sotorasib in the CS, is even narrower than that specified in the NICE scope, namely *"KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per the RECIST* [Response Evaluation Criteria In Solid Tumours] *1.1 criteria, and had ECOG* [Eastern Cooperative Oncology Group] *performance status of 0 or 1"*.[1]

2.1.1 Previous treatment

It is unclear why only platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy are considered while other lines of therapy are not.

2.1.2 Population in CodeBreaK100

The CS did not define the 1-3 prior lines of anticancer therapy. Real-world data show there is a variety of first-line treatment strategies (checkpoint inhibitor ± chemotherapy, platinum + pemetrexed, platinum + taxanes, or other chemotherapy) and a variation in second-line treatment regimens while a proportion of patients also receive third-line treatment.[3]

It is unclear why patients with ECOG performance status of 0 to 2 (on a 5-point scale, with higher numbers indicating greater disability) were eligible in phase I of the CodeBreaK100 trial, whereas phase II of the trial only included phases 0 to 1 (less severe).[4, 5]

Of note, the population in CodeBreaK100 appears to be not only narrower than the NICE scope but also than the anticipated marketing authorisation, e.g. in regards to the ECOG status of included participants. According to the CS, an application for UK marketing authorisation for sotorasib was submitted to the Medicines and Healthcare Products Regulatory Agency (MHRA) in January 2021 with a proposed indication for use as monotherapy for treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated. A conditional licensing approval via the Project Orbis regulatory route in the UK is anticipated [redacted]*.[1]

In response to the request for clarification, the company stated that *"the exclusion of patients with ECOG PS* [performance status] *2 from the CodeBreaK100 trial should not preclude the use of sotorasib within its licensed indication in such patients in clinical practice. Sotorasib should be an option available to clinicians for use in patients with ECOG PS 2 when clinically relevant"*.[6] However, no evidence was provided to support this statement.

2.1.3 Generalisability of trial population

As discussed in Section 3.2.1, the participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the United Kingdom (UK). The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample; see Table 4 of the CS).[1]

Table 2.2: Key issue 1. Population narrower than NICE scope

Report Section	**2.1**
Description of issue and why the ERG has identified it as important	There is a discrepancy of populations 1) defined in the NICE scope, 2) addressed in the CS decision problem, and 3) included in CodeBreaK100, providing the primary clinical trial evidence: 1. Adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC 2. Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated 3. Adult patients with KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of zero or one Of note, the anticipated marketing authorisation is for the "treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemo-therapy and/or anti PD-1/PD-L1 immunotherapy, unless contra-indicated". The ERG would bring this issue to the attention of the committee as it potentially limits the population for which a decision is made.
What alternative approach has the ERG suggested?	Further evidence should be gathered to cover the population defined in the NICE scope.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should be gathered to cover the population defined in the NICE scope.
CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; RECIST = Response Evaluation Criteria in Solid Tumours	

Table 2.3: Key issue 2. Generalisability / lack of UK participants

Report Section	**2.1.3, 3.2.1**
Description of issue and why the ERG has identified it as important	The participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the UK. The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample).

Report Section	2.1.3, 3.2.1
What alternative approach has the ERG suggested?	Further analyses of countries similar to the UK would be informative.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further analyses of countries similar to the UK would be informative.
ERG = Evidence Review Group; UK = United Kingdom	

2.2 Intervention

The intervention (AMG 510/LUMYKRAS™) is in line with the scope.

Sotorasib is administered orally at a dose of 960 mg (given as 8x 120 mg tablets) once daily until disease progression, no further clinical benefit is expected, unacceptable toxicity, withdrawal of consent, or death.[1] Sotorasib is a small molecule that specifically inhibits KRAS G12C amino acid substitution (G12C) in advanced solid tumours through a unique interaction with the P2 pocket of the switch II region.[5]

ERG comment: Participants in the CodeBreaK100 trial used a combination arm with sotorasib and anti PD-1/L1 or midazolam at phase I.[7] It is not clear how these participants were handled in the analyses given that sotorasib was outlined as monotherapy as per NICE scope.[2] This might have an impact on the results of effectiveness as well as cost effectiveness analyses.

2.3 Comparators

The description of the comparators in the NICE scope includes eight different treatments for non-squamous NSCLC, seven treatments for squamous NSCLC; and atezolizumab combination, lorlatinib, brigatinib, ceritinib, osimertinib, pemetrexed with carboplatin, platinum doublet chemotherapy (with or without pemetrexed maintenance), and established clinical management without sotorasib for people with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1), see Table 2.1 and NICE scope.[2]

The CS listed two comparators, docetaxel monotherapy as the primary comparator and nintedanib + docetaxel as the secondary comparator.[1]

In response to the request for clarification, the company confirmed that other comparators have not been considered to be relevant comparators for sotorasib.[6]

ERG comment: The primary comparator selected by the company, docetaxel monotherapy, was listed as a comparator for non-squamous NSCLC in the NICE scope.[2] However, it is outside the NICE scope for people with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1).[2]

It should be noted that Peter Clark (The Clatterbridge Cancer Centre NHS Foundation Trust; NHS England Cancer Drugs Fund (CDF) clinical lead) highlighted that *"KRAS 12C mutations are mutually exclusive to other targetable mutations"*.[8]

The company selected nintedanib in combination with docetaxel as the secondary comparator which is in line with NICE technology appraisal (TA) 347 for patients with adenocarcinoma and in line with the NICE scope.[2, 9]

Following advice by Peter Clark, the ERG considers the main comparator to be second-line docetaxel monotherapy and would consider the secondary comparator, nintedanib + docetaxel as a scenario analysis.[8]

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- Overall survival
- Progression-free survival
- Response rates
- Duration of response
- Adverse effects of treatment
- Health-related quality of life.

ERG comment: Most of these outcomes were included in the decision problem addressed in the CS as well as assessed in the CodeBreaK100 trial except time to treatment discontinuation (TTD).[1]

However, as stated in the response for the request for clarification, TTD was used to inform the economic model.[6] However, as discussed in Section 4.2.6 of the report, TTD was based on progression-free survival (PFS) using a hazard ratio (HR).

As detailed in Section 3.2.4.5, the ERG is concerned with the high number of treatment-emergent adverse events (TEAEs).

As detailed in Section 3.2.4.6, health-related quality of life (HRQoL) was only summarised descriptively; and changes from baseline using mixed effects models for repeated measures are tested.

2.5 Other relevant factors

According to the company, sotorasib is highly innovative and has been granted an Innovation Passport under the Innovative Licensing and Access Pathway; and addresses a significant unmet need in patients with *KRAS p.G12C* -mutated NSCLC (Section B.2.12 of the CS).[1] The drug also received accelerated approval by the United States Food and Drug Administration (FDA) on 28 May 2021 under its Real-Time Oncology Review (Section B.1.2 of the CS).[1]

Sotorasib is offered at an undiscounted price of [redacted] per patient per treatment (Table 2 of the CS).[1] The company highlighted that sotorasib may be a candidate for the CDF.

Sotorasib might fulfil the end of life criteria as specified by NICE. However, as discussed in Section 7, the ERG has concern regarding the validity of the data used to inform the second criterion, extension of life of ≥3 months.

According to the company, *"no specific equality considerations are anticipated"* (SectionSection B.1.4 of the CS).[1]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

Appendix D of the CS detailed a systematic literature review (SLR) conducted to identify trial data for systemic drug therapies used in the management of patients with KRAS mutant NSCLC.[1]

The SLR search strategy was based on a published SLR by Schulz et al. which was conducted in the pre-treated NSCLC population.[10] As this review set out to include publications reporting outcome data for a KRAS mutant population, any studies identified as relevant by Schulz et al. were included as well as all relevant studies published during or after 2015 as identified by replicating the Schulz et al. strategy.[10]

Searches were run in June 2020 and updated on 26 January 2021. In addition to a search for randomised controlled trials (RCTs), searches were also conducted to identify single arm trials with KRAS mutant NSCLC. These searches were undertaken on 24 July 2019 and updated on 10 March 2021. A summary of sources searched is provided in Table 3.1.

Table 3.1: A summary of sources searches to identify trial data

Resource	Host/Source	Date Ranges	Dates searched
RCT searches			
Electronic Databases			
Embase	Ovid	1980 – present	25 June 2020 26 January 2021
MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions®	Ovid	1946 – present	12 June 2020 26 January 2021
CDSR CENTRAL	Ovid		12 June 2020 26 January 2021
Conference proceedings			
ASCO	https://www.asco.org/	January 2017 – January 2021	
ESMO	http://www.esmo.org/		
IASLC World Congress on Lung Cancer	https://wclc2019.iaslc.org/		
AACR	https://www.aacr.org/Pages/Home.aspx		
Clinical trial registries			
Clinicaltrials.gov	www.clinicaltrials.gov		

Resource	Host/Source	Date Ranges	Dates searched
NCI clinical trial database	https://www.cancer.gov/	January 2017 – January 2021	
UKCCCR Register of Cancer Trials	http://www.ctu.mrc.ac.uk/ukcccr/		
ISRCTN Register	https://www.isrctn.com/		
EORTC	https://www.ukctg.nihr.ac.uk/		
UK Clinical Trials Gateway	https://www.ukctg.nihr.ac.uk/		
mRCT	http://www.isrctn.com/page/mrct		
Searches for single-arm trials			
Electronic databases			
Embase	Ovid	2014 - 2019	24 July 2019
		2019 - 2021	10 March 2021
MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, MEDLINE Daily and MEDLINE	Ovid	2014 – 2019	24 July 2019
		2019 – 2021	10 March 2021
CDSR	Ovid	2014 - 2019	24 July 2019
DARE		2019 – 2021	10 March 2021
CENTRAL			
NHS EED			
HTA Database			
ACP Journal Club			
Conference proceedings			
ASCO	https://www.asco.org/	2017 – 2021	24 July 2019 10 March 2021
ESMO	http://www.esmo.org/		
WCLC	https://wclc20190iaslc.org/		
ELCC	https://www.esmo.org/Conferences/ELCC-2019-European-Lung-Cancer-Congress		
Clinical trials registries			
ClinicalTrials.gov	https://clinicaltrials.gov	January 2017 – January 2021	
NIH	https://www.nih.gov/		
World Health Organization ICTRP	http://www.who.int/ictrp/en/		
ANZCTR	http://www.anzctr.org.au/		
EU CTR	https://www.clinicaltrialsregister.eu/		

AACR = American Association of Cancer Research; ACP = American College of Physicians; ANZCTR = Australian New Zealand Clinical Trials Registry; ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials;

Resource	Host/Source	Date Ranges	Dates searched
DARE = Database of Abstracts of Reviews of Effects; EED = Economic Evaluations Database; ELCC = European Lung Cancer Congress; EORTC = European Organization for Research and Treatment of Cancer; ESMO = European Society for Medical Oncology; EU CTR = European Clinical Trials Register; HTA = Heath Technology Assessment; IASLC = International Association for the Study of Lung Cancer; ICTRP = International Clinical Trials Registry Platform; ISRCTN = International Standard Randomized Controlled Trial Number; mRCT = metaRegister of Controlled Trials; NCI = National Cancer Institute; NHS = National Health Service; NIH = National Institutes of Health; RCT = randomised controlled trial; UK = United Kingdom; UKCCCR = United Kingdom Coordinating Committee on Cancer Research; WCLC = World Conference on Lung Cancer			

ERG comment: The CS provided sufficient details for the ERG to appraise the literature searches. A range of databases, conference proceedings and clinical trials registries were searched. Both the original and update searches were overall well conducted and documented, making them transparent and reproducible. A date limit was applied to the searches but this was justified as a previous SLR on pre-treated NSCLC population had been undertaken by Schulz et al.[10] A separate search for single-arm studies was undertaken without an RCT filter to pick up adverse events to any treatments for NSCLC. Searches for single-arm studies were restricted to English language only.

In response to the request for clarification, the company explained that an English language limit had been applied for pragmatic reasons as most high-quality studies are generally published in English.[6] To avoid language bias and to increase precision, the Centre for Reviews and Dissemination (CRD) guidance recommends that English language limits should not be applied at the searching stage.[11] Study design filters were applied to RCT searches but were not appropriately referenced. In response to the request for clarification, the company confirmed that a validated search filter had not been used and that the SLR was built upon the one conducted by Schulz et al.[6, 10] The ERG believes a validated RCT filter would have increased the comprehensiveness of the searches.

The CS reported that searches were modified between databases to account for differences in syntax and thesaurus headings. However, the ERG noticed that the RCT filter applied to MEDLINE searches had not been modified and many of the terms in the RCT filter did not map across automatically. The ERG requested that the company re-run MEDLINE searches with the correct medical subject headings (MeSH) terms to ensure that nothing had been inadvertently missed which the company did.[6] An additional 13 records were identified and screened. Only the population was searched for in both RCT searches and searches for single-arm studies. This seemed appropriate considering the sparsity ofthe literature.

An RCT filter was applied to searches of CDSR and CENTRAL which are already pre-filtered databases and therefore the use of a filter is considered to be overly restrictive. In response to the request for clarification, the company argued that the additional use of study filters in their experience did not significantly increase the risk of relevant studies being excluded.[6] However, this is against the explicit recommendation of the Cochrane Handbook for Systematic Reviews of Interventions which states that CENTRAL *"aims to contain only reports with study designs possibly relevant for inclusion in Cochrane Reviews, so searches of CENTRAL should not use a trials 'filter' or be limited to human studies"*.[12]

A wide range of conference proceedings and clinical trials registries were searched. Search terms were not provided in the CS but were supplied in response to clarification questions.[6] The ERG was satisfied that the search terms were sufficient. The reference lists of included publications and relevant SLRs and network meta-analyses (NMAs) were scanned for further studies.

3.1.2 Inclusion criteria

The eligibility criteria for RCTs and non-RCTs is presented in Table 3.2. However, it was initially unclear if inclusion screening was completed in duplicate or how consensus was reached. The company clarified that this stage had been completed in duplicate.[6]

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Description	Justification
Inclusion criteria		
Population	• Subject had provided informed consent prior to initiation. • Men or women ≥18 years old. • Pathologically documented, locally-advanced or metastatic stage IIIB-IV NSCLC with, KRAS p.G12C mutation or any other KRAS mutation (KRASm)) identified through DNA sequencing. • Subjects must have received (at least) prior standard therapy appropriate for their tumour type and stage of disease, or in the opinion of the investigator would be unlikely to tolerate or derive clinically meaningful benefit from appropriate standard of care therapy. • Subjects were willing to provide archived tumour samples or willing to undergo pre-treatment tumour biopsy (Part 1 Dose Exploration). • Subjects were willing to undergo pre-treatment tumour biopsy. Subjects can be allowed to enrol without undergoing a tumour biopsy upon agreement with Investigator and the Medical Monitor if a tumour biopsy was not feasible. • Measurable or evaluable disease per RECIST 1.1 criteria. • ECOG performance status of ≤2 (phase 1) or ≤1 (phase 2).	N/A
Interventions	• Sotorasib • Any therapies licensed in the United States or European Union for the second or later line treatment of patients with NSCLC • Any anti-cancer drugs, any line of treatment or no treatment	Consistent with final scope
Comparator	Any or none	Consistent with final scope

	Description	Justification
Outcomes	• Objective response rate • Partial response • Complete response • Duration of response • Disease control rate or clinical benefit rate • Treatment duration and dosing • Disease control rate • Time to response • Progression free survival • Progression after next line of therapy (PFS2) • Time to progression • Time to next treatment • Event-free survival • Overall survival • Patient-reported outcomes • HRQoL • All-grade treatment-emergent AEs • Treatment related Grade 3 or 4 AEs • Treatment related SAEs • Tolerability: dose reductions and interruptions, discontinuation (any reason), discontinuation (due to AEs)	Consistent with final scope
Study design	• Prospective randomised controlled trials (for the RCT search) • Non-RCTs, i.e. experimental/interventional, not observational (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs. ERG comment: It is unclear why phase I studies were excluded as they comprise useful for safety data. Also, unclear why non-randomised clinical trials were ineligible since the CodeBreaK100 was a non-randomised trial.
Language restrictions	English language only	To reduce number of hits and to identify studies in patient populations relevant to the UK setting

	Description	Justification
Exclusion criteria		
Population	• Subjects with active brain metastases from non-brain tumours • Paediatric and adolescent (<18 years) patients • Patients with cancers other than NSCLC • Early-stage NSCLC patients (Stage<IIIB) • Trials studying safety and efficacy of treatment administered in adjuvant setting • Treatment naïve patients	As specified by final scope
Interventions	• Treatments specifically targeting EGFR/ALK or ROS 1 mutations or other targetable mutation • Radiotherapy or surgery	Not relevant to final scope
Outcomes	Non-clinical outcomes	Not relevant to final scope
Study design	• Non-RCTs (for the RCT search) • RCTs (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs
Language restrictions	Abstracts published in non-English language	To reduce number of hits and to identify studies in patient populations relevant to the UK setting

Based on Tables 1, 2, and 6 of Appendix D of the CS[13]
AE = adverse event; ALK = Anaplastic lymphoma kinase; CS = company submission; DNA = deoxyribonucleic acid; ECOG = Eastern Cooperative Oncology Group; EGFR = Epidermal Growth Factor Receptor; ERG = Evidence Review Group; G12C = G12C amino acid substitution; HRQoL = health-related quality of life; KRAS = Kirsten rat sarcoma viral oncogene homolog; N/A = not applicable; NSCLC = non-small cell lung cancer; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; ROS = proto-oncogene tyrosine-protein kinase; SAE = serious adverse event; UK = United Kingdom

ERG comment: The inclusion criteria noted the exclusion of non-randomised trials, despite the CodeBreaK100 study being of a non-randomised design. In response to the request for clarification, the company stated that the inclusion criteria do not include searches of comparative trials that were not randomised, based on the assumption that few comparative studies are likely to be non-randomised.[6]

3.1.3 Critique of data extraction

Information provided in the CS regarding data extraction was limited. In the response to the request for clarification, the company stated that each stage of the systematic review process was completed in duplicate.[6]

3.1.4 Quality assessment

The critical appraisal of the non-randomised study was completed using the Risk Of Bias in Non-randomised Studies of Interventions (ROBINS-I) tool.[14] The SELECT-1 trial and the LUME-Lung 1 trial were reported to be assessed using the NICE single technology appraisal user guide. However, the

CS noted that aspects of the CRD guidance had been utilised.[11] , In response to the request for clarification, the company stated that this is in line with the NICE STA user guide.[6]

3.1.5 Evidence synthesis

According to Section B.2.8 of the CS, *"no meta analyses have been conducted" "as current efficacy data for sotorasib in the treatment of KRAS p.G12C-mutated NSCLC are based on a phase 2 single-arm trial"*.[1]

ERG comment: The ERG agrees that meta-analysis would not have been helpful.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

3.2.1 Design of CodeBreaK100 trial

The CodeBreaK100 trial is an ongoing phase 1/2 study, in which the phase 2 portion is a multicentre, non-randomised, open-label study.[1]

The population was comprised of adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria and had an ECOG performance status of 0 or 1. The trial locations were located in 47 centres, with none of these being based in the United Kingdom.

The intervention was comprised of 960 mg of sotorasib, which is meant to be administered orally once per day without interruption until either disease progression, intolerance, withdrawal of consent, or death. There was no listed comparator. Statistical analyses are shown in Table 3.3.

In response to the request for clarification regarding the *"blinded independent central review"*, the company noted that the blinded independent central review referred to the assessment of response per RECIST 1.1 criteria by central review, rather than investigators.[6]

Table 3.3: CodeBreaK100: study design

Study	CodeBreaK100 (NCT03600883)
Study Design (n)	Single-arm, phase 2 trial conducted in 47 centres (N=126)
Population	Adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per Response Evaluation Criteria In Solid Tumours (RECIST) 1.1 criteria, and had ECOG performance status of 0 or 1.
Intervention	Sotorasib 960 mg administered orally once per day without interruption (i.e., no planned off-treatment days) until disease progression, intolerance, withdrawal of consent or death.
Comparator	None.
Reported outcomes specified in the decision problem	• Objective response rate assessed by blinded independent central review • Overall survival • Duration of response • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life (EORTC QLQ-C30, QLQ LC13, EQ-5D-5L)

Study	CodeBreaK100 (NCT03600883)
All other reported outcomes	• Disease control • Time to release • 6- and 12-month PFS • 12-month OS • Patient-reported outcomes (NSCLC SAQ, FACT-G, PRO-CTCAE) • PK parameters and biomarkers (not further discussed in the CS)
Duration of study and follow-up	The CodeBreaK100 trial is ongoing. A primary analysis of efficacy, safety and patient-reported outcomes (PROs) data was conducted in September 2020. An updated analysis of efficacy and safety data for regulatory purposes was conducted on 1 December 2020. A phase 3 randomised controlled trial (RCT) comparing sotorasib against standard of care docetaxel in patients with NSCLC is ongoing with first results anticipated in 2022.
Countries	47 centres in the United States, Australia, Austria, Belgium, Canada, France, Germany, Japan, South Korea, and Switzerland.

Based on Table 4 of the CS[1]
CS = company submission; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Levels; KRAS = Kirsten rat sarcoma viral oncogene homolog; NSCLC = non-small lung cancer; PFS = progression-freesurvival; PRO = patient-reported outcome; QLQ = Quality of Life Questionnaire; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours

ERG comment: In response to the request for clarification regarding the generalisability of CodeBreaK100 to the clinical practice in England and Wales, the company stated that *"five UK clinical experts at an Amgen Advisory board considered that the population of patients enrolled in the CodeBreaK100 trial was reflective of patients in UK clinical practice who would meet the anticipated licensed indication"*.[6] However, the ERG wishes to emphasise that the CodeBreaK100 study did not include a single UK centre. Furthermore, at phase I, participants in the CodeBreaK100 trial used a combination of sotorasib and anti PD-1/L1 or midazolam. It is not clear how these participants were handled in the analyses and this can potentially impact on the results of effectiveness as well as cost effectiveness analyses.

3.2.2 Baseline characteristics of CodeBreaK100 trial

The baseline characteristics of the CodeBreaK100 trial are presented in Table 3.4. The participants in the phase 2 study were not randomised. The mean age of the participants was 62.9 years with a range of 37 to 80 years. The majority of the participants were white while they were evenly split among male and females. The CodeBreaK100 participants were largely comprised of people with advanced disease stages and who were either current or former smokers.

Table 3.4: Baseline characteristics of subjects in CodeBreak100, phase 2

Sotorasib 960 mg (N=126)	
Sex - n (%)	
Male	63 (50.0)
Female	63 (50.0)
Race - n (%)	
Asian	19 (15.1)
Black or African American	2 (1.6)

Sotorasib 960 mg (N=126)	
White	103 (81.7)
Other	2 (1.6)
Age (years)	
Mean	62.9
SD	9.3
Median	63.5
Min, Max	37,80
Smoking history - n (%)[a]	
Never	6 (4.8)
Current or former	117 (92.9)
NSCLC stage – n (%)	
III	5 (4.0)
IV	121 (96.0)
Metastases – n (%)	
Brain (non-active)	26 (20.6)
Liver	26 (20.6)
NSCLC histology – n (%)	
Non-squamous	125 (99.2)
adenocarcinoma	120 (95.2)
Squamous	1 (0.8)
ECOG performance status – n (%)	
0	38 (30.2)
1	88 (69.8)
Prior lines of systemic anticancer therapy – n (%)	
1	54 (42.9)
2	44 (34.9)
3	28 (22.2)
Types of prior systemic anticancer therapy[b] – n (%)	
Platinum-based chemotherapy	113 (89.7)
PD-1 or PD-L1 inhibitors	115 (91.3)
Platinum-based chemotherapy and PD1/L1 inhibitors	102 (81.0)

Based on Table 6 of the CS[1]
[a] smoking status missing for 3 participants; [b] prior systemic anticancer therapy also included targeted biologics (23.8%), targeted small molecules (7.1%), and other (0.8%)
CS = company submission; ECOG = Eastern Cooperative Oncology Group; N = number of participants in the analysis set; n = number of participants in the corresponding category; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death protein 1; PD-L1 = programmed death-ligand 1; SD = standard deviation

3.2.3 Quality of CodeBreaK100 trial

The critical appraisal of this single-arm, non-randomised study was conducted utilising the ROBINS-I, tool, see Table 3.5.

Table 3.5: Quality assessment of CodeBreaK100 using the ROBINS-I tool

Domains of risk of bias assessment							
1. Bias due to confounding	2. Bias in selection of participants into the study	3. Bias in classification and intervention	4. Bias due to deviations from intended intervention	5. Bias due to missing data	6. Bias due to measurement of outcomes	7. Bias in selection of reported result	Overall bias
1.1.PY[a] 1.2.PN[b] 1.3.PN[b] 1.4.PN[b] 1.5.PN[b] 1.6.NI[b] 1.7.NI[b] 1.8.PN[b] RoB: Serious[g]	2.1.PN[c] 2.2.N/A[b] 2.3.N/A[b] 2.4.PY[d] 2.5.N/A[b] RoB: Low	3.1.PY[d] 3.2.Y 3.3.PN RoB: Low	4.1.PN[c] 4.2.N/A[b] 4.3.N/A 4.4.PY 4.5.PY 4.6.NI[b] RoB: Moderate[g]	5.1.PY[d] 5.2.PN[c] 5.3.PN[c] 5.4.N/A[b] 5.5.N/A[b] RoB: Low	6.1.PN 6.2.PY[d] 6.3.PY[e] 6.4.PY[a] RoB: Serious[g]	7.1.PN[c] 7.2.PN[c] 7.3.PN[c] RoB: Low	Serious[f]

ERG's own assessment (please also see Table 7 of Appendix D of the CS)[13]
Response categories: N = No; N/A = Not Applicable; PN = Probably No; PY = Probably Yes; Y = Yes; NI = no information
[a] Rated PN in CS; [b] Not rated in CS; [c] Rated N in CS; [d] Rated Y in CS; [e] Rated N/A in CS; [f] Rated as "low to moderate" in CS; [g] Rated as "low" in CS
Responses in Red indicate potential marker for a serious risk of bias
Responses in Green indicate potential markers for low risk of bias
Response in Moderate indicate potential markers for moderate risk of bias
RoB = risk of bias

ERG comment: The ERG considers that this tool has not been appropriately used as there were 14 missing entries to the signalling questions in the CS.[13] Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.[13]

Hence the ERG undertook its own assessment, concluding that there was a high risk of bias related to baseline confounding, i.e. lower ECOG performance status of 0-1 at baseline favoured sotorasib. The ERG also considers that there was a high risk of bias in classification of interventions. Furthermore, the ERG concluded that appropriate methods to control for confounders such as stratification, regression, or probability weighting were not employed. In addition, there was a serious risk of bias in measurement of outcomes, i.e. outcome assessors were probably aware of the intervention received by the participants in the CodeBreaK100 trial.

In summary, the study has some important limitations as it has been judged by the ERG to be at a serious risk of bias in two (out of seven) domains of the ROBINS-I assessment tool.[14]

Table 3.6: Key issue 3. High risk of bias of CodeBreaK100

Report Section	3.2.3
Description of issue and why the ERG has identified it as important	Using the ROBINS-I tool, the company rated overall risk of bias of CodeBreaK100 to "low to moderate". However, the ERG re-assessed the study and rated the risk of bias to be "serious". Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.
What alternative approach has the ERG suggested?	Further evidence should aim to minimise the risk of bias.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should aim to minimise the risk of bias.
CS = company submission; ERG = Evidence Review Group; ROBINS-I = Risk Of Bias in Non-randomised Studies of Interventions	

3.2.4. Results of CodeBreaK100 trial

The results presented in the CS were reported from a primary analysis, in which the data cut off was 1 September 2020, along with updated analyses with data cuts of 1 December 2020 and 15 March 2021. In the response to a request for separate results for participants with and without adenocarcinoma, the company provided the information presented in Table 3.7.

Table 3.7: Efficacy in CodeBreaK100 by adenocarcinoma histology (15 March 2021 data cut, post hoc analysis)

ORR	PFS				OS			
Events/ Subjects (%) (95% CI)	**Events/ Subjects**	**Median (Months) (95% CI)**	**6 months KM Estimate (95% CI) (%)**	**12 months KM Estimate (95% CI) (%)**	**Events/ Subjects**	**Median (Months) (95% CI)**	**6 months KM Estimate (%) (95% CI)**	**12 months KM Estimate (%) (95% CI)**
Adenocarcinoma								
44/118 (37.3) (28.6 to 46.7)	82/118	6.8 (5.1 to 8.2)	52.2 (42.3 to 61.2)	26.9 (18.6 to 36.0)	62/120	12.0 (10.0 to NE)	74.2 (65.2 to 81.2)	50.5 (40.9 to 59.3)
No adenocarcinoma								
2/6 (33.3) (4.3 to 77.7)	5/6	6.2 (1.2 to NE)	50.0 (11.1 to 80.4)	33.3 (4.6 to 67.6)	2/6	NE (6.6 to NE)	100.0 (NE to NE)	66.7 (19.5 to 90.4)
Based on response to question A18 in response to the request for clarification[6] CI = confidence interval; CS = company submission; KM = Kaplan-Meier; NE = not estimable; PFS = progression-free survival; ORR = objective response rate; OS = overall survival								

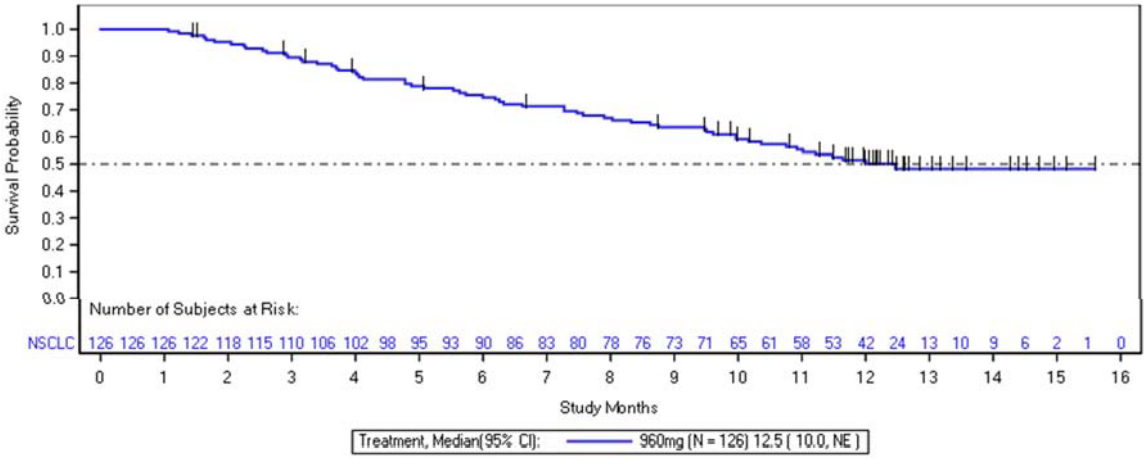
3.2.4.1 Objective response rate

ORR was the primary endpoint of the CodeBreaK100 study and was defined as the proportion of subjects with best overall response of complete response or partial response as assessed by RECIST 1.1.[1] The response was assessed by the blinded independent central review (BICR). The complete response and partial response required confirmatory computerised tomography (CT) or magnetic resonance imaging (MRI) repeat assessment at least 4 weeks after the first detection of response. According to the CS, clinical relevance was determined by the lower bound of the 95% CI excluding a prespecified benchmark of 23%.[1]

3.2.4.2 Overall survival (OS)

The median OS, as presented in the CS, was 12.5 (95% confidence interval (CI) 10.0 to not estimable) months.[1] The Kaplan-Meier (KM) estimate of survival was presented as 75.5% (95% CI 66.8 to 82.2) at 6 months and 51.4% (95% CI 41.9 to 60.1) at 12 months (Figure 3.1). Roughly half of the patients (46.8%) had experienced death at the time of the cut-off. The CS emphasises that theCodeBreaK100 study was not specifically powered for survival outcomes.[1]

Figure 3.1: Kaplan-Meier plot of overall survival (safety analysis set)



Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Death is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-003-001-eff-km-osm-nsclc-p2saf.rtf (Date Generated: 28JAN21:12:38:17) Source: adam.adsl, adam.adtte

Based on Figure 6 of the CS[1]

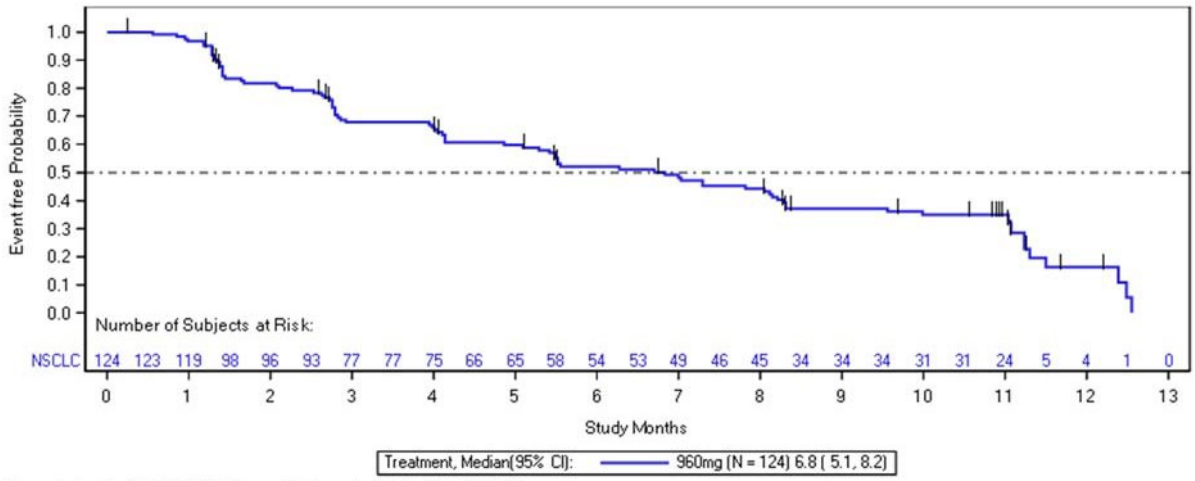
3.2.4.3 Duration of response

According to the CS, among the 46 responders who had NSCLC, the Kaplan-Meier estimate of median duration of response was 10 months (95% CI 6.9 to 11.1 months).[1] The company noted that27 subjects (58.7%) were censored. The CS also stated that 20 of the 46 objective responders were still receiving treatment without disease progression.[1]

3.2.4.4 Progression-free survival (PFS)

The median PFS was reported to be 6.8 months (95% CI 5.1 to 8.2 months) at the time of the cut-off.[1] The KM estimate of survival was 52.2% (95% CI 42.6 to 60.9) at 6 months and 16.3% (95% CI 7.4 to 28.2) at 12 months. According to the CS, 56.5% of the patients had experienced disease progression, while 10.5% of patients experienced death.[1] Of note, 41 patients were censored, which comprised of 25 patients who were on the study without disease progression, seven who started new anticancer therapy, five who missed more than one consecutive assessment, and three who withdrew their consent.[1]

Figure 3.2: Kaplan-Meier plot of progression-free survival



Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Radiological Progression or Death (whichever occurs earlier) is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-002-001-eff-km-pfs-nsclc-p2fas.rtf (Date Generated: 28JAN21:12:38:15) Source: adam.adsl, adam.adtte

Based on Figure 5 of the CS[1]

3.2.4.5 Adverse effects of treatment

The CS provided the frequencies of treatment-emergent adverse events (TEAEs) experienced in the CodeBreaK100 study.[1] As presented in Table 3.8, nearly all participants in the CodeBreaK100 study (99.2%) experienced TEAEs.

Table 3.8: Summary of overall adverse events in NSCLC subjects in CodeBreaK100

Sotorasib 960 mg daily (N=126), n (%)	
All treatment-emergent adverse events	**125 (99.2)**
Grade ≥2	110 (87.3)
Grade ≥3	75 (59.5)
Grade ≥4	23 (18.3)
Serious adverse events	63 (50.0)
Leading to discontinuation of sotorasib	11 (8.7)
Serious	7 (5.6)
Non-serious	5 (4.0)
Fatal adverse events	20 (15.9)
Treatment-related treatment-emergent adverse events	**88 (69.8)**
Grade ≥2	49 (38.9)
Grade ≥3	26 (20.6)
Grade ≥4	1 (0.8)
Serious adverse events	10 (7.9)
Leading to discontinuation of sotorasib	9 (7.1)
Serious	4 (3.2)
Non-serious	5 (4.0)
Fatal adverse events	0 (0.0)

Sotorasib 960 mg daily (N=126), n (%)
Based on Table 17 of the CS[1] Coded using MedDRA version 23.1. Severity graded using CTCAE version 5.0 CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities; N = number of participants in the analysis set, n = number of participants with observed data

The most commonly reported TEAE in ≥10% of NSCLC patients in the CodeBreaK100 trial included diarrhoea, nausea, fatigue, and elevations in alanine and aspartate aminotransferase, see Table 3.9 and Table 3.10 for TEAEs occurring in >5% of participants. The company noted that sotorasib appeared to be well tolerated and the adverse events were determined to be manageable.[1] As of the01 December 2020 data cut-off, 37.3% of patients with NSCLC experienced events relating to hepatotoxicity or renal toxicity. However, this did not result in dose interruption or discontinuation.

Table 3.9: Treatment-emergent adverse events of any severity occurring in ≥10% NSCLC patients in the CodeBreaK100 trial

Phase 2 NSCLC 960 mg daily (N = 126), n (%)	
Preferred Term	
Diarrhoea	62 (49.2)
Nausea	38 (30.2)
Fatigue	32 (25.4)
Aspartate aminotransferase increased	27 (21.4)
Alanine aminotransferase increased	26 (20.6)
Dyspnoea	24 (19.0)
Arthralgia	23 (18.3)
Vomiting	23 (18.3)
Constipation	22 (17.5)
Back pain	20 (15.9)
Anaemia	17 (13.5)
Blood alkaline phosphatase increased	17 (13.5)
Oedema peripheral	17 (13.5)
Cough	16 (12.7)
Decreased appetite	15 (11.9)
Pleural effusion	13 (10.3)
Based on Table 22 of Appendix F of the CS[13] Coded using MedDRA version 23.1; rows are sorted by preferred term in descending order of frequency CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; N = number of participants in the analysis set, n = number of participants with observed data	

Table 3.10: Treatment-related adverse events occurring in ≥5% of NSCLC subjects in CodeBreaK100

Treatment-related adverse events (TRAEs) occurring in > 5%, n (%)	Any Grade N = 126	Grade 3+ N = 126
Any event	**88 (69.8)**	**25 (19.8)**
Diarrhoea	39 (31.0)	5 (4.0)

Treatment-related adverse events (TRAEs) occurring in > 5%, n (%)	Any Grade N = 126	Grade 3+ N = 126
Nausea	24 (19.0)	0
ALT increase	19 (15.1)	8 (6.3)
AST increase	19 (15.1)	7 (5.6)
Fatigue	14 (11.1)	0
Vomiting	10 (7.9)	0
Blood alkaline phosphatase increase	9 (7.1)	1 (0.8)
Maculopapular rash	7 (5.6)	0
Based on Table 18 of the CS[1] ALT = alanine aminotransferase; AST = aspartate aminotransferase; CS = company submission; N = number of participants in the analysis set, TRAE = treatment-related adverse event		

Adverse events (AEs) of any grade, regardless of attribution, were observed in all but one patient (99.2%). The most common AEs included diarrhoea, nausea, fatigue, arthralgia (joint pain), increase in aspartate aminotransferase (ASP) or the alanine aminotransferase levels (ALT). Treatment-related AEs (TRAE) leading to dose modification (dose interruption, reduction, or both) happened in 28 patients (22.2%).[5]

ERG comment: The ERG is concerned with the high number of TEAEs, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.

Table 3.11: Key issue 4. High number of serious adverse events observed in CodeBreaK100

Report Section	3.2.4.5
Description of issue and why the ERG has identified it as important	The ERG is concerned with the high number of treatment-emergent adverse events, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.
What alternative approach has the ERG suggested?	None. The ERG wants to highlight the issue for the committee.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Potential guidance should reflect this issue.
AE = adverse event; ERG = ERG = Evidence Review Group; NSCLC = non-small cell lung cancer	

3.2.4.6 Health-related quality of life (HRQoL)

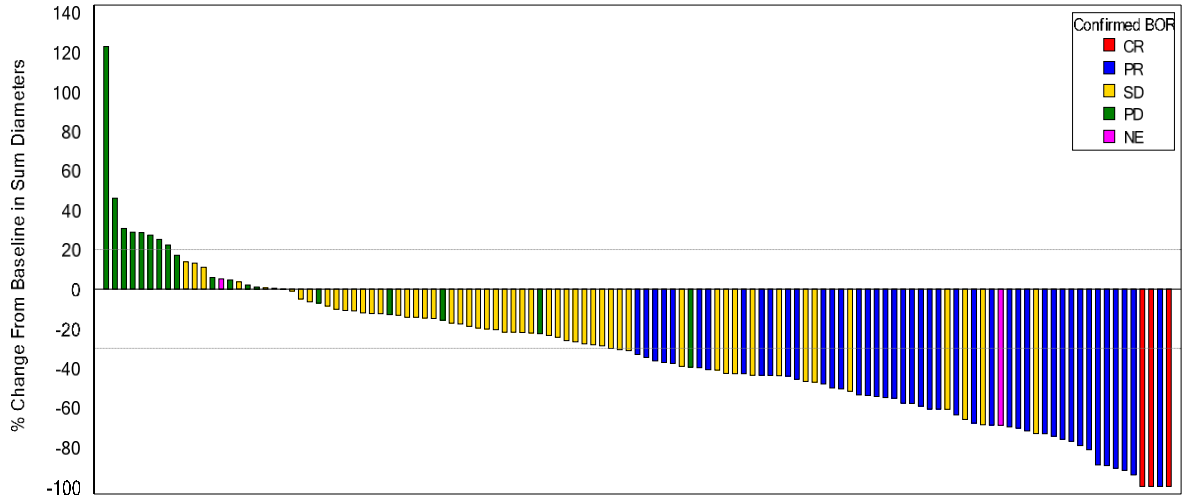
Information related to HRQoL was addressed as an exploratory analysis. For the purpose of the present CS, the European Quality of Life-5 Dimensions-5 Levels (EQ-5D-5L) was used to evaluate the generic health status of the participants.[1] At baseline, most of the participants either reported no problems or slight problems across the EQ-5D-5L health dimensions. However, 33% of participants who had reported either moderate or severe problems or were unable to perform the activity in the pain/ discomfort health dimension.

ERG comment: It is unclear why point estimates were unavailable for HRQoL in general (or PFS specific) for participants on chemotherapy that could be applied in scenario analyses.

3.2.4.7 Disease control rate

According to the CS, the disease control rate comprises of the complete response, partial response, or stable disease.[1] The disease control rate was determined to be high at 80.6% (95% CI 72.6 to 87.2). The CS noted that the percentage of subjects with stable disease was 43.5%. The company also emphasised that not all patients with advanced NSCLC have tumour shrinkage after cancer therapies.[1] Figure 3.3 depicts the tumour shrinkage by best overall response to sotorasib.

Figure 3.3: Waterfall plot of best tumour shrinkage



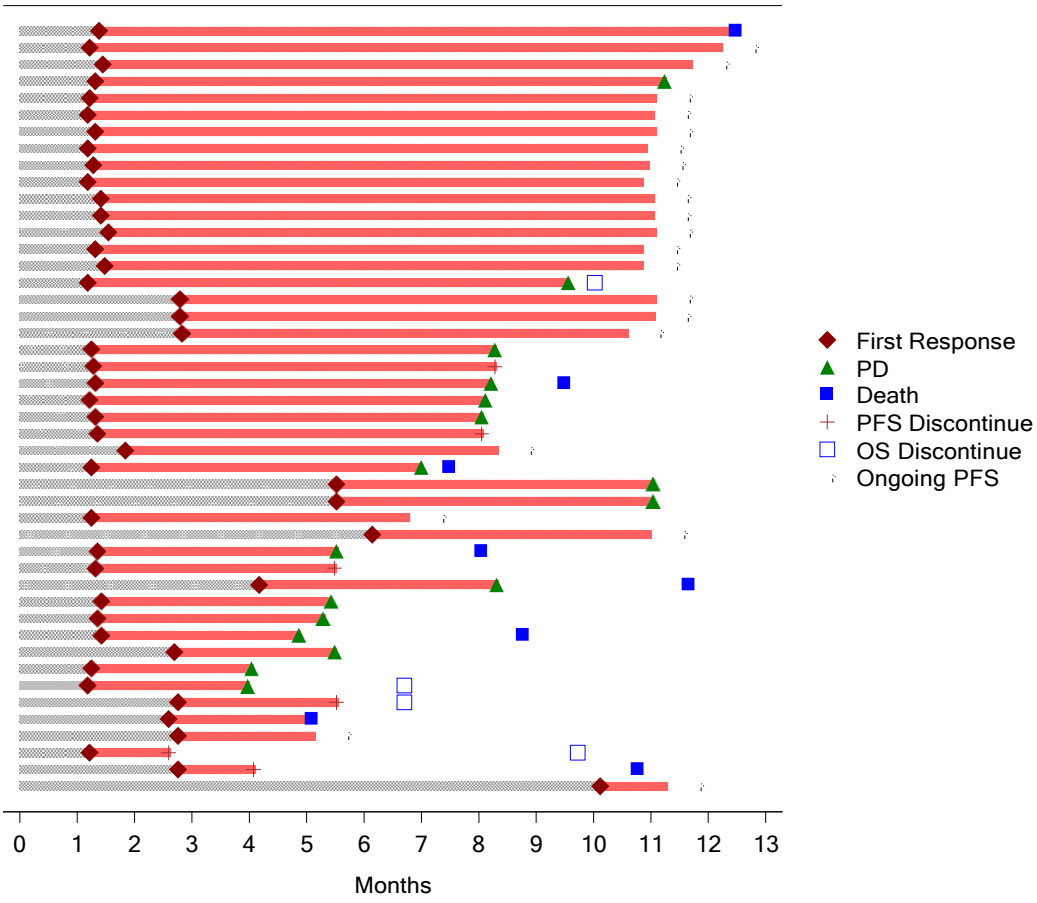
Phase 2 data cut-off date 01DEC2020.
Percent change from baseline in sum of diameters only considers tumor assessments prior to and include the 1st assessment where timepoint response is progressive disease, and prior to start of next anti-cancer therapy.
Two subjects without baseline target lesions and 3 subjects without post-baseline percent changes are not shown.

Based on Figure 3 of the CS[1]

3.4.2.8 Time to response

Among the 46 responders in the NSCLC group, the reported median time to response was 1.35 months within a range of 1.25 to 2.69 months.[1] Figure 3.4 depicts the duration and time to response. However, this is based on the December 2020 data cut-off.

Figure 3.4: Swimmer plot of duration and time to response



Phase 2 data cut-off date 01DEC2020.
'PFS Discontinue' indicates PFS censor due to no post-baseline assessment, withdrew consent, started of new anti-cancer therapy, missed two or more consecutive tumor assessments, off study due to sponsor decision, or lost to follow-up.
'OS Discontinue' indicate OS censor due to withdrew consent, completed study, off study due to sponsor decision, or lost to follow-up.

Based on Figure 3 of the CS[1]
CS = company submission; OS = overall survival; PD = progressive disease; PFS = progression-free survival

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

As detailed in Section 2.3, the CS considered docetaxel monotherapy as the primary comparator for sotorasib (referred to as the primary comparison) while docetaxel in combination with nintedanib was considered as a secondary comparator in patients with adenocarcinoma (referred to as the secondary comparison). As detailed in Section B.2.9 of the CS, the company expects the anticipated conditional approval of sotorasib to be based on the single-arm CodeBreaK100 trial, see Section 3.2 for details of the trial.[1]

The CS identified two studies, SELECT-1 and LUME-Lung 1, as relevant studies to inform an unanchored indirect treatment comparison (ITC) of sotorasib and docetaxel monotherapy and docetaxel combined with nintedanib, respectively, see Table 3.12 for details of the studies of the studies used for ITCs.[1]

Table 3.12: Overview of study designs of CodeBreaK100, SELECT-1 and LUME-Lung 1

Study characteristics	Sotorasib (CodeBreaK100)[15]	Docetaxel monotherapy (SELECT-1)[16]	Docetaxel + nintedanib (LUME-Lung 1)[17]
Blinding	Open label	Double-blinded	Double-blinded
Inclusion criteria	• Male or female patients (≥18 years) • Histologically confirmed locally advanced or metastatic NSCLC • KRAS p.G12C mutation identified through molecular testing • ECOG Performance Status 0 to 1 • ≥1 prior line of systemic anticancer therapy	• Male or female patients (≥18 years) • Histologically confirmed locally advanced or metastatic NSCLC • KRAS-mutation identified through molecular testing • WHO Performance Status 0 to 1 • 1 prior line of systemic anticancer therapy	• Male or female patients (>18 years) • Histologically confirmed locally advanced or metastatic NSCLC • ECOG Performance Status 0 to 1 • 1 prior line of systemic anticancer therapy
Key exclusion criteria	• Active brain metastases • Anti-tumour therapy including chemotherapy, antibody therapy, molecular targeted therapy, retinoid therapy within 28 days of study day 1	• Brain metastases • Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC • Prior treatment with a MEK inhibitor or any docetaxel-containing regimen (prior treatment with paclitaxel is acceptable)	• Active brain metastases • Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC • Prior treatment with a VEGFR inhibitor (other than bevacizumab) or docetaxel
Primary endpoint	Centrally assessed ORR	Investigator-assessed PFS	Centrally assessed PFS
Key secondary endpoints	• Centrally assessed PFS • Investigator-assessed PFS • OS	OS	OS

Based on Table 9 of the CS[1]
CS = company submission; ECOG = Eastern Cooperative Oncology Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; MEK = mitogen activated protein kinase; NSCLC = non-small cell lung cancer; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; VEGFR = vascular endothelial growth factor receptor; WHO = World Health Organization

The CS summarised similarities and differences of these studies:

"CodeBreaK100, SELECT-1 and LUME-Lung 1 were all multicentre studies that recruited patients with confirmed locally advanced or metastatic NSCLC (stage IIIB to IV) who had failed prior therapy. CodeBreaK100 specifically enrolled patients with KRAS p.G12C mutations, whereas SELECT-1 enrolled patients with KRAS mutations at codon 12, 13 or 61.[16] LUME-Lung 1 did not specify KRAS mutations as an enrolment criterion and did not record KRAS mutations among the participants; however, in the subpopulation of interest (licensed population of patients with adenocarcinoma) the proportion of patients with KRAS p.G12C mutations is likely close to the prevalence of KRAS p.G12C mutations in the general non-squamous population (~13%). CodeBreaK100 enrolled patients with 1 to 3 prior therapies, whereas SELECT-1 and LUME-Lung 1 included patients with 1 prior therapy. All studies excluded subjects with active brain metastases, although CodeBreak100 and LUME-Lung 1 permitted inclusion of stable brain metastases.

All three studies reported PFS and OS as primary or secondary endpoints. PFS was assessed by investigators in SELECT-1, by both independent central review and by investigator in CodeBreaK100 and by independent central review in LUME-Lung 1".

Table 3.13 gives an overview of the baseline characteristics of these studies.

Table 3.13: Comparison of baseline characteristics in CodeBreaK100, SELECT-1 and LUME-Lung 1 trials

Baseline characteristics[a]	Sotorasib (CodeBreaK100) N=126[15]	Docetaxel monotherapy (SELECT-1) (N=256)[16]	Docetaxel + nintedanib (LUME-Lung 1) (N=322)[i17]
Age	62.9 (mean)	60.9 (mean)	58.5 (median)
Gender (% female)	50%	43%	37%
Brain metastases (%)	21%	NR[c]	8%
Performance status (ECOG or WHO; % PS 1 [vs PS 0])	70%	59%	70%
Race (% white)	82%[d]	95%	NR[g]
% KRAS p.G12C-mutated	100%	42%[b]	NR[h]
Anti-PD-(L)1 in prior line(s)	91%	0%	0%
Number of prior lines (% with 1/2/3 prior lines)	43%/35%/22%	100%/0%/0%	Mostly 1 prior line[i]
Metastatic disease at baseline	96%	96%	90%
Histology (% non-squamous)	99%	95%	100%[j]
Smoking status (% ever smoker)	93%[e]	92%	64%
Other targetable mutations (EGFR, ALK, BRAF, ROS-1)	3%	NR[f]	NR
PD-L1 expression at baseline (<5% [vs >5%])	48%	58%	NR

Baseline characteristics[a]	Sotorasib (CodeBreaK100) N=126[15]	Docetaxel monotherapy (SELECT-1) (N=256)[16]	Docetaxel + nintedanib (LUME-Lung 1) (N=322)[i17]
Based on Table 10 of the CS[1] [a] all reported baseline characteristics in SELECT-1 and other key characteristics; [b] the rest of the population has KRAS mutations other than G12C; [c] not reported for SELECT-1. All studies had exclusion criteria for active brain metastases; [d] 15 percentage points of the 18% remaining correspond to Asian patients; [e] 2 percentage points of the remaining 7% are missing data; [f] probably very low due to KRAS mutant; [g] Race was not reported, the trial was non-US based and run mainly in Europe (71% of patients) as well as Asia; [h] LUME-Lung 1 did not enrol by or record genetic mutations; the % of KRAS p.G12C is likely close to the prevalence of KRAS p.G12C mutations in the general non-squamous population (~13%); [i] LUME-Lung 1 included patients with a prior platinum-based therapy and allowed adjuvant/neoadjuvant as line of therapy; [j] Based on the subpopulation of interest (adenocarcinoma) ALK = anaplastic lymphoma kinase; BRAF = B-Raf Proto-oncogene; CS = company submission; ECOG = European Co-operative Oncology Group; EGFR = epidermal growth factor receptor; KRAS = Kirsten rat sarcoma viral oncogene homolog; NR = not reported; PD-L1 = programmed death-ligand 1; PS = performance status; ROS = proto-oncogene tyrosine-protein kinase; WHO = World Health Organization			

The CS highlighted that *"as LUME-Lung 1 enrolled patients with mixed histology,[17] but nintedanib in combination with docetaxel is only licensed for use in patients with adenocarcinoma,[18] only the characteristics of the adenocarcinoma subpopulation of LUME-Lung 1 are considered"*.[1]

Overall, the distribution of patients between the three trials is similar in terms of age, disease stage and histology, and the majority of patients had ECOG/WHO performance status of 1.

However, the CS highlighted a few differences between the studies which arose from the different time at which these were conducted:[1]

1. G12C KRAS mutation status, i.e. 100% in CodeBreaK100, 42% in SELECT-1 (remaining patients had other KRAS mutations), and not reported for LUME-Lung 1.
2. CodeBreak100 included patients taking 1-3 prior therapies and a high proportion of patients who had prior use of PD(L)-1 inhibitors, reflecting the current treatment pathway for patients with KRAS p.G12C -mutated NSCLC in the UK. In contrast, the SELECT-1 and LUME-Lung 1 trials, which were both conducted before the evidence base supported front-line use of immunotherapy, included patients taking 1 prior therapy only and no PD(L)-1 inhibitors.
3. Based on inclusion criteria and/or a lack of recording, it is also not possible to compare for the presence of (non-active) brain metastases in SELECT-1, for the PD-1 expression in LUME-Lung 1, or for the presence of other targetable mutations in either of these comparator trials.
4. It is also of note that LUME-Lung 1 recruited fewer females, fewer prior smokers and patients with fewer brain metastases than CodeBreaK100.

According to the CS, UK clinical experts considered these were the best and most relevant sources of data available with which to make indirect comparisons for sotorasib in patients with KRAS p.G12C mutated NSCLC.[1, 19]

The company also used the Flatiron study as an alternative data source for the primary comparison.[20] The reason given for this being only supplementary was that docetaxel was only used in a minority of patients and that about 24% of patients had received prior first-line immunotherapy.

ERG comment: The ERG noted the differences between these studies. As discussed in more detail in Section 3.4, it is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC.

It is not entirely clear that SELECT-1 was a better data source than Flatiron. This is not least because individual participant data were available to perform what the company called a 'Propensity Score Weighted Analysis (PSWA)', which appears to be an Inverse Probability Weighting (IPW) analysis according to technical support document (TSD) 17, for the latter such that the comparator data could be adjusted to be more like the intervention population.[21] However, the size of reduction in any bias would depend on the degree to which prognostic factors could be identified, either from the CodeBreaK100 data for the MAIC or the Flatiron data for the PSWA.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

3.4.1 Matching of prognostic patient characteristics

As noted in Section 3.3, there are differences between the three studies considered for indirect comparison, CodeBreaK100, SELECT-1 and LUME-Lung 1.

For the primary comparison (vs. docetaxel), *"an MAIC was used to compare changes in OS and PFS with sotorasib versus docetaxel monotherapy"* (see Section B.2.9.3.1 of the CS).[1]

The company rejected the use of a MAIC for the secondary analysis (vs. nintedanib plus docetaxel in adenocarcinoma). The reasons given was that: *"…the differences in patient characteristics and data availability for matching would present significant challenges, including reducing the effective sample size and precision for relative treatment effect estimates and introducing a further population that is less closely aligned with CodeBreak100 and not aligned with the SELECT-1 trial population to which sotorasib recipients in CodeBreaK100 had already been matched"* (see Section B.2.9.3.2 of the CS).[1]

According to the CS, *"further examination of the LUME-Lung 1 data indicated that a piecewise approach to hazard ratio estimation would be required, and estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel could only be made following implementation of these data within the economic model"* (further details in Section B.3.3.5 of the CS).[1]

3.4.1.1 Choice of covariables for MAIC

In response to question A27 of the the request for clarification, the company gave details on the choice of covariables for the MAIC, stating that *"1) the populations being compared are defined similarly in terms of ECOG performance status at baseline and 2) adjustment based on ECOG score is performed to ensure the populations being compared are balanced. Other factors that were indicated by the physicians to be very important by a majority of physicians to assess the prognosis or response to treatment of patients included presence of brain metastases (ideally distinguishing between active and controlled brain metastases), disease stage at baseline (stage IIIb/c vs IV or IIIb-IVa vs IVc). A number of other factors were considered as being at least somewhat important for prognosis and should be considered when the information is available. Finally, age and gender, although not consistently considered as prognostic or predictive factors, were mentioned as key covariates to include in an adjusted comparative effectiveness analysis"*.[6]

Table 11 of the CS presented the starting list of prognostic covariates (five classified as "very important", 13 classified as "somewhat important", and three "additional covariates reported in other MAIC analyses").[1]

"Of these 21 potential covariates, 8 were selected for inclusion in the MAIC analysis based on data availability in SELECT-1, their prognostic importance for patients receiving sotorasib and docetaxel, and the feasibility of matching whilst preserving the effective sample size", as detailed in Table 12 of the CS, namely:

1. ECOG (% PS 1 [vs. PS 0])
2. Age (mean)
3. Metastatic disease stage at baseline
4. Smoking status (% ever smoker)
5. PD-L1 expression level
6. Gender (% female)
7. Histology (% Non-squamous)
8. Race (% white).[1]

ERG comment: It should be noted that some factors identified by the clinical experts, such as presence of brain metastases and disease stage at baseline, as well as other factors such as KRAS p.G12C mutation status were not considered for the MAIC comparing CodeBreaK100 and SELECT-1.

3.4.1.2 KRAS p.G12C mutation status

The CS stated that *"a propensity score weighted analysis approach such as MAIC requires the matching of prognostic patient characteristics to generate robust comparative treatment effect estimates. Due to missing data or other differences between the trials it would not be possible to matchacross all trials for KRAS p.G12C mutation status, brain metastases, prior lines of therapy or prior useof PD-L1 inhibitors. Given that PFS and OS outcomes are similar in the absence of targeted therapies,irrespective of KRAS status (see section B.2.9.2.1* [of the CS]*), the inability to match by specific KRASstatus is unlikely to lead to biased estimates"*.[1]

In response to the request for clarification (question A24), the company replicates Table 3 of the CS to support the view that *"given the OS and PFS for All NSCLC patients are highly consistent to those for patients with KRAS mutations (and the KRAS mutation datasets are included in the All NSCLC dataset),it is reasonable to conclude that the results are consistent for those with and without KRAS mutations"*.[1, 6]

However, the company (in response to question A26b), highlighted that *"as targeted therapy for KRAS mutated NSCLC did not exist at the time of the LUME Lung 1 trial, and screening for KRAS mutant NSCLC was not routine practice, we have no means of knowing the KRAS mutant status of patients enrolled in the LUME Lung 1 trial. It was for this reason that it was not possible to match the LUME Lung 1 trial participants and CodeBreaK100 trial participants in an MAIC"*.[6]

ERG comment: Table 3 of the CS does show that median OS and PFS do appear to vary little between KRAS p.G12C-mutated and KRAS-mutated (non-p.G12C) NSCLC, e.g. first line: 12.0 (9.6, 15.3) vs. 12.2 (10.5, 14.4) and 5.0 (4.4, 5.8) vs. 5.6 (5.4, 6.0), respectively.[1, 6] However, despite the consequent increase in uncertainty by using only the 42% of patients with KRAS p.G12C mutation, an analysis with mutation status as covariate could be informative.

3.4.1.3 Brain metastases

The CS stated that not matching on brain metastases between CodeBreaK100 and SELECT-1 is *"unlikely to introduce significant bias"*.[1] In response to the request for clarification (question A28), the company elaborated on this point:[6]

"The proportion of patients with brain metastases was higher in CodeBreaK100 (21%) than in LUME Lung-1 (8%). The proportion with brain metastases in SELECT-1 was not reported. However, all three trials excluded patients with active (or symptomatic) brain metastases. (…) *As CodeBreaK100 enrolled a high proportion of patients with brain metastases, and somewhat higher than in patients recruited to LUME Lung 1, it is a reasonable assumption that SELECT-1 did not include a higher proportion of patients with non-active brain metastases than CodeBreaK100. Any negative influence on survival of the presence of brain metastases would therefore impact on the CodeBreak100 population to a greater extent than on the populations in LUME Lung 1 or SELECT-1. Therefore, the results of the comparison of sotorasib (from CodeBreaK100) vs nintedanib plus docetaxel (from LUME Lung 1) or docetaxel monotherapy (from SELECT-1) would favour the comparators. On this collective basis, our inability to match for brain metastases between CodeBreaK100 and SELECT-1 is unlikely to introduce bias in favour of sotorasib and is more likely to be conservative".*[6]

ERG comment: Although active brain metastases were excluded from all three trials, the presence of brain metastases did seem to affect prognosis as indicated by the subgroup analyses reported in Appendix E.[13]

In particular, median OS was not estimable for no metastases and percentage surviving to 12 months was 55.5 (44.8, 64.9) compared to 35.3 (23.4, 48.4) for presence of metastases. The company claim that, because the percentage was a lot higher for CodeBreaK100 than LUME-Lung 1 then it must also be higher than for SELECT-1, so that not adjusting for brain metastases is favourable to the comparator.

However, the ERG would regard this as speculation and therefore there is no way of knowing the effect of not adjusting for brain metastases on outcome.

3.4.1.4 Other baseline characteristics

In response to the request for clarification (questions A25a and A25b), the company confirmed that a number of factors, such as country of origin, socio-economic status, comorbidities, year of recruitment and number/severity of metastases as well as age, gender, smoking status, geographic region/ethnicity/race, body mass index/weight or history of alcohol abuse, were not considered as, based on a physician's assessment, none of these factors was found to be *"very important to consider"* (i.e. *"at least 4 of the 6"* physicians highlighting the importance).

The ERG noted differences in the smoking rate in LUME-Lung 1 study compared to CodeBreaK100 and SELECT-1 (64% versus 93% and 92%) and asked for clarification (question A26). In response, the company stated to *"not know the reasons for why the smoking history of patients enrolled in LUME Lung 1 was different to that in CodeBreaK100 and SELECT-1; however, it can be seen in the LUME Lung 1 trial results that progression-free survival (PFS) and overall survival (OS) were not significantly different between patients with or without a history of smoking (see Figure 4 in Reck 2014,which refers to the adenocarcinoma population)".*[6]

ERG comment: The ERG noted that a number of factors were not considered. While the clinical experts consulted by the company agreed with that approach, there remains uncertainty to the impact matching these factors would have had.

3.4.1.5 Standard of care

In response to a request to clarify whether the standard of care for NSCLC is likely to be equivalent between the studies (question A25c), the company stated that *"CodeBreaK100 subjects were more heavily pre-treated than SELECT-1 subjects, with SELECT-1 and LUME-Lung 1 patients receiving*

only 1 previous line of systemic anticancer therapy", concluding that *"this is a conservative limitation for the comparative effectiveness as a more heavily pre-treated population is generally associated with poorer clinical outcomes"*.[6] **ERG comment:** The assessment by the company is likely to be correct, however, this adds to the uncertainty linked to the ITCs.

3.4.2 PSWA using Flatiron

The company stated that they used a PSWA, which most closely resembles an IPW according to TSD 17.[21] The details of the method are reported in Appendix D.[13]

The weights were applied only to the comparator data, which effectively implies the estimation of average treatment effect of the treated (ATT) as opposed to average treatment effect (ATE), thus limiting applicability to the population of patients who received sotorasib as opposed to any in the index population.[21]

Firstly, Flatiron patients were selected to align with the CodeBreaK100 eligibility criteria:

- Diagnosis of advanced NSCLC between 01 January 2011 and index date
- First positive test for KRAS mutation no later than 21 days after index date (to avoid introducing immortal time bias in the analyses)
- Age 18 years or older at index date
- Started the selected line of treatment on/before 31 March 2020 (to allow sufficient opportunity for a follow-up time of at least 6 months)
- Structured electronic health record activity in the first 90 days after the date of advanced NSCLC diagnosis
- Previous treatment with at least one prior line of therapy containing anti-PD-1 or anti-PD-L1 immunotherapy and/or platinum-based chemotherapy
- Selected line of therapy does not contain a clinical study drug
- Selected line of therapy is not the patient's first line of treatment containing an anti-PD-(L)1 component
- Baseline ECOG performance status ≤1

In addition, the following selection rule was applied to determine which line of therapy was considered for the control cohort:

- If a patient had received between 2 and 4 (inclusive) lines of therapy on or before 31 March 2020, the latest line of therapy which met the inclusion criteria was selected.
- If a patient had received more than 4 lines of therapy on or before 31 March 2020, the 4th line was selected (unless that line of therapy did not meet the inclusion criteria, in which case the most recent eligible treatment line was included).
- If no line of treatment met the inclusion/exclusion criteria, the patient was not included in the analysis.

All platinum-based chemotherapy patients and not only those who had taken docetaxel monotherapy were included. As shown in Table 12, Appendix D, there were about 31% of the former and 10% (n=21) of the latter in the KRAS mutant population with about 29% and 13% (n=11) respectively in the KRAS p.G12C mutant population.[13]

The process of covariate selection started with the same set as for the MAIC, as shown in Table 11 of the CS:[1]

- 'Very important' covariates were included expect PD-L1 status due to 98.7% of values being missing in Flatiron.
- 'Somewhat important' covariates were included (except for eGFR, again due to missing data (38.7%)) on the basis of a "*stepwise variable selection algorithm*", which was not clearly explained.

The list of included covariates is shown in Table 10 of Appendix D, which showed the effect of adjustment.[13] Figure 6 showed the standardised differences in covariates between CodeBreaK100 and Flatiron.[13] This shows that adjustment reduced those differences to close to zero in the KRAS mutant population. However, the standardised differences remained above 0.1 for several covariates and above 0.2 for liver metastases, one prior line of therapy and two age groups in the KRAS p.G12C population. This and the small effective sample size were the reasons given for preferring the KRAS mutation population.

ERG comment: Estimation of ATT as opposed to ATE might be an issue depending on the degree of heterogeneity of treatment effect and applicability of the CodeBreaK100 trial. An analysis applying the propensity score weights to all patients could be informative. Also, there are methods other than IPW, such as regression adjustment (RA) or doubly robust (RA plus IPW), that could have been employed and so scenario analyses using these methods could also be informative.[21]

Selecting patients to align with the CodeBreaK100 trial is in principle a good idea. However, given that sotorasib is to be positioned for 2nd line or later, it is not clear to the ERG why patients only at 4th line were selected. Although patient numbers are small, it might have been informative to see results for the docetaxel monotherapy population.

The process of covariate selection was not entirely clear and would therefore benefit from further explanation. It did appear that better balance was achieved for the KRAS population and, as discussed above, it might be reasonable to consider the prognosis similar to the KRAS p.G12C population.

In conclusion, the ERG considers that there might be reasons to believe that the results of the PSWA (using Flatiron) are less biased than those of the MAIC (using SELECT-1) given that:

1. The PSWA adjusted the Flatiron data to make more comparable to the CodeBreaK100 population: the benefit of this lies in CodeBreaK100 being more applicable to the patients that might be treated in the UK with sotorasib, which is uncertain
2. Very little difference in effective sample size (104.8 for Flatiron in the KRAS population in the PSWA vs. OS/PFS 108.8/106.1 for CodeBreaK100 in MAIC primary analysis)
3. The MAIC primary analysis only adjusted for four covariates, which excluded brain metastasis, as opposed to 13 in the PSWA, which included brain metastasis

However, there remains considerable uncertainty in the effectiveness of sotorasib vs. docetaxel that might be to some extent addressed by further analysis using the Flatiron data as described above.

Table 3.14: Key issue 5. Validity of ITC

Report Section	3.3, 3.4
Description of issue and why the ERG has identified it as important	The ITC is unanchored i.e. no common comparator. Therefore, there are potentially relevant differences in prognostic factors between the studies included in the ITCs (CodeBreaK100, SELECT-1, LUME-Lung 1), e.g. regarding G12C KRAS mutation status, prior therapies, presence of brain metastases, and factors like sex and smoking history. It is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC. The company chose a MAIC for their primary analysis of the main comparison with docetaxel, which is particularly prone to bias given lack of identification of all relevant prognostic factors and clinical experts identified factors to be "very important", e.g. brain metastases and disease stage at baseline. However, these, alongside G12C mutation status, were not considered for the MAIC comparing CodeBreaK100 and SELECT 1. Also, because only summary statistics were available from SELECT-1, the CodeBreaK 100 had to be adjusted to match the SELECT-1 population. The company also conducted a supplementary analysis using the Flatiron study, which, using a method of adjustment, referred to as PSWA that appears to involve IPW allowed the comparator data to match the CodeBreaK100 population. A richer set of individual patient data also afforded a greater number of potential prognostic factors. In addition to the underlying uncertainty introduced by an indirect comparison of treatments (compared to a direct comparison), the differences between studies, the choice of baseline variables for matching, the choice of underlying data source and adjustment method can be questioned and the ERG would have liked to see further analyses.
What alternative approach has the ERG suggested?	1. For the MAIC, an analysis with mutation status as covariate could be informative 2. For the PSWA, methods other than IPW, such as RA or doubly robust (RA plus IPW), could have been employed and so scenario analyses using these methods could be informative 3. For the PSWA, limiting to the docetaxel only population could be informative 4. In principle, evidence directly comparing treatments would provide more robust evidence.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	See suggestions above.

ERG = Evidence Review Group; G12C = G12C amino acid substitution; IPW = inverse probability weighting; ITC = indirect treatment comparison; KRAS = Kirsten rat sarcoma viral oncogene homolog; MAIC = matching adjusted indirect comparison; PSWA = propensity score weighted analysis; RA = regression adjustment

3.4.3 Results of indirect comparison

3.4.3.1 Primary analysis – MAIC using CodeBreaK100 and SELECT-1

Results for the primary analysis were reported in Section B.2.9.4.1.[1] Table 3.15 provides an overview of results while Figures 3.5 and 3.6 show Kaplan-Meier plots of the primary MAIC analysis for OS and PFS, respectively.

Table 3.15: Results of MAIC for primary comparison of sotorasib vs docetaxel monotherapy

Analyses	CodeBreaK100 N (OS / PFS)	CodeBreaK100 ESS (OS / PFS)	Median OS Sotorasib vs. Docetaxel	Median PFS Sotorasib vs. Docetaxel
Unadjusted	126	126	[redacted]	[redacted]
MAIC Model: "all variables of prognostic importance" (Primary analysis)	123/ 121	108.8/ 106.1	[redacted]	[redacted]
MAIC Model: "all available covariates" (sensitivity analysis)	98/ 96	53.3/ 53.1	[redacted]	[redacted]

Based on Table 14 of the CS[1]
* Median OS not reached, OS was 50.4% at 12.5 months; ¶ Median OS not reached, OS was 52.5% at 12.0 months
CI = confidence interval; CS = company submission; ESS = effective sample size; HR = hazard ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival

Figure 3.5: Kaplan-Meier plot for primary MAIC analysis of OS for sotorasib and docetaxel monotherapy

Based on Figure 7 of the CS[1]
CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching adjusted indirect comparison; OS = overall survival

Figure 3.6: Kaplan-Meier plot for primary MAIC analysis of PFS for sotorasib and docetaxel monotherapy

Based on Figure 8 of the CS[1]
CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching adjusted indirect comparison; PFS = progression-free survival

3.4.3.2 Supplementary primary comparison – Propensity score weighting analysis using CodeBreaK100 and Amgen Flatiron Health real-world evidence study

As described in Section B.2.9.4.1, *"this supplementary analysis was undertaken to explore an alternative data source and method of estimating relative treatment effects for sotorasib vs docetaxel monotherapy (using the basket of standard of care chemotherapy regimens in the Amgen Flatiron real-world evidence cohort as a proxy for docetaxel monotherapy)"*.[1] Results are presented in Table 3.16.

Table 3.16: Results of supplementary primary comparison using propensity score weighting analysis

Outcome	Flatiron N before adjustment	KRAS mutant		KRAS-p.G12C mutant subgroup	
		ESS	Median HR (95% CI)	ESS	Median HR (95% CI)
Overall survival	206	104.8	[redacted]	17.8	[redacted]
Progression-free survival	206	104.8	[redacted]	17.8	[redacted]

Based on Table 15 of the CS[1]
CI = confidence interval; CS = company submission; ESS = effective sample size; HR = hazard ratio; KRAS = Kirsten rat sarcoma viral oncogene homolog

3.4.3.3 Secondary comparison implemented in the economic model

According to Section B.2.9.4.2 of the CS, an *"estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel was implemented in the economic model"*.[1] Table 3.17 presents the results for the secondary comparison while Section 4.2.6.6 provides a critique of the approach.

Table 3.17: Results of secondary comparison implemented in the economic model

	Sotorasib	Nintedanib plus docetaxel	Increment
Mean OS (months)*	[redacted]	[redacted]	[redacted]
Mean PFS (months)*	[redacted]	[redacted]	[redacted]

Based on Table 16 of the CS[1]
* Derived from economic model with 20-year time horizon, undiscounted values (see Section B.3.3.5 of the CS for details on the implementation)
CS = company submission; OS = overall survival; PFS = progression-free survival

3.5 Additional work on clinical effectiveness undertaken by the ERG

As detailed in Section 3.2.3, the ERG re-assessed the risk of bias of the CodeBreaK100 study using ROBINS-I.[14]

3.6 Conclusions of the clinical effectiveness Section

As the clinical effectiveness searches were run in June 2020 and updated on 26th January 2021, the ERG considers it likely that all potentially relevant studies were included in the systematic review. However, the ERG remains concerned about the application of English language restrictions and a lack of validated search filter for RCTs which both could negatively impact on the comprehensiveness and precision of the company's clinical effectiveness review.

The ERG has identified some inconsistencies in the study selection process that potentially introduce bias. For instance, exclusion of non-RCTs or phase I trials is questionable and based on the company's assumption that the evidence-base is limited. The ERG did not identify any issues with regards to data extraction.

The CodeBreaK100 study was a single arm, multicentre, non-randomised, open-label, phase II study.[5] Therefore, due to the absence of a comparator arm, the interpretation of the results is problematic. The study did not include a single centre from the UK which indicates generalisability of the CodeBreaK100's findings into clinical practice in England and Wales. It is not clear how participants at phase I of the trial were handled in the analyses as they used a combination of sotorasib and anti PD-1/L1 or midazolam. The ERG also undertook its own risk of bias assessments and found some serious limitations in the CodeBreaK100 study.

As there was no comparative trial data, the only available analysis was an unanchored ITC between sotorasib and a) docetaxel monotherapy (SELECT-1) and b) docetaxel + nintedanib (LUME-Lung1).[16, 17] The ERG highlighted a few dissimilarities between the studies and stressed that it is not possibleto match for all of these differences which potentially impacts validity of the findings of any ITC. The ERG also believes that the results of the MAIC (using SELECT-1) are potentially more biased than an alternative approach using PSWA (based on Flatiron data).

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

This Section pertains mainly to the review of cost effectiveness analysis studies. However, the search Section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the CS. Therefore, the following Section includes searches for the cost effectiveness analysis 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 cost effectiveness Section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.

Searches for cost effectiveness analysis review

Appendix G of the CS detailed a SLR conducted to identify published studies evaluating cost effectiveness, costs and resource use and HRQoL for treatments in NSCLC.[13] Searches were undertaken on 20 February 2020 and updated on 29 January 2021. Searches for costs and healthcare resource use were restricted to 2009 onwards. An English language restriction was reported but this was not applied at the searching stage. A summary of the sources searched is provided in Table 4.1.

Table 4.1: A summary of the sources to identify cost effectiveness studies

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations	Ovid	1946 to present	20 February 2020 29 January 2021
MEDLINE Daily, MEDLINE and Versions			
Embase		1974 to present	20 February 2020 29 January 2021
• CDSR • DARE • CENTRAL • CMR • NHS EED • HTA Database			20 February 2020 29 January 2021

Resource	Host/Source	Date Ranges	Dates searched
• ACP			
Congress searches			
ASCO	https://www.asco.org/	2017 - 2020	
ESMO	http://www.esmo.org/		
WCLC	https://wclc2019.iaslc.org/		
AACR	https://www.aacr.org/Pages/Home.aspx		

AACR = American Association of Cancer Research; ACP = American College of Physicians; ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CMR = Cochrane Methodology Register; DARE = Database of Abstracts of Reviews of Effects; EED = Economic Evaluations Database; ESMO = European Society for Medical Oncology; HTA = Heath Technology Assessment; NHS = National Health Service; WCLC = World Conference on Lung Cancer

ERG comment: Searches were undertaken for a SLR to identify all cost effectiveness, HRQoL and cost and resource use studies. The CS provided sufficient details for the ERG to appraise the literature searches. A range of databases and conference proceedings were searched as well as previous NICE submissions for disease management costs.

The search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use. A date limit of 2009 was applied to searches for health economics but this was considered appropriate. As for clinical effectiveness searches, the strategies between Embase, MEDLINE and the Cochrane Library were not modified in all cases to take account for differences in thesaurus headings. However, the ERG was satisfied that the sufficient use of free-text terms compensated for this failure.

The use of filters in NHS EED may have been overly restrictive as this database is topic specific. However, as NHS EED is no longer being updated, the ERG is satisfied that anything of relevance is unlikely to have been missed.

4.1.2 Inclusion/exclusion criteria

Inclusion and exclusion criteria for the review on cost effectiveness studies, HRQoL studies and costs and resource use studies are presented in Table 4.2.

Table 4.2: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	NSCLC patients with KRAS mutated (further specification not required) with a primary interest in $KRAS^{G12C}$	Known KRAS mutation-negative status
Intervention	Any	Drug targeted to ALK, BRAF, EGFR, NTRK, or ROS1 (unless a KRAS mutated NSCLC comparator group is included)
Comparator	Any or none	N/A
Outcomes(s)	- Health-related quality of life - Quality-adjusted life-years gained	Any other

	Inclusion criteria	Exclusion criteria
	- Progression-free life-years gained - Life-years gained - Treatment cost by stage of disease (e.g., pre-progression vs. post-progression), including healthcare resource use, cost of care, cost of illness - Health state utilities - Economic evaluations	
Study design	Any	Animal/in vitro studies, case studies, and case reports
Date restrictions	***Costs/healthcare resource use*** - 2009 to present ***HRQoL and economic evaluation*** - No limit	
Language restrictions	English language	
Publication type	All primary publications and systematic reviews	Non-systematic reviews, editorials, notes, and letters
Country	Not restricted	
Based on Appendix G of the CS[13] ALK = anaplastic lymphoma kinase; BRAF = B-Raf Proto-oncogene; CS = company submission; EGFR = epidermal growth factor receptor: HRQoL = health-related quality of life; KRAS = Kirsten rat sarcoma viral oncogene homolog; N/A = not applicable; NSCLC = non-small cell lung cancer; NTRK = neurotrophic tyrosine kinase; ROS = proto-oncogene tyrosine-protein kinase		

ERG comment: The eligibility criteria used by the company provided sufficient detail and appeared to be appropriate.

4.1.3 Conclusions of the cost effectiveness review

Searches were undertaken for a SLR to identify all cost effectiveness, HRQoL and cost and resource use studies. The CS provided sufficient details for the ERG to appraise the literature searches. The search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use. As for clinical effectiveness searches, the strategies between Embase, MEDLINE and the Cochrane Library were not modified in all cases to take account for differences in thesaurus headings.

No published economic studies were identified in the SLR which examined the cost effectiveness of interventions for the management of patients with KRAS p.G12C mutation-positive locally advanced or metastatic NSCLC or for KRAS mutation in general. Also, no relevant studies on HRQoL to inform the decision problem were identified.

The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies. Eligibility criteria were suitable for the SLR performed.

According to the CS, the SLR identified 14 studies reporting costs and healthcare resources used in patients with NSCLC and a KRAS p.G12C mutation.[1] Of these, 13 were on costs associated with biomarker testing which was not considered relevant for this appraisal. The company concluded that

the studies identified in the SLR on costs and healthcare resource use did not provide adequate costs and resource use valuations which were useful to a UK clinical setting, although it was not clear from Appendix I and the CS why the 14th study was not relevant.[1, 13]

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Direct health effects for patients included
Perspective on costs	NHS and PSS	NHS and PSS
Type of economic evaluation	Cost utility analysis with fully incremental analysis	Cost utility analysis, two separate analyses for two comparators – hence no full incremental analysis was performed, because the populations were not comparable.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The time horizon of 20 years is considered long enough to reflect all relevant differences in costs and outcomes.
Synthesis of evidence on health effects	Based on a systematic review	Systematic review conducted to identify additional evidence on health effects beyond trial data. However, none of the studies found pertained to the KRAS p.G12C mutation.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Health effects were expressed in QALYs. Quality of life was measured with EQ-5D-5L and mapped to the EQ-5D-3L.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Reported directly by patients.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Crosswalk – representative sample of the UK population.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No equity issues have been identified.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be	The model includes the costs that relate to NHS and PSS

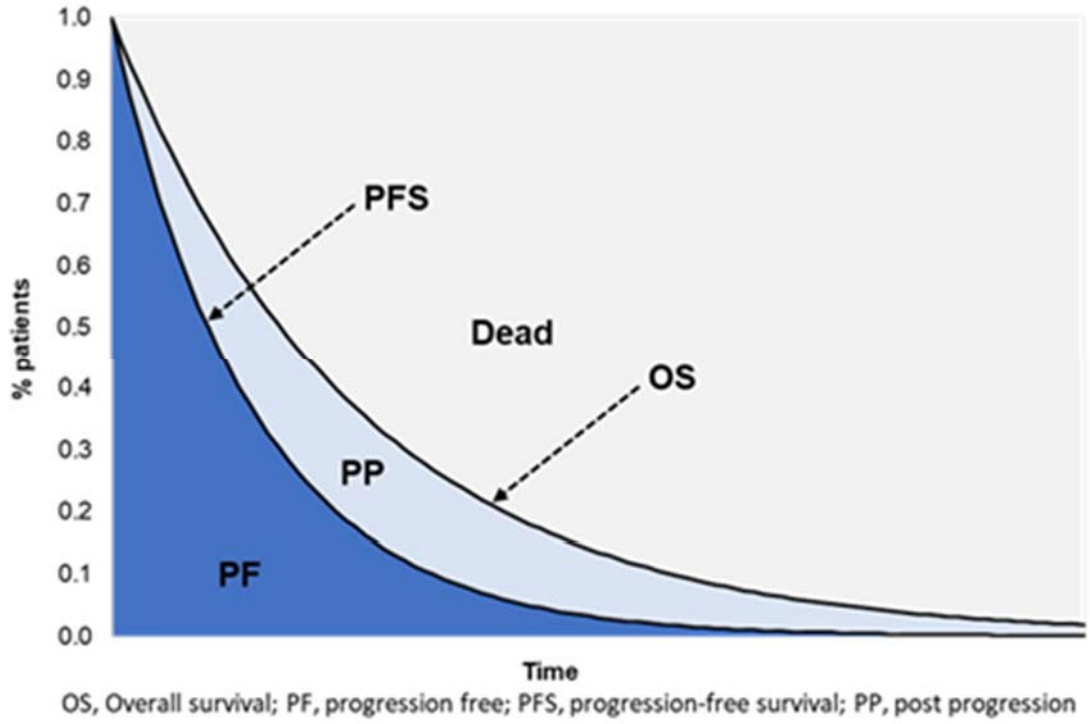
Element of health technology assessment	Reference case	ERG comment on company's submission
	valued using the prices relevant to the NHS and PSS	resources, valued using the prices relevant to the NHS and PSS.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at 3.5%.
EQ-5D = European Quality of Life-5 Dimensions; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom		

4.2.2 Model structure

The company developed a cohort-level partitioned survival model (PSM) in Microsoft Excel, with the following three health states: progression free (PF), post progression (PP) and death. The proportions of patients in each health state at the beginning of each model cycle are calculated from the PFS and OS curves from relevant clinical trials. In the model, all patients start in the progression free health state and on treatment.

Figure 4.1 shows the model structure of the partitioned survival model.

Figure 4.1: Model structure



Based on Figure 9 of the CS[1]
CS = company submission

ERG comment: The main concern of the ERG relates to the use of a PSM without a state transition model (STM) alongside it to validate the model structure. The company stated that the model structure applied was fully aligned with the primary objectives of treatment in oncology and NSCLC, namely avoiding disease progression and prolonging life, and that all relevant health states were included.[6]

The ERG considers this to be not an exclusive feature of a PSM, an STM would have aligned fully with these objectives as well and could have included the same health states. Therefore, the ERG requested the company to provide an STM as a scenario for validation purposes, as recommended in NICE Decision Support Unit (DSU) TSD 19.[22] In response to the request for clarification, the company stated that they considered it to be a recommendation in TSD 19 that an STM should be accompanying a PSM for validation.[6] The company also stated they believed an STM would not overcome the potential downsides of a PSM and that the scenarios provided would explore these sufficiently.

The ERG acknowledges that every model approach has its limitations but is still concerned that the consequences of choice of model structure may not be fully overseen because all choices and scenarios implemented follow this chosen structure. Size and direction of bias (if any) associated with choice of model structure cannot be estimated in the absence of alternative approaches.

Table 4.4: Key issue 6. Partitioned Survival Model structure not validated or justified

Report Section	**4.2.2**
Description of issue and why the ERG has identified it as important	The company used a partitioned survival model without elaborate justification and without an accompanying scenario implementing an STM to validate the results
What alternative approach has the ERG suggested?	The ERG did not suggest an alternative approach other than the STM
What is the expected effect on the cost effectiveness estimates?	The expected effect cannot be predicted
What additional evidence or analyses might help to resolve this key issue?	The ERG recognises that it is difficult and intensive to provide results from a model with an alternative structure.
ERG = Evidence Review Group; STM = state transition model	

4.2.3 Population

Consistent with the NICE scope, the population considered in the CS (Table 1 of the CS) was adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC.[1] The anticipated licensed indication of sotorasib is: for the treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated.[23]

The phase 2 trial evidence for sotorasib, i.e. the single-arm CodeBreaK100 study, focused on safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy. The population in the CodeBreaK100 study is defined as: adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per Response Evaluation Criteria In Solid Tumors (RECIST) 1.1 criteria, and had ECOG performance status of 0 or 1.[1]

Subgroup analyses were not included in the cost effectiveness analysis given there were no subgroups observed with substantially different efficacy compared to the whole population. Further, the relatively small population enrolled in CodeBreaK100 (N=126) would limit the sample size available and interpretability of any subgroup analyses.

The key baseline patient characteristics in the economic model are listed in Table 4.5 below and were obtained directly from CodeBreaK100.

Table 4.5: Key baseline patient characteristics of CodeBreaK100 used in the economic model

Patient characteristic	Mean / %	Source
Age at baseline (years)	62.9	CodeBreaK100 CSR, Table 9.2[24]
Gender (female)	50%	CodeBreaK100 CSR, Table 9.2[24]
Weight (kg)	71.1	CodeBreaK100 CSR, Section 9.3[24]
Body Surface Area (BSA, m^2)	1.81	Calculation - Mosteller formula[24]
Based on Table 22 of the CS BSA = body surface area; CS = company submission; CSR = clinical study report; SD = standard deviation		

ERG comment: As already mentioned in Section 2.1.2, the population in CodeBreaK100 and therefore also the population in the economic model, appears to be narrower than that defined in the NICE scope. In addition, the population in the secondary comparison (docetaxel plus nintedanib) may be different from the population in the primary comparison.

The company did not perform a full incremental analysis to compare all three treatment strategies in this appraisal. In their response to the question of the ERG in the clarification phase whether this was because of non-matching populations, the company stated that *"a minority who are eligible for docetaxel will have an add-on nintedanib"* and that there was no easy way to produce a relative treatment effect between sotorasib and nintedanib plus docetaxel.[6]

The ERG considers that the absence of a full incremental analysis for the three treatment options negatively impacts the validity of the comparison and the generalisability of results to UK clinical practice.

4.2.4 Interventions and comparators

The intervention considered in the CS was sotorasib, a $KRAS^{G12C}$ inhibitor. Sotorasib is administered once daily as oral monotherapy, at a dose of 96 0mg (8x 120 mg tablets). The comparators considered were docetaxel monotherapy, or nintedanib for patients with adenocarcinoma. As discussed in Section 2.3, the NICE scope listed the following comparators:

Non-squamous NSCLC:

- pemetrexed with carboplatin
  with or without pemetrexed maintenance
- other platinum doublet chemotherapy
  with or without pemetrexed maintenance
- nintedanib with docetaxel (adenocarcinoma histology)
- docetaxel monotherapy
- atezolizumab
- nivolumab (subject to ongoing CDF review)
- pembrolizumab (PD-L1-expressing tumours)
- best supportive care

Squamous NSCLC:

- gemcitabine with carboplatin or cisplatin
- vinorelbine with cisplatin or carboplatin
- docetaxel monotherapy
- pembrolizumab (PD-L1-expressing tumours)

- atezolizumab
- nivolumab
- best supportive care

People with *KRAS p.G12C* mutation and another driver mutation (including EGFR-TK, ALK or ROS1):

- Established clinical management without sotorasib, including:
    - atezolizumab combination (after EGFR-TK or ALK-targeted therapies)
    - lorlatinib (after ALK-targeted therapies)
    - brigatinib (after ALK-targeted therapies)
    - ceritinib (after ALK-targeted therapies)
    - osimertinib (EGFR T790M mutation-positive after EGFR-TK targeted therapies)
    - pemetrexed with carboplatin
    - platinum doublet chemotherapy
    - with or without pemetrexed maintenance
    - nintedanib with docetaxel (adenocarcinoma histology)
    - nivolumab (subject to ongoing CDF review)

The company justified the limited number of comparators as follows:

- For immunotherapy and combination radiotherapy: re-challenge is not routine clinical practice according to clinical expert opinion obtained from a UK advisory board.
- Co-occurrence of KRAS p.G12C next to another driver mutations, is very rare (<1%).[25]
- Docetaxel monotherapy is considered a key second- and subsequent-line option in NSCLC.[26, 27]
- For adenocarcinoma patients eligible for docetaxel, a combination of nintedanib and docetaxel may be administered in some regions in the UK.

Additionally, the CS states that the use of docetaxel monotherapy as the comparator was agreed upon in scientific advice from NICE and EUnetHTA.[1]

In the company's response to clarification, the company mentioned that the PSWA of chemotherapy regimens from the Flatiron database compared to sotorasib, could be used as a proxy of using platinum doublet chemotherapy as a comparator.[6] These cost effectiveness results were explored in scenario analyses (Section B.3.7.3.1, Table 46 of the CS).[1]

Sotorasib dose reductions are recommended in case of adverse reactions. The first reduction brings the total dosage to 480 mg (four tablets) and the second reduction to 240 mg (two tablets), taken once daily. If patients are unable to tolerate 240 mg daily, treatment should be discontinued. Dose modifications related to adverse events are displayed in the draft SmPC provided by the manufacturer.[23]

ERG comment:

a) The number of comparators included in the cost effectiveness analysis is limited compared to the initial scope set out by NICE. Importantly, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority.[8]
b) The ERG does not consider the suggestion made by the company in their response to clarification that Table 46 in the CS (the analysis using Flatiron data) could be used as a pragmatic reflection of sotorasib versus platinum-based chemotherapy, to be supported by the information presented in the

CS.[1, 6] No conclusions regarding the cost effectiveness of sotorasib versus platinum doublet therapy should be drawn from the analysis presented by the company.

Table 4.6: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line

Report Section	4.2.4
Description of issue and why the ERG has identified it as important	Compared to the final scope for this appraisal, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority
What alternative approach has the ERG suggested?	The ERG has no alternative approach as adding the comparator to the model would require structural and substantial changes which are outside the scope of work for the ERG.
What is the expected effect on the cost effectiveness estimates?	Could potentially have a substantial impact on the cost effectiveness, direction unknown.
What additional evidence or analyses might help to resolve this key issue?	Implementing platinum-based chemotherapy in the model as an additional comparator would help to resolve the issue and reduce uncertainty.
ERG = Evidence Review Group	

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is one week with a 20-year time horizon and a half-cycle correction is applied.

ERG comment: In the CS, the company states a 20-year time horizon was used, at what point <1% of the patients is expected to be alive.[1] This was considered to represent a lifetime time horizon. The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness for sotorasib is the CodeBreaK100 trial (updated from the initial submission to include data up to 15 Match 2021).[28]

Treatment effectiveness of the comparators is derived from the SELECT-1 trial for docetaxel and from the LUME-Lung 1 trial (adenocarcinoma subgroup) for nintedanib plus docetaxel. An additional analysis was provided using real-world data from the Flatiron cohort, in which a basket of standard-of-care chemotherapy was used. As no head-to-head trial was performed comparing sotorasib to its comparators, all analyses are indirect analyses, the methods of which are described in Section 3.4.

4.2.6.1 Sotorasib versus docetaxel

For the base-case estimation of the OS for sotorasib versus docetaxel, an HR of [redacted] was derived from the MAIC indirect analysis and an HR of [redacted] for PFS. Several parametric distributions were fit to the data from the CodeBreaK100 trial and the adjusted data from the SELECT-1 trial. In the CS a joint-fit restricted lognormal model fit best to the OS data, considering the Akaike information criterion (AIC) and Bayesian information criterion (BIC) and sensitivity analyses with the other models

were also provided. For the PFS, a restricted joint fit using a generalised gamma model fit best to the data when the AIC was considered and a lognormal fit best to the data when the BIC was considered.

4.2.6.2 Sotorasib versus nintedanib plus docetaxel

For the nintedanib plus docetaxel comparator, no patient-level data were available, instead, pseudo-patient level data was generated from the published results of the LUME-Lung 1 trial, which compared nintedanib plus docetaxel to placebo plus docetaxel. No MAIC was performed, as the patient population in the LUME-Lung 1 trial was deemed to differ too much from the CodeBreaK100 trial population. Nintedanib was modelled by applying time-dependent HRs to the data for docetaxel patients from SELECT-1. For the OS comparison between docetaxel plus nintedanib versus docetaxel plus placebo, piecewise HRs were used: for 0-6 months [redacted] for 6- to 26 months [redacted], and for 26 months andover [redacted] Piecewise HRs were also considered for the PFS: for 0-2 months [redacted], for 2-6 months [redacted], and for 6 months and over [redacted].

4.2.6.3 Flatiron real-world data

An alternative analysis was provided in the CS using the Flatiron real-world dataset. In this analysis, the sotorasib data from the CodeBreaK100 trial was compared to a basket of standard-of-care chemotherapy: 21 out of 206 patients in this dataset were on docetaxel monotherapy. Also 85 out of 206 participants had a KRAS p.G12C mutation. Using a propensity score analysis described inSection 3.4, the HR for OS was estimated at [redacted] for the KRAS p.G12C mutant subgroup and the HR for the PFS was estimated at [redacted]. For both the OS and PFS a restricted joint fit lognormal model provided the best fit considering the AIC and BIC.

4.2.6.4 Waning of treatment effect

In the base-case of the CS, sotorasib was extrapolated for the full time horizon of the analysis.[1] In the original CS, a scenario analysis was provided to limit the treatment effect of sotorasib to 5 years and in the company's response to the ERG clarification questions, seven additional scenario analyses were provided.[1, 6] Two methods were used to incorporate treatment effect waning (TEW): gradual TEW and immediate TEW.

In the gradual TEW, the sotorasib effects gradually decrease for 5 years starting in year 2, 3, 4 or 5; for the immediate TEW, the sotorasib HRs were immediately set to 1 (meaning no benefit compared to the comparator) from year 2, 3, 4 and 5. In response to the request for clarification, the company noted that *"TEW is a very blunt tool and in an ideal world its use should be limited to cases where there is no (or very little) available external data to compare or adjust long term extrapolations with"* and provides several reasons why TEW should not be applied in this case.[6]

4.2.6.5 Treatment duration

TTD for sotorasib was estimated by applying an HR of [redacted] to the PFS curve. A sensitivity analysis was included where treatment discontinuation was modelled using separate parametric models, which is an approach in line with the methods used to model OS and PFS for sotorasib and docetaxel. According to the CS, for docetaxel no robust data were available, and TTD was assumed to be equal toPFS. For nintedanib plus docetaxel, treatment duration was also set to be equal to PFS, which is a conservative estimate according to the CS, as in a previous NICE submission (TA347) the PFS rate washigher than the discontinuation rate.[9]

4.2.6.6 ERG comment

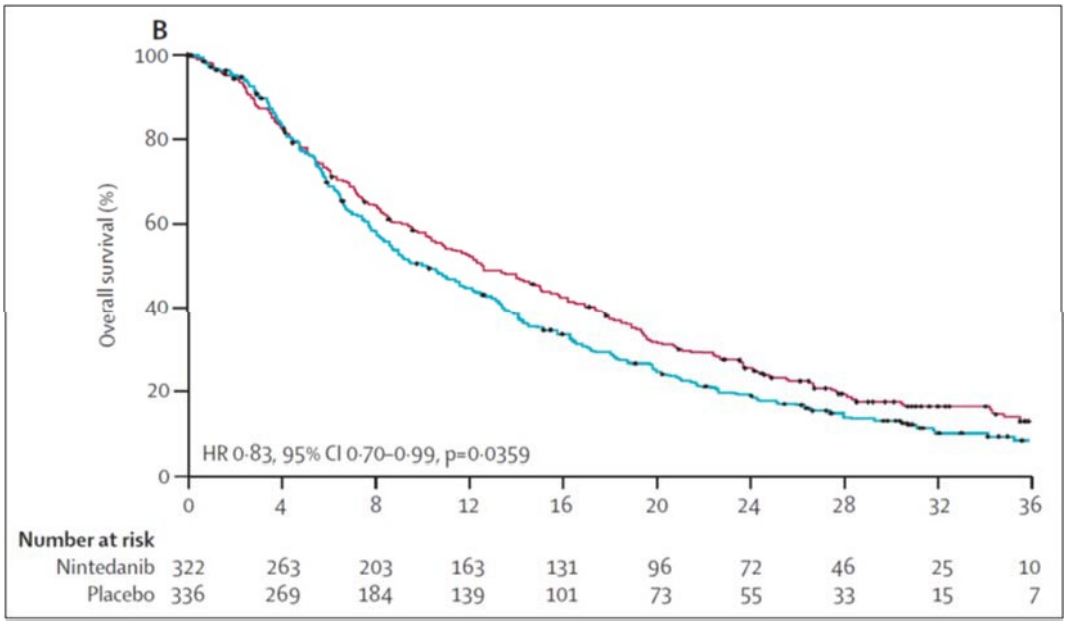
a) The methods to extrapolate the treatment effect of sotorasib versus docetaxel using the CodeBreaK100 and SELECT-1 trial data are well explained in the CS and the decisions made are clear. It should be noted however, that the decision for a specific parametric model remains somewhat arbitrary and can have a major influence on the model outcomes. As provided in the company's response to clarification questions, the deterministic incremental cost effectiveness ratio (ICER) ranges from £30,112 to £62,123 per QALY depending on the chosen PFS and OS functions. As the presented OS and PFS curves are the results of an indirect analysis, the ERG expects additional uncertainty regarding the chosen model, mainly for the comparator. Since the generalised gamma distribution provided the best fit for PFS considering the AIC, the ERG considers the use of this distribution to be an important scenario to include; next to the base-case in which the lognormal distribution was used which provided the best fit considering the BIC.

b) The modelling of nintedanib plus docetaxel is subject to considerable uncertainty: first, the LUME-Lung 1 trial data is used to compare docetaxel plus nintedanib to docetaxel plus placebo, then the resulting HRs are applied to the SELECT-1 data, which are then used for the indirect analysis using the same methods as for the sotorasib versus docetaxel comparison.
In addition to the uncertainty introduced by this method, the ERG has major concerns regarding the clinical plausibility of the resulting OS curve. First of all, the patient populations of the SELECT-1 trial and LUME-Lung-1 trial differ mainly in terms of smoking and performance status (Table 10 of CS) and the CS does not report any adjustments for these differences.[1] Additionally, the resulting HR of [redacted] for the first 6 months results in a major rise in mortality (see Figure 4.2, copied from the economic model provided by the company). The ERG finds it implausible that adding nintedanib to docetaxel treatment would result in a major rise in mortality and does not consider the resulting OS curve to be in line with the Kaplan-Meier-curve reported in the LUME-lung-1 trial (see Figure 4.3). There was no expert opinion provided in the CS to support Figure 4.2. Note that the titles of Tables 30 and 31 in the CS contain an error, as the HRs provided are for docetaxelplus nintedanib versus docetaxel plus placebo; not for nintedanib plus docetaxel versus sotorasib.[1]

Figure 4.2: Modelled OS curves taken from the economic model

[redacted]

Figure 4.3: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line)



On visual inspection of the OS and PFS survival curves provided in the company's response to clarification questions, none of the fitted curves have a particularly good fit. A piecewise analysis was used, with two cut-off points, for the OS at 6 months and at 26 months. The ERG does not agree with the cut-off point at 26 months and the company failed to justify this approach both in the initial CS and in the company's response to clarification questions. The ERG suggests reducing the number of cut-off points to one at month 6.

c) The company did not consider any waning of the treatment effect, and in their response to the clarification questions, the company noted that *"TEW is a very blunt tool and in an ideal world its*

use should be limited to cases where there is no (or very little) available external data to compare or adjust long term extrapolations with".[6] According to the ERG, there is no external data available in this case, as the only data regarding the treatment effects of sotorasib come from the CodeBreaK100 trial, with a limited follow-up time and no comparators.

The ERG does agree with some of the points made in the company's response, e.g. that the impact of discontinuation is already somewhat "baked" into the model. On the other hand, it may not be reasonable to expect that patients continue to benefit from the treatment indefinitely, even after they have stopped treatment. Considering that only 18 months of CodeBreaK100 trial data have been collected and there is no additional information available for the sotorasib treatment effects beyond this, the ERG thinks it is a feasible approach to introduce a gradual TWE after 24 months, for which a waning period of 5 years can be used; the period is suggested in the company's response to clarification questions.

d) The TTD was modelled by applying an HR to PFS from CodeBreaK100. The company explored an alternative approach in a sensitivity analysis where the weights generated from the MAIC analysis were applied to the CodeBreaK100 discontinuation data and parametric models were fitted to extrapolate the treatment duration. However, the company considered this approach to be more complex and ultimately dependent on the variable selection in the MAIC analysis. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, i.e. based on MAIC. Also, by basing the TTD on the PFS, TTD would still be, via PFS, ultimately dependent on the variable selection in the MAIC. Moreover, in Figure 36 of the CS the company presents Kaplan-Meier data for TTD alongside the modelled curve and the ERG believes this to be a poor fit, potentially underestimating true TTD in the long run.

Table 4.7: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The indirect way of estimating OS and PFS for the secondary comparator docetaxel plus nintedanib leads to worse survival for docetaxel plus nintedanib compared to docetaxel plus placebo in the first 6 months of the OS curve.
What alternative approach has the ERG suggested?	The ERG prefers to assume that the HR for docetaxel plus nintedanib versus docetaxel plus placebo cannot go above 1
What is the expected effect on the cost effectiveness estimates?	Lowering the HR for docetaxel plus nintedanib versus docetaxel plus placebo will increase the ICER for sotorasib versus docetaxel plus nintedanib
What additional evidence or analyses might help to resolve this key issue?	Direct evidence for this comparison
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression-free survival	

Table 4.8: Key issue 9. No waning of treatment effect

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The company's assumption of continued effect of sotorasib does not seem justified and is difficult to maintain given immature evidence.

What alternative approach has the ERG suggested?	The ERG suggested to start waning of the treatment effect at the 2-year timepoint and have it gradually decreased to an HR of 1 over a period of 5 years (with exploratory scenario analyses for 3 and 7 years).
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on lasting treatment effect.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio	

Table 4.9: Key issue 10. TTD modelling approach

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The TTD was modelled by applying a hazard ratio to PFS from CodeBreaK100. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, fitting a parametric curve on TTD data using weights based on the MAIC.
What alternative approach has the ERG suggested?	The ERG suggested to use the company's alternative approach, based on the MAIC, in the base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on observed treatment duration in sotorasib and comparator arms
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation	

4.2.7 Adverse events

The company included grade 3+ adverse events with an incidence of ≥5% in any of the comparator arms in the analysis, considering data from the CodeBreaK100, SELECT-1 and LIME-Lung 1 trials.[1] Only TRAEs were included in the analyses, as only these were available from the LUME-Lung 1 trial. Treatment-emergent adverse events were included in a scenario analysis for the comparison of sotorasib to docetaxel.

Disutilities related to adverse events were included in the analysis, the values of which are provided in Table 36 of the CS.[1] If no disutility value could be identified, this was assumed to be 0. This is the case for: decreased neutrophils, increased AST, and pleural effusion.

ERG comment:

a) The inclusion of only TRAEs could negatively impact the validity of the assessment, as the quality of life of patients may not be captured well if TEAEs are excluded. However, the company provided

a scenario analysis including TEAEs for the comparison of sotorasib to docetaxel which increased the ICER from £43,660 to £44,116 per QALY, which the ERG considers a minor impact.

b) Disutilities were assumed to be 0 if no disutility value could be identified. The CS states that: *"This assumption could potentially be conservative given the generally increased frequency of these AEs in the comparator arms versus sotorasib"*.[1]
The ERG does not agree with this statement, although it seems to be reasonable for the nintedanib plus docetaxel comparison, it is not reflected by the data for docetaxel monotherapy. Within the sotorasib arm, the incidence of decreased neutrophils was 0.8% and the increased AST incidence was 5.6%, while this was 0.0% and 0.0% respectively for the docetaxel arm. As no disutility was applied to these adverse events, this is expected to favour the cost effectiveness of sotorasib compared to docetaxel. In contrast, as decreased neutrophils are more prevalent in the nintedanib plus docetaxel comparison, it may negatively impact the cost effectiveness of this comparator.

4.2.8 Health-related quality of life

The utility values were estimated for the following health states: progression-free, and post-progression, via a disutility subtracted from the progression-free utility. Notably, these health state utilities were only used in a sensitivity analysis as the approach taken in the base-case was to use time to death utilities.

4.2.8.1 Utility values

In the absence of studies from the SLR (see Section 4.1.3), the primary source of HRQoL values in the model was CodeBreaK100.[1] HRQoL was collected in CodeBreak100 using the EuroQoL-5D-5L instrument.[29] This instrument was completed on the first day of cycle 1, on every first day of subsequent cycles until cycle 7, and then on the first day of every second cycle until end of treatment. The company defined various datasets, see Table 4.10 for details. Using mixed models with repeated measures (MMRM), utilities were estimated using two approaches: time to death and health states. Theanalysis included several combinations of datasets and covariates.

Table 4.10: Datasets used for HRQoL analysis

	Original	AN01	AN02
Safety analysis set	N=126	N=122[#]	N=86
Full analysis set	N=123	N=119[*]	N=84

Based on pages 111 and 112 of the CS[1]
AN01 = patients who completed at least one EQ-5D-5L questionnaire in line with study protocol with all fields of the questionnaire completed; AN02 = patients who completed EQ-5D-5L at baseline visit per protocol and at least one other completed EQ-5D visit
[#] used for time to death utilities analysis in the model; [*] used for health state utilities analysis in the model
CS = company submission; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Levels; HRQoL = health-related quality of life

Health-state utility values

For the health-state utilities, the CS presented results for both the AN01 and the AN02 full dataset.[1] Although the company did find that a model including both progression status and baseline utility score as covariates fitted best, this would require the use of the AN02 dataset since in the AN01 dataset not all subjects had completed the baseline questionnaire. And so, *"to account for all information available"*, as the company stated, the MMRM with only progression status based on AN01 was used to inform the model.[1]

Time-to-death utility values

For the time-to-death utility analysis, the AN01 dataset was used as well, but based on the safety analysis set instead of full analysis set, which implied a few additional subjects were included compared to the health state utility values, see Table 4.11.

The company provided mean utility scores visually in Figure 40 of the CS and the final time-to-death utility scores used to inform the model in Table 35 of the CS, which were updated in the addendum accompanying the response to clarification to reflect the latest data cut-off (15 March 2021).[1, 30] These updated time-to-death utilities were used in the company's base-case, and preferred over the health-state utility scores. The company stated this to be, amongst other reasons, because the time-to-death approach reflects the findings of studies which have shown NSCLC patients to have markedly decreased utilities towards the end of life.[1, 6]

A summary of all utility values used in the cost effectiveness analysis is provided in Tables 4.11 and 4.12.

Table 4.11: Health-state utility values - used in sensitivity analysis

Health state	Utility value (mean and 95% CI)	Reference
Progression-free	0.734 (0.700 to 0.769)	CodeBreaK100[30, 31a] and UK crosswalk tariffs[32]
Disutility in progressed disease	0.064 (0.097 to 0.031)	
Post-progression	0.670	Calculation

Based on addendum to clarification response, Table 7[30]
[a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-47.702 and subsequent analyses
CI = confidence interval; UK = United Kingdom

Table 4.12: Time-to-death utilities - used in CS base case

Health state	Utility value (mean and 95% CI)	Reference
Utility more than 6 months to death	0.762 (0.698, 0.767)	CodeBreaK100[30, 31a] and UK crosswalk tariffs[32]
Disutility between 3 and 6 months to death (versus more than 6 months)	0.047 (0.090, 0.004)	
Disutility between 1 and 3 months to death (versus more than 6 months)	0.125 (0.176, 0.074)	
Disutility less than 1 month to death (versus more than 6 months)	0.233 (0.312, 0.153)	
Utility between 3 and 6 months to death	0.715	Calculated
Utility between 1 and 3 months to death	0.637	Calculated
Utility in last month of life	0.529	Calculated

Based on addendum to clarification response, Table 8[30]
[a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-47.702 and subsequent analyses
CI = confidence interval; CS = company submission; UK = United Kingdom

4.2.8.2 Disutility values

In the absence of reported utility data for the comparators, the company included a disutility to express the implications of a hospital-based intravenous (IV) administration and increased cytotoxicity of docetaxel and nintedanib plus docetaxel.[1] The utility decrement was set at 0.025 (per cycle on treatment), based on a previous study in advanced NSCLC and disutility associated with IV administration.[33] This previous study (published in 2010) was on the cost effectiveness of erlotinib versus docetaxel and reported utilities of 0.451 and 0.426 for oral therapy and IV therapy respectively, in the progression free health state. These utilities were determined by having 154 members of the general population from four UK sites filling out a visual analogue scale (VAS).[33]

Disutilities of adverse events were discussed in Section 4.2.7.

The utilities in the economic model were not adjusted for age and sex. In response for a scenario analysis including age related decrements, the company added a scenario applying an adjustment to utilities based on the sex-matched general population utilities, to ensure that the estimated patient utilities never exceed that of the general population.[6] The company did however not apply an age-related decrement in this scenario since the TTD utility values were considered to already account for aging.

ERG comment: The main concerns of the ERG relate to a) the choice of TTD utilities over health state utilities as the TTD utilities seem less well informed; b) the treatment disutility applied for the comparators; and c) the absence of an age-related decrement:

a) Although using a time to death approach to utility scores may be justified in this population, the ERG has concerns about the data underlying the estimates used in the model. Firstly, by relying on the AN01 dataset, all patients that at least filled out one EQ-5D questionnaire were included in the analysis. Given that there were at maximum 14 timepoints available at which patients could have completed a questionnaire (see Figure 38 of the CS), one questionnaire seems the bare minimum and this raises questions about representativity of the sample.[1]
   In the AN02 dataset, patients had to have completed at least 2 questionnaires, one of which at the baseline visit. Using the AN02 dataset may have been more valid and stable, but the company discarded the AN02 *"to account for all information available"* even though the mixed model including baseline utility score as a covariate had a better fit than the model with progression status alone.
   Then, for the time to death analysis, the AN01 dataset was again preferred over the AN02 dataset, seemingly because the company wanted to align with the health state utility analysis (but nevertheless did decide to use the safety analysis dataset here instead of the full analysis dataset). Although for the health state utility approach some results of the AN02 dataset were presented, for the TTD approach no information on AN02 analyses were provided. In addition, the TTD utility scores presented in Table 35 of the CS and the final TTD utilities used in the model (see Table 4.12) above do not seem to match very well with the visual representation of mean utilities shown in Figure 40 of the CS.[1]
   Also apparent from Figure 40 of the CS is that numbers of distinct patient underlying the scores were quite small, i.e.., 86, 30, 31 and 12 for the more than 6 months, 3 months to 6 months, 1 month to 3 months, and less than 1 month to death categories.
   Altogether, the ERG considers the TTD utilities not reliable and therefore prefers the utilities by health state approach.
b) The disutility applied for IV administration of docetaxel. In the clarification phase, the ERG asked whether the disutility applied was appropriate for use in this case, given that the value for this disutility was derived by using a VAS instrument in a general population (so not EQ-5D, and

therefore not officially utilities) and the values obtained for the progression free health states in this study were vastly lower than observed in CodeBreaK100 (i.e. 0.426 and 0.451 compared to 0.74, respectively).[34] The company, in their response, made a case for treatment-specific utilities when comparing targeted therapy vs. chemotherapy.[6]

The ERG agrees that the use of treatment-specific utilities may be justified but considers the source used for disutility in the company base-case to be a questionable one. Apart from this, the company did not provide a response to the questions of the ERG how one day of IV infusion per 3 weeks would compare, in terms of quality of life, to taking eight tablets every day, as is the case for sotorasib treatment and for which no disutility was applied.. The company provided two alternative scenarios, which both effectively increase the disutility compared to the company base-case.[1, 6]

c) The ERG considers the fact that utilities in the model were not adjusted for age to be a potential source of bias. Although the company did provide a scenario where utilities in the model could not exceed the sex-adjusted utilities in the general population, the utilities in the model could then in theory still exceed the age-adjusted utilities in the general population, even though TTD utilities would decrease over time. The ERG would have liked to see a scenario as requested, including age-related utility decrements, to estimate the impact of such a scenario.

Table 4.13: Key issue 11. Time-to-death utilities do not seem well-informed

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	The time to death utilities which the company used in the base-case did not seem well-informed. The data underlying the estimates were sparse, and increasingly so for the closer to death states.
What alternative approach has the ERG suggested?	The ERG suggested to use utilities based on disease progression as base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Fully specified models using also AN02 dataset should be provided to see which approach is most appropriate. But given that even AN02 probably has many missing data this may still not be ideal.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 4.14: Key issue 12. Disutility for IV administration not well justified

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	A disutility for IV administration of docetaxel is applied without sufficient justification for the size of the disutility or the exclusion of the potential disutility for taking eight tablets of sotorasib daily.
What alternative approach has the ERG suggested?	The ERG suggested to exclude the IV disutility in the base-case
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses	Comparative evidence on (observed) health state utilities in sotorasib and comparator arms could resolve this

Report Section	4.2.8
might help to resolve this key issue?	
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; IV = intravenous	

4.2.9 Resources and costs

The cost categories included in the model were treatment acquisition costs, medical costs (treatment administration, monitoring and disease management, subsequent treatments, and terminal care), and costs of managing adverse events.

Unit prices were based on the NHS reference prices, British National Formulary (BNF), Personal Social Services Research Unit (PSSRU) and Electronic Market Information Tool (eMIT).[35-38] All costs, where necessary were inflated to the 2018/2019 cost year to remain consistent with the latest available NHS Reference Costs using the PSSRU Hospital and Community Health Services (HCHS) and the NHS Cost Inflation Index inflation indices (NHSCII).[37]

4.2.9.1 Treatment costs (with patient access scheme (PAS))

Drug acquisition costs for the intervention and comparators are presented in Table 4.15. The sotorasib dose of 960 mg per day is consistent with the anticipated license and the dosing regimen in CodeBreaK100.[5, 39, 40] Dosage for docetaxel and docetaxel plus nintedanib is aligned with UK clinical practice and informed by NHS treatment protocols.[41]

Estimates of relative dose intensity (RDI) as observed in respective clinical trial programmes were applied to calculate total monthly costs.[9, 15, 39, 41, 42] RDI for sotorasib was slightly lower at 89% compared to docetaxel and nintedanib (90.3% and 921.1% respectively). In response to the request for clarification the company stated that there would be no reason to assume that RDI is truly lower for sotorasib, and that the differences in these observations may reflect random sampling error.[6]

Drug wastage was not discussed as such in the CS but from Table 48 in the CS it is apparent that the base-case assumption was zero wastage and that a scenario was run to test the impact of potential drug wastage in clinical practice by estimating drug acquisition costs based on total packs as opposed to treatments received.[1] In their response to the request for clarification, the company stated to maintain their base-case assumption of zero wastage.[6] This was justified with arguments on the ability to implement dose reductions and the single strength formulation of sotorasib, which would allow the pharmacist to optimise the dose without wastage and provide the appropriate supply of drugs to patients until disease progression is recorded. The company stated that they believe the scenario analysis including wastage would significantly overestimate the true drug utilisation.[6]

Table 4.15: Unit drug costs

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)
Sotorasib	240x 120 mg tablets	[redacted]	[redacted]	960 mg per day	89%	[redacted]
Docetaxel	160 mg per vial	17.95	eMIT[38]	75 mg/m^2 on day of treatment	90.3%	19.93

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)
Nintedanib	120x 100 mg tablets	2,151.10	BNF[36]	400 mg per day (21-day cycle)[a]	92.1%	1,926.28

Based on Table 38 of the CS[1]
[a] Nintedanib administered on days when docetaxel is not taken, i.e. 20 days per 21 day cycle
BNF British National Formulary; CS = company submission; eMIT = electronic Market Information Tool, RDI = relative dose intensity

Treatment administration costs were assumed to be zero for sotorasib and nintedanib as these are both taken orally. For docetaxel, administration costs were based on NHS reference costs for the administration of simple parenteral chemotherapy.[35] See also Table 38 in the CS.[1]

4.2.9.2 Health state and event costs

Costs of monitoring and disease management were largely based on assumptions used and accepted in previous NICE STAs, in particular NICE TA347 on nintedanib plus docetaxel.[9] Apart from a per-cycle cost per health state, a one-off cost was applied at treatment initiation and at progression. A one-off cost was also applied at the moment of dying to reflect the cost of terminal care, based on the values used in the NICE multiple technology appraisal (MTA) for erlotinib and gefitinib, see Table 4.16 for an overview.[43]

Table 4.16: Disease management and terminal care costs

		Source
Health state	**Cost per cycle (£)**	
Progression-free	77.04	NHS reference costs 2018/2019[35]; PSSRU[37]; aligned with NICE TA347[9] and TA428[44]
Post-progression	39.98	
Event	**Cost (£)**	
At treatment initiation	834.25	NHS reference costs 2018/2019[35]; PSSRU[37]; aligned with NICE TA347[9] and TA428[44]
At progression	116.53	
Terminal care	3,759.73	Appendix L of CS[13]

Based on Tables 39 and 42 of the CS[1]
CS = company submission; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSSRU = Personal Social Services Research Unit; TA = Technology Appraisal

4.2.9.3 Adverse event costs

The unit costs related to the management of adverse events were mainly derived from a previous NICE MTA for erlotinib and gefitinib.[43]

4.2.9.4 Subsequent treatment costs

The costs of subsequent treatment were included in the economic model as a one-off cost at disease progression, see Table 4.17. The distribution of subsequent treatments for docetaxel and docetaxel plus nintedanib was informed by previous STAs as was treatment duration. The distribution of subsequent treatments for patients who progress on sotorasib was informed by UK clinical experts. In response to the request for clarification, the company also provided data on observed treatment mix in 44 patients

receiving subsequent treatments in CodeBreaK100, which revealed that patients would often receive more than one subsequent treatment, i.e. the 44 patients in the sample altogether receive[redacted] subsequent treatments, see Table 4.18.[6]

Table 4.17: Subsequent treatment costs

Subsequent treatment / Original treatment	BSC	Platinum-based	Docetaxel	Source
Sotorasib (%)	50%	10%	40%	Assumption based on clinical expert feedback
Docetaxel (%)	70%	30%	0%	NICE TA 347 – assumption[9]
Nintedanib + docetaxel (%)	70%	30%	0%	NICE TA 347 – assumption[9]
Treatment duration (weeks)	14	14	14	NICE TA347, TA428[9, 44]
Cost of subsequent treatment (£)	0	2,835	1,219	Calculation – appendix L[13]

Based on Table 41 of the CS[1]
BSC = best supportive care; CS = company submission; NICE = National Institute for Health and Care Excellence; TA = Technology Appraisal

Table 4.18: Treatment mix for 44 patients receiving subsequent treatments

Treatment	N	Proportion of 77 treatments (%)	Proportion of 44 patients (%)
Pemetrexed or docetaxel	[redacted]	[redacted]	[redacted]
Platinum based chemotherapy	[redacted]	[redacted]	[redacted]
Others* or non-interventional therapy	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Based on page 67 of the response to the request for clarification[6]
* Other includes novel treatments assessed in clinical trial settings and other treatments not relevant UK clinical practice or unknown
UK = United Kingdom

ERG comment: The main concerns of the ERG relate to:

a) The RDI of sotorasib being lower than for the comparators while the company have stated in their response to clarification that there is no reason to assume this. The ERG believes it would be reasonable to set the RDI for sotorasib, docetaxel, and docetaxel plus nintedanib at 90.5% which is the average of the observed RDIs for the three interventions considered.
b) The assumption of zero wastage, which the ERG considers to be overly optimistic. Although an oral drug at a fixed dose will be associated with less wastage than IV treatment which is dosed based on BSA, the ERG does not believe it to be likely that packs of sotorasib, once delivered to the patient and opened, will be returned and later administered to other patients. Hence, some wastage will always occur, no matter how precise and short-term the dosing.
c) Subsequent treatment costs for sotorasib are likely underestimated by assuming that patients would receive only one subsequent treatment while data from CodeBreaK100 suggests otherwise. The ERG considers the percentage of actual patients receiving docetaxel and platinum-based

chemotherapies is more relevant here than the mix between therapies. Notably, the percentage of patients receiving platinum-based chemotherapies may be underestimated in the model.

Table 4.19: Key issue 13. Relative dose intensity and wastage assumption not justified

Report Section	**4.2.9**
Description of issue and why the ERG has identified it as important	In their base-case, the company assumed a lower RDI for sotorasib than for comparators, which was not justified. The company also assumed zero wastage for sotorasib, which the ERG also considered not justified.
What alternative approach has the ERG suggested?	The ERG proposed to take the average RDI as base-case, and to include wastage based on opened packs.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	For the wastage, the company would have to make a convincing case that opened packs, when not used, would be returned for usage by other patients, i.e. a specific program would have to be in place.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; RDI = relative dose intensity	

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base-case cost effectiveness results from the updated model indicated that sotorasib is both more effective and more costly than docetaxel, which resulted in an ICER of £43,660 per QALY gained (Table 5.1).[1] When comparing sotorasib to the secondary comparator, docetaxel plus nintedanib,the deterministic ICER was £33,628 per QALY gained (with additional costs of £[redacted], incremental QALYs [redacted] and life years gained (LYG) [redacted]).

It should be noted that in the original CS, the ICER for sotorasib vs. docetaxel was £47,176 per QALY gained.[1] After updating the model with the new data cut-off point of 15 March 2021 for CodeBreaK100 (the original CS was submitted with data cut-off of 01 December 2020), the ICER decreased by 7.5%.[15] The increase in OS of sotorasib compared to docetaxel was the main driver for the lowered ICER compared to the original submission.

Table 5.1: Deterministic base-case results: sotorasib vs. docetaxel

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]				
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	43,660
Based on updated company model ICER = incremental cost effectiveness ratio; LYG = life years gain; QALY = quality-adjusted life years							

Overall, the technology is modelled to affect QALYs by:

- Increasing survival, which accrues in PF [redacted] vs [redacted] months) as well as in PP [redacted] vs [redacted] months).
- Increased QoL because of the longer survival, and because of treatment-related disutility for docetaxel.

Overall, the technology is modelled to affect costs by:

- The higher cost of sotorasib compared to docetaxel [redacted] vs. £17.95).
- Early treatment discontinuation for sotorasib compared to docetaxel.

ERG comment: To test the effect of extreme values on the model, the weight was set to zero and there were small changes on ICER. The reason was that the treatment of sotorasib was not dependant on weight. In the updated model after clarification, weight was removed.

5.2 Company's sensitivity analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSAs), deterministic sensitivity analyses (DSAs) as well as scenario analyses.[1] The PSA included probabilistic parameters that were used to estimate QALYs and costs. The PSA was run for 1,000 iterations. The 2.5th and 97.5th percentiles for the probabilistic incremental costs and QALYs were £[redacted] to £[redacted] and [redacted] to [redacted] respectively. The PSA shows consistency with the deterministic results with an ICER of £43,183 per QALY gained. The probability of sotorasib being cost effective against docetaxel is [redacted] (Figure 5.1).

The PSA for the secondary comparator (docetaxel plus nintedanib) is more favourable towards sotorasib with a probability of being cost effective of [redacted] The 2.5th and 97.5th percentiles for the probabilistic incremental costs and QALYs were £[redacted] to £[redacted] and [redacted] to [redacted] respectively. The probabilistic ICER of £33,368 per QALY gained is consistent with the deterministic results.

Based on the DSA of sotorasib versus docetaxel, the parameters that have the greatest effect on the ICER are the following:

- the HR applied to PFS to model sotorasib treatment duration (TTD)
- the time to death utility for >6 months prior to death
- disease management costs per week in the progression free health state

Based on the DSA of sotorasib versus docetaxel plus nintedanib, the parameters that have the greatest effect on the ICER are the following:

- HR of OS in the third period (from week 113 until week 261)
- HR of OS in the first period (up to week 26)
- HR of OS in the second period (from week 26 until week 113)

Consistently, modelling assumptions that relate to these parameters likely have the greatest effect on the ICER. This is illustrated by the following CS scenarios that have a substantial impact on the ICER (Table 48 of the CS):[1]

- Limiting treatment effects to 5 years (ICER: £46,684 per QALY gained)
- Applying health state utilities by progression status (ICER: £47,208 per QALY gained)
- Including drug wastage (ICER: £46,387 per QALY gained)
- Excluding RDI (ICER: £48,944 per QALY gained)
- MAIC-adjusted TTD curve from CodeBreaK100 (ICER: £44,496 per QALY gained)
- Generalised gamma distribution selected to estimate long-term PFS (ICER: £45,123 per QALY gained)
- Joint (unrestricted) lognormal distribution selected to estimate long-term PFS (ICER: £47,917 per QALY gained)

ERG comment:

a) Patient characteristics (age, sex, BSA) should not be included in PSA.
b) A scenario assuming TTD for sotorasib was equal to PFS (like for the comparators) was not included.

Figure 5.1: The cost effectiveness acceptability curve for sotorasib versus docetaxel

QALY = quality-adjusted life year; WTP = willingness to pay

5.3 Model validation and face validity check

Some aspects of validation were discussed by the company in the validation Section of theCS (Section B.3.9).[1] The clinical plausibility of the parametric models used was evaluated by comparing modelled median PFS and OS to the reported medians in the MAIC adjusted CodeBreaK100trial and the docetaxel arm of the SELECT-1 study (CS Section B.3.9.1).[1] Also, the predicted OS landmark results at the 1 year, 5 year and 10 year points for the various parametric models were evaluated based on clinical expert opinion. The base-case jointly fitted (restricted) log-normal distribution was considered to be clinically valid for the population under consideration.

The real-world Flatiron Health database was used to test the robustness of the results generated by the MAIC. Using this data, the ICER of the base-case scenario would be £38,279 per QALY gained which is 12.3% less than the ICER of the base-case using CodeBreaK100 data.[1] The main difference was caused by the longer OS and PFS when using Flatiron instead of CodeBreaK100 (see Table 5.2). The company considered these results to be consistent with the conclusion of the MAIC analysis and underlining the robustness of the analyses presented.[1]

Lastly, quality control of the economic model was performed by systematic examination of calculations, extreme value analysis and tracing of calculations. The company used a verification checklist to guide this, but details of this checklist were not made available.

Table 5.2: Disaggregated results by using Flatiron and CodeBreaK100

	Flatiron		CodeBreaK100		Difference between Flatiron and CodeBrak100	
	Sotorasib	**Docetaxel**	**Sotorasib**	**Docetaxel**	**Sotorasib**	**Docetaxel**
PFS, mean (months)	[redacted]					
OS, mean (months)	[redacted]					
LYG in PFS	[redacted]					
LYG in OS	[redacted]					
QALYs	[redacted]					
Costs (£)	[redacted]					
ICER(£/QALY)	38,279		43,660		-5,381	

Based on updated company model
ICER = incremental cost effectiveness ratio; LYG = life year gained; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year

ERG comment: The ERG considers the validation as described by the company to be minimal. As discussed in Section 4.2.2, the company did not provide a scenario with a state transition model (STM) and so validating the model structure in this way was not possible. The Flatiron analysis shows some rather distinct changes (in PFS, OS) for mainly the docetaxel arm compared to the CodeBreaK100 analysis.

In the absence of suitable clarification for these differences, the ERG does not agree with the company that the results from the Flatiron analysis underline the robustness of the analyses presented. However, if there is a lack of correspondence between the results based on the MAIC (using SELECT-1) and the PSWA (using Flatiron), this might be because the latter provides estimates that are less biased, although, as discussed in Section 3.4.2, this is very uncertain.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020:[45]

- 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 confidence intervals, small sample sizes, or immaturity of data
- Bias & indirectness, e.g. there is a mismatch between the decision problem 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 cost effectiveness, whether it is reflected in the ERG base-case 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 ERG report, the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous Sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016):[46]

- Fixing errors (FE; correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV; correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ; amending the model where the ERG considers that reasonable alternative assumptions are preferred)

6.1.1 Explanation of ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as a starting point) are listed below. Section 6.2 shows the impact of each adjustment plus the combined effect of all abovementioned adjustments simultaneously, in the deterministic, probabilistic and scenarios analyses.

6.1.1.1 Fixing errors

No errors were found in the CS model.

6.1.1.2 Fixing violations

1. Patient characteristics included in the PSA (Section 5.2).
   The ERG corrected this.

6.1.1.3 Matters of judgement

2. Key issue 10 (Section 4.2.6)
   TTD modelling approach for sotorasib: the ERG used the approach based on the MAIC fitting parametric models instead of HR applied to PFS.

3. Key issue 11 (Section 4.2.8)
   Method for health state utilities; the ERG used utilities based on disease progression instead of time to death.
4. Key issue 13 (Section 4.2.9)
   Relative dose intensity (RDI); the ERG assumed these to be equal (at average) for all interventions.
5. Key issue 13 (Section 4.2.9)
   Method to calculate treatment costs: the ERG preferred to calculate treatment costs on a per-opened-pack basis.
6. Distribution of subsequent treatments (Section 4.2.9)
   The ERG changed the distribution of subsequent treatments based on total patients receiving these.
7. Key issue 12 (Section 4.2.8)
   The ERG excluded the utility decrement for IV infusion.
8. Key issue 9 (Section 4.2.6)
   The ERG implemented a limit to the treatment effect at the 2 year timepoint, with a subsequent gradual waning of the effect over 5 years.
9. Key issue 8 (Section 4.2.6)
   For the secondary comparison (docetaxel plus nintedanib), the ERG assumed OS for docetaxel plus nintedanib could not be worse than OS for docetaxel plus placebo (i.e. where HR exceeded 1 it would be set equal to 1).

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base-case. The main concern for the ERG was extrapolating the effectiveness of sotorasib and the comparators.

6.1.2.1 Exploratory scenario analyses

1. AE disutilities (Section 4.2.7)
   For AEs where disutility was zero, a disutility of 0.05 was assumed.
2. Treatment-emergent vs. treatment-related AEs (Section 4.2.7)
3. PFS distribution (Section 4.2.6)
   Assuming a generalised gamma distribution instead of lognormal distribution for PFS.
4. Gradual treatment waning (Section 4.2.6)
   Assuming gradual waning of treatment effect over 3 years (instead of 5 years).
5. Gradual treatment waning (Section 4.2.6)
   Assuming gradual waning of treatment effect over 7 years (instead of 5 years)
6. Piecewise HR for docetaxel plus nintedanib vs. docetaxel (Section 4.2.6)
   Assuming constant HR of OS and PFS for nintedanib from the second period (from week 113) onwards

6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1)

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
Key issue 6. Model structure	4.2.2	Methods	State transition model to validate current PSM results	+/-	No	No
Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in the 2nd line	4.2.4	Methods	Amend model to include platinum-based comparator	+/-	No	Yes
Key issue 8. Docetaxel plus nintedanib modelling approach	4.2.6	Bias & indirectness	Assumed that HR of docetaxel plus nintedanib versus docetaxel cannot exceed 1	+	Partly	Yes
Key issue 9. Treatment waning	4.2.6	Unavailability – immature data	Assumed gradual waning of treatment effect over 5 years, starting at 2-year timepoint	+	Partly	Yes
Key issue 10. TTD modelling approach	4.2.6	Bias & indirectness	Assumed alternative approach using MAIC and parametric distributions	+	Partly	Yes
Key issue 11. Health-related quality of life approach	4.2.8	Unavailability/missing data/small sample sizes	Assumed utilities based on disease progression	+	Partly	Yes
Key issue 12. Disutility for IV infusion	4.2.8	Unavailability of comparative HRQoL data	Excluded disutility	+	Partly	Yes
Key issue 13. RDI and wastage assumption	4.2.9	Unavailability of evidence for the company's assumption	Equal RDI and costs based on opened packs	+	Partly	Yes

[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 ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator; [b] Explored

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
ERG = Evidence Review Group; HR = hazard ratio; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; IV = intravenous; RDI = relative dose intensity; TTD = time to treatment discontinuation						

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.2.1 The results of deterministic ERG preferred base case scenario

In Section 6.1, the ERG base-case was presented, based on various changes compared to the company base-case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously for the primary comparator (docetaxel). The largest impact on the ICER was caused by limiting the treatment effect of sotorasib at 2 years with a gradual waning effect over 5 years after (MJ 8), which resulted in a 10.7% increase of the ICER compared to the CS base-case (£48,332 per QALY vs. £43,660 per QALY), mainly due to a decrease in LYG. The ERG base-base, combining all proposed adjustments, was 33.8% higher than the CS base-case (£58,415 per QALY vs. £43,660 per QALY). The main reasons for this difference were higher drug acquisition costs for sotorasib ([redacted] vs. [redacted] and lower LYG in the post-progression health state ([redacted]).

The impact of each individual change and the combined effect of all changes simultaneously for the secondary comparator (docetaxel + nintedanib) was presented in Table 6.3. Changing the HR forOS (MJ 9) had the largest impact on the ICER, increasing it with 33.7% compared to the CS base- case (£44,969 per QALY vs. £33,628 per QALY). The ERG base-case ICER was 54.8% higher than the CS base-case (£52,051 per QALY vs. £33,628 per QALY). The main reasons for this difference were higher drug acquisition costs for sotorasib (** vs [redacted]d lower LYG gained in thepost-progression health state [redacted] for docetaxel + nintedanib.

Table 6.2: ERG base-case adjustments (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	43,660
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,394
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,496
MJ 4: Including drug wastage (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	46,387
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	47,208
MJ 6: Subsequent treatments based on alternative distribution					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	43,825
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,339
MJ 8: gradual waning of treatment effect over 5 yrs, starting at 2-year timepoint (key issue 9)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	48,332
ERG base-case					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,415

Based on CS updated model

CS = company submission; ERG = Evidence Review Group; FV = fixing violations; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation

Table 6.3: ERG base-case adjustments (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,628
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,909
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,692
MJ 4: Including drug wastage (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,673
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	35,990
MJ 6: Subsequent treatment based on alternative distribution					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,839
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel + nintedanib	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	34,087
MJ 8: gradual waning of treatment effect over 5 yrs, starting at 2-year timepoint (key issue 9)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	33,618
MJ 9: Assuming HR of 1 for OS for nintedanib for the first period (key issue 8)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	44,969
ERG base-case					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,051
Based on CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violations; HR = hazard ratio; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; OS = overall survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation					

6.2.2 The results of ERG sensitivity and scenario analyses

The sensitivity and scenario analyses were performed based on the ERG base-case. The results of the PSA for the ERG base-case were in line with deterministic results for both primary and secondary comparator (Tables 6.4 and 6.5). The probability of sotorasib being cost effective against docetaxel and docetaxel + nintedanib was [redacted]**, respectively. Figures 6.1 to 6.4 show the cost effectiveness plane and cost effectiveness acceptability curve for the docetaxel and docetaxel + nintedanib comparisons.

- The first ERG scenario had a small impact on ICER compared to ERG base-case for both comparators. The reason was that adding disutility of "decreased neutrophils" and "increased aspartate aminotransferase", led to a very minor decrease in incremental QALYs.
- The second ERG scenario with assuming treatment-emergent instead of TRAEs, resulted in slightly higher ICERs compared to the ERG base-case.
- The third ERG scenario slightly increased the ICER for both comparators. The incremental QALYs did not change in this scenario compared to the ERG base-case, however, people spend more time on treatment which led to an increase in costs.
- The fourth and fifth ERG scenario explored different periods for the waning effect of sotorasib. The ICER of sotorasib vs. docetaxel increases when shortening the waning period to 3 years and decreases when applying a waning effect over 7 years. For the secondary comparator however (sotorasib vs. docetaxel plus nintedanib) both the 3 year and the 7 year scenario result in an increase in the ICER. The reason for this is that in the company model, and also in the ERG analysis, the waning effect is applied to nintedanib as well.
- The sixth ERG scenario was explored only for the secondary comparator (docetaxel plus nintedanib). Since in this scenario the HR for OS of docetaxel plus nintedanib versus docetaxel is on average higher than the ERG base case [redacted] vs. ****), the QALYs decreased slightly for docetaxel plus nintedanib which made the ICER just below £50,000 per QALY.

Table 6.4: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	57,567
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,444
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	58,986
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	60,809
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	60,428
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	57,206

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

Table 6.5: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	50,249
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	51,874

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,733
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,851
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,179
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	52,074
ERG scenario 6: Assuming constant HR of OS and PFS for nintedanib from 2nd period onwards					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	49,664

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; QALY = quality-adjusted life year

Figure 6.1: Cost effectiveness plane for ERG base-case (Comparator: docetaxel)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.2: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel)

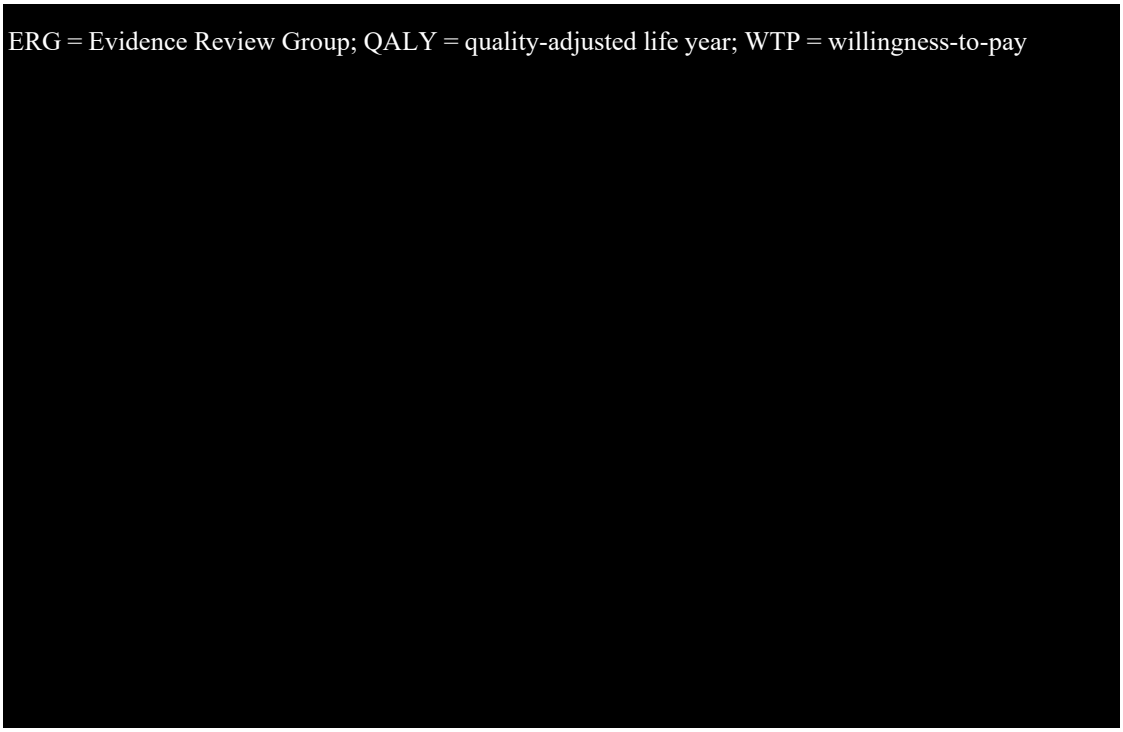


Figure 6.3: Cost effectiveness plane for ERG base-case (Comparator: docetaxel + nintedanib)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.4: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel + nintedanib)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

6.3 ERG's preferred assumptions

In Section 6.2, the results based on the ERG preferred assumptions were presented. The estimated ERG base-case ICERs were £58,415 and £52,051 per QALY gained for sotorasib versus docetaxel and docetaxel + nintedanib, respectively which was 33.8% and 54.8% higher than the CS base-case. The probabilistic ERG base-case analyses indicated that the probability of sotorasib being cost effective was [redacted] against docetaxel and docetaxel + nintedanib, respectively, at a willingness-to-pay (WTP) threshold of £50,000 per QALY gained. Comparing sotorasib to docetaxel, the most influential adjustment in the ERG base-case was limiting the treatment effect to 2 years with a waning effect over 5 years. Comparing sotorasib to docetaxel + nintedanib, the most influential adjustment was setting the HR of OS to 1 for docetaxel plus nintedanib versus docetaxel. Concerning exploratory scenarios, using a generalised gamma distribution for PFS was the most influential scenario, driving the ICER upwards, for both comparisons.

6.4 Conclusions of the cost effectiveness Section

As discussed in Section 4.1.1, the search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use.

Separate sets of searches were conducted to identify cost effectiveness studies, HRQoL studies and healthcare resource use evidence. The eligibility criteria used by the company provided sufficient detail and were suitable to fulfil the company's objective to identify cost effectiveness studies.

The CS was largely in line with the NICE reference case. The CS partly deviated from the scope, however, where it concerned the comparators modelled. More specifically, platinum-based chemotherapy in the 2nd line was excluded as a comparator, while expert opinion indicated that it is a relevant treatment option for a substantial part of the population. Also, the company did not perform a full incremental analysis but instead presented two pairwise comparisons.

Although the ERG agreed that a partitioned survival model seemed appropriate for the decision problem, they would have liked to see a state transition model as a scenario to validate the results of the company's partitioned survival model.

The ERG considered the absence of any waning of the treatment effect in the company model not well justified. Data from the CodeBreaK100 trial are not sufficiently mature to assume a continuous effect of sotorasib. Given the available follow-up in CodeBreaK100 of 18 months (with many patients censored) the ERG believes that implementing a gradual waning of the treatment effect over 5 years, starting from the 2 year point, is a fair and maybe even already optimistic scenario.

The ERG was concerned about the approach taken to estimate treatment duration. Instead of taking a similar approach as for OS and PFS, TTD was linked to PFS via a HR. The ERG was not convinced by the rationale of the company to choose this approach and felt it more consistent to take the same approach for TTD as for OS and PFS, which is to fit parametric models to CodeBreaK100 discontinuation data (weighted based on the MAIC).

A major concern of the ERG was the validity of the modelling approach in the secondary comparison, sotorasib versus docetaxel plus nintedanib. The two-step approach taken potentially introduces bias, of which the fact that modelled OS for docetaxel plus nintedanib was initially below OS for docetaxel may be only one symptom. The ERG believes that the docetaxel plus nintedanib comparison is subject to large uncertainty, beyond what the ERG was able to take into account in their ERG base-case analysis.

Some comments on the incorporation of adverse events in the economic model were made by the ERG, but these could be resolved in the ERG analyses and were of minor importance for overall cost effectiveness results.

With respect to the implementation of health state utility values in the mode, the ERG had some major concerns. Firstly, the datasets used (AN01 and AN02) contained a much smaller number of EQ-5D observations than could have been expected based on the sample size and number of timepoints available for collecting these data, and so the mixed models were based on a sample that may not be representative of the whole population. Furthermore, the time-to-death approach based on AN01 preferred by the company was not justified by statistical arguments, while results for alternative approaches (with AN02 data for instance) were not presented. And because of the preferred time-to-death utilities, the company considered it not necessary to apply an age-related decrement. The company then also proposed to apply a disutility for IV infusion of docetaxel but did not discuss the potential disutility of having to take eight tablets daily for sotorasib. The ERG considered this approach altogether not well justified and feels that substantially more evidence on comparative HRQoL is necessary to be able to resolve these issues.

The ERG considered the company's assumption of no wastage for sotorasib to be unrealistic. Without a specific program in place that would guarantee that opened packs could be returned by the patient and then used by another patient, the cost calculation based on opened packs seems closest to daily practice. The values for RDI and subsequent treatments were deemed to slightly favour sotorasib while not entirely justified, so the ERG adjusted these to be more conservative. For a reliable estimate of subsequent treatments provided after sotorasib, more evidence is warranted.

The ERG made various adjustments to the company base-case. The probabilistic ERG base-case ICER for sotorasib versus docetaxel was [redacted] per QALY gained (based on 1,000 iterations). For sotorasib versus docetaxel plus nintedanib, the ICER was **. The most influential scenario for both comparators was where the generalised gamma distribution for PFS was used instead of the lognormaldistribution, driving the ICER upwards.

In conclusion, cost effectiveness estimates of sotorasib compared with docetaxel and with docetaxel plus nintedanib are subject to considerable uncertainty, mainly because of immaturity of data and lack of comparative evidence in various areas. Even when all the ERG preferred assumptions were implemented in the model, uncertainty remained on a number of issues, such as whether all relevant comparators were included in the analysis, treatment duration and long-term efficacy of sotorasib, and comparative HRQoL values. The comparison for docetaxel plus nintedanib is potentially more heavily biased even because of the indirectness of the two-step approach to model OS and PFS.

7. END OF LIFE

According to the CS, sotorasib in its full anticipated licensed indication as a second- or subsequent line therapy meets the NICE criteria for an end of life medicine, see Table 7.1.[1]

Table 7.1: End of life criteria

Criterion	Data available	Reference in CS (Section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Large real world evidence studies indicate that that OS with non-targeted 2nd line therapies is <10 months, and with 3rd line therapies is <7 months. OS with 2nd line docetaxel monotherapy in the SELECT-1 study was 7.9 months.[16] OS with 2nd line nintedanib plus docetaxel in the LUME-Lung 1 study was 10.9 months.[17]	Section B.1.3.1.2, pages 19-21
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	A robust MAIC indicates sotorasib provides at least an additional [redacted] in median OS compared with docetaxel monotherapy based on available trial data. The economic model estimates that sotorasib plausibly provides an additional undiscounted mean OS of [redacted] months compared with docetaxel monotherapy and [redacted] months compared with nintedanib plus docetaxel*.	Section B.2.9.4.1, page 57 Section B.2.9.4.2, page 60

Based on Table 20 of the CS[1]
* Derived from economic model with 20-year time horizon, values undiscounted (see Section B.3.3.5 of the CS for how comparison of sotorasib vs nintedanib plus docetaxel is implemented)
CS = company submission; MAIC = matching adjusted indirect comparison; NHS = National Health Service; OS = overall survival

ERG comment: The ERG considers the first criterion, life expectancy less than 24 months, to be met.

Regarding the second criterion, extension of life of ≥3 months, the ERG agrees that, based on the data cited by the company, the criterion has been met. However, as discussed in Section 3.3 and 3.4, the ERG has concerns regarding the validity of the indirect comparisons referred to by the company, see key issue 5.

8. REFERENCES

[1] Amgen Ltd. *Sotorasib for previously treated KRAS p.G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*. Cambridge: Amgen, 2021. 140p.

[2] National Institute for Health and Care Excellence. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer. Final scope (Appendix B) [Internet]*. London: NICE, 2021 [accessed 20.7.21]. 6p. Available from: https://www.nice.org.uk/guidance/gid-ta10639/documents/final-scope

[3] Sebastian M, Eberhardt WEE, Hoffknecht P, Metzenmacher M, Wehler T, Kokowski K, et al. KRAS G12C-mutated advanced non-small cell lung cancer: a real-world cohort from the German prospective, observational, nation-wide CRISP Registry (AIO-TRK-0315). *Lung Cancer* 2021;154:51-61.

[4] Hong DS, Fakih MG, Strickler JH, Desai J, Durm GA, Shapiro GI, et al. KRAS(G12C) Inhibition with Sotorasib in Advanced Solid Tumors. *N Engl J Med* 2020;383(13):1207-1217.

[5] Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, et al. Sotorasib for lung cancers with KRAS p.G12C mutation. *N Engl J Med* 2021;384(25):2371-81.

[6] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: response to request for clarification from the ERG*: Amgen, 2021. 77p.

[7] Amgen. A phase 1/2, study evaluating the safety, tolerability, PK, and efficacy of AMG 510 in subjects with solid tumors with a specific KRAS mutation (CodeBreaK 100). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2018- [cited 2021 Sep 1]. Available from: http://clinicaltrials.gov/show/NCT03600883. NLM Identifier: NCT03600883

[8] Clark P. *ID3780 Sotorasib for previously treated KRAS G12C NSCLC - comparator query. Clinical expert opinion [Personal communication]*.

[9] National Institute for Health and Care Excellence. *Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer: technology appraisal guidance TA347 [Internet]*. London: NICE, 2015 [accessed 13.8.21]. 49p. Available from: www.nice.org.uk/guidance/ta347

[10] Schulz C, Gandara D, Berardo CG, Rosenthal R, Foo J, Morel C, et al. Comparative Efficacy of Second- and Subsequent-line Treatments for Metastatic NSCLC: A Fractional Polynomials Network Meta-analysis of Cancer Immunotherapies. *Clin Lung Cancer* 2019;20(6):451-460 e5.

[11] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[12] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019) [Internet]*: Cochrane, 2019 [accessed 6.7.20] Available from: https://training.cochrane.org/handbook

[13] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: appendices*. Cambridge: Amgen, 2021. 156p.

[14] Sterne JA, Hernan MA, Reeves BC, Savovic 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.

[15] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC With KRAS p.G12C mutation (CodeBreaK 100). Clinical study report CodeBreak100 efficacy data (01 December data cut). Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 17 February 2021. 105p.

[16] Janne PA, van den Heuvel MM, Barlesi F, Cobo M, Mazieres J, Crino L, et al. Selumetinib Plus Docetaxel Compared With Docetaxel Alone and Progression-Free Survival in Patients With KRAS-Mutant Advanced Non-Small Cell Lung Cancer: The SELECT-1 Randomized Clinical Trial. *JAMA* 2017;317(18):1844-1853.

[17] Reck M, Kaiser R, Mellemgaard A, Douillard J-Y, Orlov S, Krzakowski M, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol* 2014;15(2):143-155.

[18] Boehringer Ingelheim Limited. Vargatef 100 mg soft capsules [Internet]. electronic Medicines Compendium (eMC), 2021 [accessed 1.9.21]. Available from: https://www.medicines.org.uk/emc/product/3647

[19] Amgen Inc. *Data on File: Amgen NSCLC UK Advisory Board - 11th February 2021*, 2021

[20] Amgen Inc. *Data on file: a retrospective cohort study of patient characteristics and treatment outcomes among patients with advanced non-small cell lung cancer in the Flatiron Health-Foundation Medicine Clinico-Genomics Database (Study 20200132)*, 2020

[21] 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, 2015 [accessed 1.9.21] Available from: http://www.nicedsu.org.uk

[22] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU technical support document 19. Partitioned survival analysis for decision modelling in health care: a critical review [Internet]*. Sheffield: Decision Support Unit, 2017 [accessed 9.7.21]. 72p. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf

[23] Amgen Inc. *Draft Summary of Product Characteristics: LUMYKRAS (Sotorasib)*, 2021

[24] Amgen Inc. *Data on File: CodeBreaK100 Safety data (01 December 2020 data cut)*. Thousand Oaks (CA): Amgen, 2021

[25] Aggarwal S, Hsu H, Tu H, Carrigan G, Wang X, Ngarmchamnanrith G, et al. Clinicopathological characteristics, treatment patterns, and outcomes in patients with KRAS p.G12C mutant advanced non-small cell lung cancer in the Flatiron Health-Foundation Medicine Clinico-Genomic Database. Poster presented at the European Society for Medical Oncology (ESMO) Virtual Congress; 19-21 September 2020. 2020.

[26] National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management. NICE guideline NG122 [Internet]*. London: NICE, 2019 [accessed 20.7.21]. 41p. Available from: www.nice.org.uk/gidance/ng122

[27] National Institute for Health and Care Excellence. Lung cancer overview [Internet]. London: NICE, 2021 [accessed 13.8.21]. Available from: https://pathways.nice.org.uk/pathways/lung-cancer

[28] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report CodeBreak100 full primary analysis Mar 2021 data cut. Data on file* Thousand Oaks (CA): Amgen, 15 March 2021

[29] EuroQoL. EuroQol Instruments - EQ-5D-3L, EQ-5D-5L, EQ-5D-Y [Internet]. 2021 [accessed 30.4.21]. Available from: https://euroqol.org/

[30] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: addendum. March 2021 data cut updates*. Cambridge: Amgen, 2021. 25p.

[31] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report supplementary PROs primary analysis (September 2020 data cut). Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 12 November 2020. 81p.

[32] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[33] Lewis G, Peake M, Aultman R, Gyldmark M, Morlotti L, Creeden J, et al. Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. *J Int Med Res* 2010;38(1):9-21.

[34] National Institute for Health and Care Excellence. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: Clarification letter*. London: NICE, 2021

[35] NHS. *National Cost Collection for the NHS [Internet]*. London: NHS, 2021 [accessed 13.8.21] Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819

[36] Medicines Complete. British National Formulary [Internet]. London: Medicines Complete, 2021 [accessed 13.8.21]. Available from: https://about.medicinescomplete.com/publication/british-national-formulary/

[37] Personal Social Services Research Unit. *Unit costs of health and social care 2019 [Internet]*. Canterbury: PSSRU, University of Kent, 2019 [accessed 25.8.21] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/

[38] Department of Health and Social Care. Drugs and pharmaceutical electronic market infrmationo tool (eMIT) [Internet]. London: Department of Health and Social Care, 2021 [accessed 13.8.21]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[39] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report CodeBreak100 full primary analysis Sept 2020 data cut. Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 10 November 2020. 107p.

[40] Amgen Inc. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Amgen study 20170543 protocol*. Thousand Oaks (CA): Amgen, 2018

[41] Northern Cancer Alliance. Lung chemotherapy protocols and prescriptions [Internet]. Newcastle Upon Tyne: Norther Cancer Alliance, 2021 [accessed 13.8.21]. Available from: https://northerncanceralliance.nhs.uk/advisory_group/lung-expert-advisory-group/lung-chemotherapy-protocols-and-prescriptions/#

[42] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.6.15]. 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

[43] National Institute for Health and Care Excellence. *Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy: technology appraisal guidance TA374 [Internet]*. London NICE, 2015 [accessed 20.7.21]. 67p. Available from: www.nice.org.uk/guidance/ta374

[44] National Institute for Health and Care Excellence. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy: technology appraisal guidance TA428 [Internet]*. London: NICE, 2017 [accessed 20.7.21]. 28p. Available from: www.nice.org.u/guidance/ta428

[45] 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.

[46] 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.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and ERG report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the ERG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on Wednesday 15 September** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and separately highlight information that is submitted as 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

Issue 1 Licence confirmed

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
If you wish, can now provide full MHRA conditional licence received on 8th Sept (e.g. on page 13 box among other places such as P31)	Replace anticipated with confirmed conditional MHRA licence wording, which is the following: *"LUMYKRAS is indicated as monotherapy for the treatment of adult patients with KRAS G12C mutated locally advanced or metastatic non-small cell lung cancer (NSCLC), who have progressed on, or are intolerant to, platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy."* Could also adjust the wording such as: *"Further evidence should be gathered to cover the population defined in the NICE scope."* It is unusual to suggest the licence and evidence should fit the NICE scope – the NICE scope cannot be beyond the licence.	There is now more certainty around licence indication wording.	Not a factual inaccuracy. The wording in the ERG report was correct at the time of writing.

Issue 2 Proportion of patients of different races

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the UK. The	The trial population for Asian patients (and other races): Race: 101 white (81.5%); 5 other (4.0%); 6 black (4.8%), 12 Asian (9.7%).	Typographical error	ERG report revised. Where applicable, the ERG now reads:

generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of participants from Japan or South Korea (15.1% of the sample).			"The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample)."

Issue 3 High number of serious adverse events

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The ERG notes the "High number of serious adverse events observed in CodeBreaK100", but this statement should be contextualised – it's not clear this would be any lower with the comparators (e.g. docetaxel) and may well be higher under chemotherapy regimens. Further context should be added to denote treatment-related specifically and the fatal adverse events were not considered related to the investigational product by the investigator.	The relevant information for sotorasib is as follows: Fifty-nine subjects (47.6%) had serious adverse events. The most frequently reported (in ≥ 4 subjects) were pneumonia (9 subjects, 7.3%); metastatic lung cancer (7 subjects, 5.6%); respiratory failure (6 subjects, 4.8%); pleural effusion and increased ALT (5 subjects each, 4.0%); and increased AST (4 subjects, 3.2%). **Twenty-one subjects (16.9%) had fatal adverse events; none were considered related to investigational product by the investigator**. Treatment-related Eighty subjects (64.5%) had treatment-related adverse events. The most frequently reported (in ≥ 4 subjects) were diarrhoea (28 subjects, 22.6%); increased ALT and increased AST (22 subjects each, 17.7%); fatigue (14 subjects, 11.3%); nausea (13 subjects, 10.5%);	Suggests serious AE profile with Sotorasib is unusually high, but it is important to contextualise between treatment-related adverse events and other types of adverse events.	Not a factual inaccuracy. The observation was based on the information presented in the company submission. However, the ERG corrected a reference to an incorrect section in the Table on key issue 4 (3.2.4.5 rather than 2.1).

	increased ALP (10 subjects, 8.1%); abdominal pain (8 subjects, 6.5%); vomiting (7 subjects, 5.6%); decreased appetite (6 subjects, 4.8%); decreased lymphocyte count (5 subjects, 4.0%); and anaemia, dry mouth, increased bilirubin, and rash (4 subjects each, 3.2%). Add context about no direct evidence this is any more than chemotherapy (table 1.5, p15 and table 3.11 and accompanying sections).		

Issue 4 Variables considered vs possible in MAIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Adapt wording around what variables were not considered for MAIC with SELECT-1. Variables can be considered, but not used because of data constraints.	Adapt the following in table 1.6 (16) and others (e.g. top and bottom of p56): "However, these, alongside G12C mutation status, were not considered for the MAIC comparing CodeBreaK100 and SELECT 1." To reflect that there were no data available to include G12C in the MAIC – there are no specific G12C subgroup baseline variables published for SELECT-1 and so inclusion is not possible but was considered. To also reflect that there is no reported proportion of brain mets in SELECT-1 and so this could not be included in the MAIC.	Text at the moment suggests these 2 variables could be included in the MAIC analyses, but in practical terms it is impossible given published data.	Not a factual inaccuracy.

Issue 5 PSWA docetaxel only analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Adjust language in table 1.6 – "For the PSWA, limiting to the docetaxel only population could be informative" – to reflect that this may not be feasible given the data and methods for PSWA.	CS appendix table 12 shows that there are 21 patients (around 10%) and around half of this number in the G12C dataset of flat iron and so a PSWA analysis with only docetaxel patients is not feasible (patient numbers would be weighted away during the analysis and many variables that are relevant could not be adjusted for). Suggest, adding a following statement reflecting data challenge with this.	Add some acknowledgement this analysis not possible or very challenging.	Not a factual inaccuracy.

Issue 6 RCT pragmatism and trial norm

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On page 31, consider caveating the statement that it is not clear why CB100 did not include ECOG2 patients, with clinical consensus that this is not unusual in NSCLC clinical trials and is a pragmatic approach (we need reasonably healthy patients so that there is enough time for meaningful survival data to be produced).	Add a caveat sentence/statement reflecting pragmatic nature of previous NSCLC trials.	Should acknowledge pragmatism needed in NSCLC trials (near end of life).	Not a factual inaccuracy.

Issue 7 Context around clinical estimates

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
In table 46 (p73) the ERG suggests expert clinical opinion suggests 40% of patients could be eligible for PDC. However, Amgen has consulted with many clinicians and although estimates vary 40% is very large.	Caveat to suggest this is only one consulted estimate and that there will be variation in estimates. If possible, also to suggest that the population of patients who have received only previous IO (and not PDC) is shrinking and conversely the population with previous IO+PDC is growing (this has been validated with a number of clinicians that Amgen consulted).	40% is high compared with the estimates of other clinicians.	Not a factual inaccuracy.

Issue 8 PFS scenario minor discrepancy 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On p75 and p88 the ERG report discusses using the generalised gamma (GG) restricted (i.e. joint fitted) model instead of log-normal (LN) for PFS, however we believe there is a minor typo or some such	The quoted ICERs for the scenario with restricted GG (£45,628 on page 88 and the scenario analysis section) are obtained by setting only the Sotorasib arm function to restricted GG distribution. But that means the Sotorasib model engine picks up the sotorasib arm of a joint/restricted GG fitted distribution, but the docetaxel engine continues to pick up the docetaxel arm of a joint/restricted LN fitted model. To obtain the right ICER you should select both restricted GG functions for both Sotorasib and docetaxel and this gives the correct ICER for a joint/restricted GG selection (=£45,123).	The quoted ICER does not fit the scenario.	ERG report revised, on page 88 it now reads: “Generalised gamma distribution selected to estimate long-term PFS (ICER: £45,123 per QALY gained).”

Issue 9 PFS scenario minor discrepancy 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On p88 the following scenario is described and ICER given: "Joint (unrestricted) lognormal distribution selected to estimate long-term PFS (ICER: £49,745 per QALY gained)" But we cannot replicate that value.	We believe the scenario described is log-normal (LN) independent fit (i.e. as opposed to joint fit model with treatment arm 0//1 dummy variable) for the Sotorasib arm and log-normal independent fit for docetaxel arm as well (for PFS only). You can obtain this by setting both arm PFS options to "unrestricted parametric distribution" and selecting LN for each and the appropriate ICER is £47,917.	The quoted ICER does not fit the scenario.	ERG report revised, on p88 it now reads: "Joint (unrestricted) lognormal distribution selected to estimate long-term PFS (ICER: £47,917 per QALY gained)"

Location of incorrect marking	Description of incorrect marking	Amended marking
Give full details of inaccurate marking - document title and page number	Give details of incorrect confidential marking	Please copy the impacted section here, with your amended marking.

(Please add further lines to the table as necessary)

NICE National Institute for Health and Care Excellence

Technical engagement response form

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments by **5pm** on **22 October 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. If you would like to comment on issues in the ERG report 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 appraisal, 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.
- Do not use abbreviations.

NICE National Institute for Health and Care Excellence

- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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 replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

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.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Amgen Limited**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	**Does this response contain new evidence, data or analyses?**	Response
Key issue 1: Population narrower than NICE scope	NO	As agreed during the technical engagement call, the initial NICE scope population can sometimes differ from the final license population, but the final NICE recommendation cannot be wider than the license indication. **This issue is considered resolved following the technical engagement call.**
Key issue 2: Generalisability / lack of UK participants	NO	As agreed in the technical engagement call it is not unusual to have small or no numbers of UK patients in trials of targeted NSCLC treatments. The CodeBreak 100 trial is broadly generalisable to the relevant population in this appraisal. Amgen has consulted clinicians, who practice NSCLC in large centres in England, and they have reported that the population demographics in CodeBreak 100, including ethnicity, is largely representative of patients treated in their clinics. **This issue is considered resolved following the technical engagement call.**
Key issue 3: High risk of bias of CodeBreaK100	NO	A serious adverse event in CodeBreak100 was defined as an adverse event that meets at least 1 of the following criteria: fatal, life threatening (places the subject at

		immediate risk of death), requires in patient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, congenital anomaly/birth defect, or other medically important serious event. **As agreed in the technical engagement call it is important to note that serious adverse events do not necessarily need to be related to treatment, and hence treatment-related adverse events are considered more relevant here.** Serious adverse reactions occurred in 50% of patients treated with sotorasib. Serious adverse reactions in ≥ 2% of patients were pneumonia (8%), hepatotoxicity (3.4%), and diarrhoea (2%). Fatal adverse reactions occurred in 3.4% of patients who received sotorasib due to respiratory failure (0.8%), pneumonitis (0.4%), cardiac arrest (0.4%), cardiac failure (0.4%), gastric ulcer (0.4%), and pneumonia (0.4%). Treatment-related adverse events are those adverse events of any grade that were considered by the investigators to be related to treatment. In CodeBreak 100 a total of 88 patients (69.8%) reported adverse events of any grade that were considered by the investigators to be related to treatment. (treatment-related adverse events). The worst grade of treatment-related adverse event was grade 3 in 25 patients (19.8%) and grade 4 in 1 patient (0.8%; pneumonitis and dyspnoea); no treatment-related adverse events of grade 5 (deaths) were reported.

NICE National Institute for Health and Care Excellence

		The most frequent treatment-related adverse events were diarrhoea (in 40 patients [31.7%]), nausea (in 24 [19.0%]), increase in the alanine aminotransferase level (in 19 [15.1%]), increase in the aspartate aminotransferase level (in 19 [15.1%]), and fatigue (in 14 [11.1%]). **This issue is considered resolved following the technical engagement call.**
Key issue 4: High number of serious adverse events observed in CodeBreaK100	NO	As agreed in the technical engagement call, the risk of bias associated with CodeBreak100 is broadly in line with other pivotal 1-arm trials in NSCLC that have been the basis of previous NICE appraisals. Issues related to concealment/blinding and confounding are inherent in 1-arm trials of this nature and hence the need for statistical methods such as MAIC and PSWA. **This issue is considered resolved following the technical engagement call.**
Key issue 5: Validity of ITC without a common comparator	NO	<u>Base-case MAIC for comparison with docetaxel (using SELECT-1 RCT)</u> Amgen conducted a MAIC that weighted CodeBreak100 patients based on BL characteristics to match the docetaxel arm of the SELECT-1 trial. The variables for inclusion were selected based on literature review and 6 individual interviews with clinicians and the resulting list is in line with previously presented MAICs in NSCLC. As agreed in the technical engagement call, there was exclusion of some variables from the MAIC which could potentially be treatment effect modifiers (brain

NICE National Institute for Health and Care Excellence

		metastases, KRAS G12C mutation status). However, as agreed given the data available it was not possible to include these in the MAIC: • SELECT-1 did not report the proportion of (inactive) brain metastases at baseline (trial reports, publications and appendices, HTA submissions) and so it was not possible to include as a matching variable in the MAIC. o All trials exclude *active* brain metastases, which clinicians suggest is more likely to be a modifier. o The proportion of patients with brain metastases was higher in CodeBreaK100 (21%) than in LUME-Lung 1 (8%). If SELECT-1 had a similar proportion of inactive brain metastases to CodeBreak100, any bias would favour the docetaxel arm and make cost-effectiveness results conservative. • It is not feasible to match on KRAS G12C in the base-case MAIC. 100% of CodeBreak100 are KRAS G12C mutation positive (42% in SELECT-1 with the remaining having KRAS mutations other than G12C) and so a MAIC would "weight away" the sample of CodeBreak100. <u>Supplementary analysis – propensity score weighting analysis (PSWA) using Amgen Flatiron RWE database</u>

		This analysis aimed to compare OS and PFS with sotorasib against standard of care chemotherapy observed in a cohort of patients with previously treated KRAS-mutated advanced/metastatic NSCLC in the Flatiron database. The PSWA used a flat-iron RWE dataset that is a basket of chemotherapies, of which around 40% received singlet docetaxel or a doublet containing docetaxel. Some of the concerns raised by the ERG were clarified in the technical engagement call: • It is not the case that only 4th line patients were included in the PSWA analysis dataset. Patients were selected based on their last line of treatment (2nd, 3rd, 4th line) and if data was available in the 5th line or later only the patient's data up to the 4th line was selected (i.e. broadly in line with inclusion criteria of CodeBreak100). • The ERG suggested exploring alternative methods for calculating a treatment effect from PSWA: ATE (average treatment effect) instead of the presented ATT (average effect of the treatment on the treated). Taking the ERGs advice Amgen can report that switching to the former made little difference in relative effectiveness and so little difference to the PSWA scenario analysis ICER. <u>Comparison of base-case MAIC and Flatiron PSWA</u> Amgen agrees with the ERG assessment that it is difficult to assess which analysis is more robust (or less biased). There are some subtle trade-offs that make it unclear and may even favour the Flatiron PSWA:

		• Both data sources are fundamentally observational (i.e. unrandomised): a single arm of a controlled trial (SELECT-1) vs. an uncontrolled (but larger sample) historical control cohort (Flatiron). • Neither source has any UK patients: CodeBreak100 is a multinational Ph2 trial and Flatiron is based on an American health record database. • The MAIC analysis must weight CodeBreak100 patients to match SELECT-1 and the assumption is made that this treatment effect translates to the CodeBreak100 population, whereas the PSWA weights the Flatiron data to match the CodeBreak100 population. • In the real-world disease progression is derived from physician notes in a clinical practice setting and may be informed by RECIST criteria in conjunction with other signs of progression. • Given data availability and richness of Flatiron data, the PSWA allowed consideration of and final inclusion of more weighting covariates and is therefore more heavily adjusted. o The same clinician elicitation exercise as for the MAIC informed selection of “very important” and “somewhat important” covariates but now the data available allowed all “somewhat important” to be included in the covariate selection algorithm (i.e. could potentially be included in the final model). o Therefore, the final analysis included several covariates not in the MAIC analysis (brain metastases, presence of non-KRAS

NICE National Institute for Health and Care Excellence

		mutations, prior lines of treatment, type of prior treatments, Albumin). **For these reasons, the PSWA is also presented alongside the base-case MAIC below for the committee's consideration.**
Key issue 6: Partitioned Survival Model structure not validated or justified	NO	As agreed in the technical engagement call, it is not always feasible (and not conventional) to present two fundamentally different modelling methodologies for validation. As argued previously, the fundamental problems of the Partitioned Survival Model are unlikely to be resolved by a State Transition Model and the data requirements of such a model are harder to meet. **This issue is considered resolved following the technical engagement call.**
Key issue 7: Exclusion of platinum-based chemotherapy as a comparator in 2nd line	YES	There is a reasonably broad consensus among consulted clinicians and NHSE that the optimal initial therapy for patients in NSCLC without current actionable mutations is anti-PD-1/PD-L1 immunotherapy (IO) in combination with platinum-based chemotherapy (i.e. platinum doublet chemotherapy or "PDC"). This is also in-line with the NICE pathway and historical NICE recommendations. The group of patients who are eligible for Sotorasib that will have received both an IO and PDC previously is growing (and those that have only received an IO shrinking). Therefore, for most patients Sotorasib will displace docetaxel in the pathway.

NICE National Institute for Health and Care Excellence

		Nevertheless, if PDC were considered a minor comparator an unanchored MAIC would not be possible. As agreed in the technical engagement call, Amgen can confirm that the SLR did not identify any KRAS population trials with a PDC arm (see appendix A below). The results of the SLR also support the proposition that patients in this population have been pre-treated with PDC (as does the baseline characteristics of the pivotal trial CodeBreak100 with 90% of patients having been pre-treated with PDC). As additional evidence, in Appendix B a retrospective analysis using data from Oncology Dynamics ™ confirms that most patients who received docetaxel recently in the UK are likely to have received IO and PDC previously. The PSWA could be considered a reasonable proxy for a comparison with PDC, given that the most common regimen in the chemotherapy basket was platinum-based chemotherapy (but still under 1/3 of patients). **This issue is considered resolved following the technical engagement call.**
Key issue 8: Docetaxel plus nintedanib modelling approach leading to worse survival	NO	A piecewise approach to generating a nintedanib treatment effect (vs. docetaxel alone) by fitting HRs to 3 periods was undertaken (via Cox PH models) in the base-case model. This is because curve diagnostics suggested that the LUME-

		Lung 1 OS curves did not satisfy the proportional hazards assumption and so a single fitted HR was deemed inappropriate (KM curves shown in Appendix C). Amgen agree with the ERG that it is less clear that the proportional hazards assumption is violated at 26 months (compared with 6 months) and so a scenario with piecewise HRs for only 2 periods is worth exploring (i.e. 0-6 and 6+ months). The ERG proposes setting the HR in the first period (0-6m) to 1 and so assuming equal survival between docetaxel and add-on nintedanib. However, Amgen find it highly irregular to invalidate measured trial data from a published 2-arm phase 3 trial. Trial data points are usually considered more valid than intuitive assumptions in the hierarchy of evidence. Sometimes the impact of a treatment is nuanced and Kaplan-Meier (KM) curves can cross, but this should be reflected in any fitted treatment effects. This is particularly so when the crossing occurs with many patients at risk (i.e. the sample size is higher and the curves more reliable at the beginning of a KM). **This issue only impacts the minor comparison with nintedanib plus docetaxel. This issue is considered unresolved following the technical engagement call and the ICER ranges presented below will reflect this.**
Key issue 9: No waning of treatment effect	NO	Amgen believe that although treatment effect waning (TEW) can be useful to explore model sensitivities, it is a relatively blunt tool and should be applied considering the particulars of each case (as requested Appendix D below presents the hazard plot for context).

		As argued previously, there are several reasons why TEW should be limited and applied carefully in this case: • To a large extent, the impact of discontinuation on OS and PFS has already been "baked" into the hazard function (and so projected survival estimates) because within the trial period (in which parametric curves have been fitted) a significant number of patients have discontinued over the period (>80%). • TEW is more intuitive and more easily defendable when the two treatments are comparable and have a similar mechanism of action and so a reasonable assumption can be made that the relationship between being on treatment and benefiting longer term are similar (e.g. two EGFR targeting TKI therapies). However, sotorasib and docetaxel are very different medicines with different actions and therefore such an assumption is more uncertain. • A case can be made that the mix of patient at the point of the March data cut (around 15 months of follow-up) in the sotorasib arm is in a better average "health state" than the docetaxel patients and so the hazards of survival will continue to be better in the former for some time. o Before the extrapolated portion (i.e. within trial period), 1/2 of Sotorasib patients have yet to progress, but for docetaxel it is only <1/6.

NICE National Institute for Health and Care Excellence

		▪ In the docetaxel arm of the model at around 15 months (i.e. the trial period of SELECT-1 and not an extrapolated portion), of the <30% alive only 4% points of patients are progression free (the remaining progressed). In contrast, by this point around half of the patients in Codebreak100 who are alive (i.e. around 40%) have yet to progress (20% points of patients). • According to Appendix E of the company submission and the related publication, by the time of the March 2021 data cut-off around 80% (81.7%) of patients have discontinued treatment, around 40% remain alive and around 20% have yet to progress. Therefore, half of patients who are alive will have remained on sotorasib treatment at this point. o Sotorasib is given in CodeBreak100 until progression or the development of unacceptable AEs and so it is inappropriate to apply TEW early when a significant proportion of those alive are still benefiting from treatment. o Applying TEW in too blunt a fashion would bias cost-effectiveness results in that sotorasib arm patients continue to accrue the costs of treatment but not the relative benefits of the treatment. **This issue is considered unresolved and the ICER ranges presented below will reflect this.**

NICE National Institute for Health and Care Excellence

Key issue 10: TTD modelling approach inconsistent with OS and PFS modelling	NO	As described in the response to CQs, the base-case approach of connecting TTD to PFS by applying a fitted HR is reasonable and consistent with the clinical use of Sotorasib: treatment is continued until progression (or development of unacceptable AEs) with the majority discontinuations in CodeBreak100 being due to progression. In a similar way and for similar reasons, applying a HR to PFS, assumptions such as TTD=PFS or adding a mean number of cycles of treatment to newly progressed patients have been accepted in previous NSCLC appraisals. TTD is also very mature with >80% of patients having discontinued by the March 2021 data cut. This relative certainty means that applying parametric curves has a limited impact on the ICER and so the ICER ranges below reflect the ERG base-case selection (fitted TTD parametric curves). **The issue is considered resolved following the technical engagement call and the ICER ranges below reflect this.**
Key issue 11: Time-to-death utilities do not seem well-informed	NO	Amgen believe that both a time-to-death approach and a health state approach with a PFS utility differential (see issue 12 below) are plausible. As argued in the response to CQs, two independent sets of interviews were conducted with clinicians to validate visually the time-to-death and health state approaches and clinicians tended to favour time-to-death as more of a driver than health state (based on RECIST defined progression). The ERG had some concerns over the utility analyses:

		• There are limitations to all datasets, but for context: of the 123 patients in the safety dataset, 122 completed at least one eq-5d questionnaire (AN01), and of these 86 completed at least two including at baseline (AN02). Therefore, AN01 includes AN02 as a subset and was used as the basis of utility analyses to maximise sample size. • The ERG is correct that including baseline utility as a covariate in the MMRM models would mean we would be using the AN02 dataset, but although this was a *statistically* significant covariate it did not have a significant impact on results when excluded (e.g. table 33 in the company submission shows it had minimal impact on estimated disutility on progression) and so the larger AN01 dataset was deemed more appropriate. o Although some trade-off exists (sample size vs variable inclusion), all MMRM models already include a patient level random effect that takes account of correlations between observations of the same patient. This in a way has already adjusted for patients with baseline utility and may explain why inclusion of the covariate had little impact. **This issue is considered unresolved following the technical engagement call. A time-to-death approach is plausible and this is reflected in the ICER ranges below.**

NICE National Institute for Health and Care Excellence

Key issue 12: Disutility for IV administration not well justified	NO	Amgen agrees that this alone seems an arbitrary reason for a utility decrement, but as explained in the technical engagement call this should be seen in the context of an overall conservative HRQoL base-case. The company submission assumed equal on treatment PFS utilities for a targeted therapy (Sotorasib) vs. chemotherapy (docetaxel) but a differential is often seen in trials and accepted by committees (e.g. NICE TA628, TA416, TA406 and TA422). • This is the norm in other appraisals for targeted therapies in NSCLC. For example, a differential of 0.02 to 0.08 has been seen for ALK targeted therapies compared with chemotherapy. o For example, see TA628: “This was found in PROFILE 1007, where utilities for the ALK TKI crizotinib (0.82, 95% CI: 0.79−0.85) were significantly greater (p<0.05) than for PDC (0.73, 95% CI: 0.70−0.79)…within the HRQoL SLR…for four of these studies, a comparison between ALK TKIs and chemotherapy was available and, in all instances, a utility decrement was applied for patients on chemotherapy compared to those receiving treatment with an ALK TKI (0.02–0.08).” o Applying AE decrements in only the first cycle of the model is not usually considered double counting. These are not expected to make a difference (compared with differential on treatment PFS

		utilities) – as expected, removing these decrements from the base-case model has virtually no impact. The HRQoL SLR did not identify any KRAS specific utility data. PFS specific utilities from SELECT-1 are not available but there is a PFS utility available from LUME-Lung 1 (0.687 which is used in TA347 and TA416). When this is applied to the PFS base-case utility it implies a decrement of 0.047 (= 0.734 – 0.687). Given this, Amgen believe that scenarios with a health state utility approach and a 0.025 or 0.04 PFS (on treatment) utility differential between arms are reasonable compromises to explore. **This issue is considered unresolved following the technical engagement call. A time-to-death approach is plausible, but a reasonable compromise may be a health state approach with PFS utility differential (scenarios with 0.025 and 0.04) and these are reflected in the ICER ranges below.**
Key issue 13: Relative dose intensity and wastage assumption not justified	NO	Amgen does not believe it is appropriate that relative dose intensity's (RDIs) should be equalised, given that this invalidates measured trial data. Trial data is usually considered more valid than intuitive assumptions in the hierarchy of evidence. **The issue of wastage is considered resolved following the technical engagement call (and inclusion of wastage is reflected in the ICER ranges**

NICE National Institute for Health and Care Excellence

		below). The RDI assumption is not resolved and the ICER ranges presented below reflect this (small impact).

Additional issues

Please use the table below to respond to additional issues in the ERG report 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 appraisal (e.g. at the clarification stage).

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: N/A	N/A	N/A	N/A

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

The table below reflects the updated confidential PAS for Sotorasib [redacted] *and explores the spread of deterministic ICERs implied by the remaining unresolved issues that are relevant to the main comparator docetaxel. The 4 mix of settings at the bottom of the table reflect the following unresolved issues:*

- *Issue 5: the alternative PSWA ICER is shown alongside the base-case MAIC ICER (last column) and this should be considered in decision making*
- *Issue 9 (TEW): the ICER ranges reflect TEW from years 2 (ERG preference), 3 and 4*
- *Issue 11 and 12 (utilities): the ICER ranges reflect health state (HS) utilities with no differential (ERG preference), HS utilities with a 0.025 PFS differential between arms, HS utilities with a 0.04 PFS differential between arms and time-to-death utilities (no decrement or differential)*
- *Issue 13 (RDI): the ICER ranges reflect assumed equal RDI (ERG preference) or RDI from trials (i.e. not equalised)*

Results for the minor comparator nintedanib+docetaxel are shown in Appendix E. Codebreak200 data is currently unavailable, but some PFS data (and potentially interim OS) may be available in 2022 depending on accrual of events (and thus unblinding) – estimated final completion is 2026 (up to 5 years of follow-up).

It should be noted that the probabilistic ICER (vs. docetaxel) is consistently lower by around £800 and this should be considered in decision making. Amgen is confident that this new PAS and the implied range of plausible ICER will allow the Committee to consider a positive recommendation (via baseline commissioning).

NICE National Institute for Health and Care Excellence

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case ICER
Original submitted company base-case (with updated Sotorasib PAS)			ICER (base-case MAIC): **£38,715** ICER (PSWA analysis): **£33,811**
Updated base-case reflecting resolved issues (with updated Sotorasib PAS)	Original submitted base-case used HR applied to PFS to calculate TTD and assumed no wastage.	TTD Parametric curves and wastage included (ERG preferences)	Change from original base-case ICER: +£3,155 ICER (base-case MAIC): **£41,870** ICER (PSWA analysis): **£33,570**
1. ERG preferred settings: health state utilities with no PFS differential, RDI assumed equalised,	Original submitted base-case used time-to-death method for utilities (with IV decrement), RDI from trials (i.e. not equalised) and no TEW.	TEW from 2 years (ERG preferred base-case)	Change from original base-case ICER: +£13,054 ICER (base-case MAIC): **£51,769** ICER (PSWA analysis): **£43,412**
		TEW from 3 years	Change from original base-case ICER: +£11,337 ICER (base-case MAIC): **£50,052** ICER (PSWA analysis): **£41,791**

NICE National Institute for Health and Care Excellence

while varying TEW		TEW from 4 years	Change from original base-case ICER: +£10,332 ICER (base-case MAIC): **£49,047** ICER (PSWA analysis): **£40,802**
2. Health state utilities with lower PFS differential (0.025), RDI not equalised, while varying TEW		TEW from 2 years	Change from original base-case ICER: +£11,317 ICER (base-case MAIC): **£50,032** ICER (PSWA analysis): **£41,680**
		TEW from 3 years	Change from original base-case ICER: +£9,690 ICER (base-case MAIC): **£48,405** ICER (PSWA analysis): **£40,164**
		TEW from 4 years	Change from original base-case ICER: +£8,737 ICER (base-case MAIC): **£47,452** ICER (PSWA analysis): **£39,238**
3. Health state utilities with higher PFS differential (0.04), RDI		TEW from 2 years	Change from original base-case ICER: +£11,554 ICER (base-case MAIC): **£50,269** ICER (PSWA analysis): **£41,757**

NICE National Institute for Health and Care Excellence

equalised, while varying TEW		TEW from 3 years	Change from original base-case ICER: +£9,935 ICER (base-case MAIC): **£48,650** ICER (PSWA analysis): **£40,258**
		TEW from 4 years	Change from original base-case ICER: +£8,987 ICER (base-case MAIC): **£47,702** ICER (PSWA analysis): **£39,343**
4. Time-to-death utilities (no IV decrement), RDI equalised while varying TEW		TEW from 2 years	Change from original base-case ICER: +£9,387 ICER (base-case MAIC): **£48,102** ICER (PSWA analysis): **£39,528**
		TEW from 3 years	Change from original base-case ICER: +£7,697 ICER (base-case MAIC): **£46,409** ICER (PSWA analysis): **£37,992**
		TEW from 4 years	Change from original base-case ICER: +£6,707 ICER (base-case MAIC): **£45,422** ICER (PSWA analysis): **£37,059**

NICE National Institute for Health and Care Excellence

Appendix A: summary of clinical SLR

Neither SLR (RCTs and single-arm trials) identified trials that have a PDC arm in KRASm patients. Please see below the summary of interventions and comparators assessed as well as the inclusion criteria for patients. The Majority of identified RCTs required patients to have prior treatment with platinum-based chemotherapy. Several single-arm trials also specified that patients should have had prior treatment with platinum-containing chemotherapy regimen(s), supporting the argument that most patients have had previous PDC.

1) RCTs

Studies reporting on an exclusively KRASm study population (n=7)

- All but one of the studies assessed kinase inhibitors. Rulli et al 2015 compared docetaxel monotherapy to erlotinib monotherapy.
- Rulli et al 2015 and Carter et al 2016 required failure following a platinum-containing doublet chemotherapy regimen
  - Rulli et al 2015 reported 90% of patients had received first-line platinum-based therapy with the remaining patients having received chemotherapy in the adjuvant therapy
  - Carter et al 2016 reported that 40–55% and 45–60% of patients had received one or two prior regimens, respectively (none had more than two).

Studies reporting data for a KRASm subgroup

Chemotherapy as a comparator (n=6)

- Treatments assessed were nivolumab, atezolizumab, ganetespib, pelareorep and erlotinib, administered either as monotherapy or in combination with chemotherapy. Docetaxel or pemetrexed monotherapy was the control treatment in all six studies.
- The study by Ramalingam et al 2015 only required patients to have progressed following first-line therapy, while the remaining studies all specified prior treatment with platinum-based chemotherapy doublets (except for Bradbury et al 2018 where doublet therapy was not required in patients aged >70 years).

Erlotinib as a comparator (n=3)

- Spigel et al 2017 assessed onartuzumab (a MET-inhibiting humanized monoclonal antibody) in combination with erlotinib; Scagliotti et al 2015 assessed tivantinib in combination with erlotinib and Karampeazis et al 2013 assessed pemetrexed monotherapy.

- All three studies required one or two prior lines of platinum doublet chemotherapy (except in patients ≥65 years in Karampeazis et al 2013).

Other interventions (n=1)

- Ciuleanu et al 2017 study assessed linsitinib (a dual IGF-1R and IR inhibitor) given as maintenance therapy in conjunction with erlotinib (versus placebo + erlotinib) in patients with stages IIIB or IV NSCLC (with ECOG performance status 0–1) and stable disease or better following four cycles of first-line platinum-based chemotherapy.

2) Single arm trials

- 4 broadly defined drug classes were assessed across identified trials: inhibitors of KRASG12C, inhibitors of the EGFR/MAPK signaling pathway, immune checkpoint inhibitors, and modulators of protein stability.
- Sotorasib trial (CodeBreak 100)
- Antroquinonol trial (NCT02047344) included patients with disease progression after 2 prior LoTs (at least 1 platinum-based) or patients who refused treatment with approved treatments
- Gerber et al. 2020 (assessing Defactinib), NCT026420 (assessing Docetaxel and trametinib) and NCT02258607 (assessing Momelotinib and trametinib) required patients to have at least 1 prior platinum-based CHT
- Gulley et al. 2017 (assessing Avelumab) included patients who had progression after platinum-based doublet CHT for metastatic disease
- Pujol et al. 2020 (assessing Abemaciclib and pembrolizumab) included CHT-naïve patients with ≥ 1% TC PD-L1 staining in Cohort A and patients with ≤ 1 prior platinum-containing CHT regimen in Cohort B

NICE National Institute for Health and Care Excellence

Appendix B: Oncology Dynamics RWE survey to validate NSCLC pathway in UK

Overall Research Design

The aim of this analysis was to ascertain for the period of the last year what proportion of patients with locally advanced and/or metastatic NSCLC receiving 2nd Line docetaxel have had previous immunotherapies or platinum-based chemotherapy (monotherapy or in combination).

This was a retrospective analysis using readily available data from Oncology Dynamics TM (IQIVIA Ltd., London, UK). The inclusion criteria were as follows: adult patients aged ≥18 years diagnosed with advanced/metastatic NSCLC (and the subset who were confirmed EGFR/ALK/ROS-1 negative) and who received treatment between Q3 2020 to Q2 2021 in Oncology Dynamics.

Data Source

Oncology Dynamics is a physician-based cross-sectional survey that collects anonymised patient level data information on drug-treated cancer patients from ten countries including France, Germany, Italy, Spain and the UK in Europe. Data are collected on cancer-diagnosed patients regardless of cancer type, stage and/or treatment modality. It is designed as repeated quarterly cross-sectional cohorts, contains more than 167,000 cancer cases per year and over 35 cancer indications and covers demographic, diagnostic and treatment information recorded at time of diagnosis, as well as treatment information at 'current' and 'previous' line of therapy (as defined by the treating physician) (Kafatos et al, 2021a). The sampling plan for the database is set up in order to be representative of physician specialties treating cancer in each country with a sample size that adequately determines patterns of cancer management. The database therefore provides a representative sample of clinical practice among currently treated cancer populations. The validity of these data has been previously demonstrated for a wide range of oncology indications (Kafatos et al, 2021b; Maroun et al, 2018; Marchetti et al, 2017; Canta et al, 2016; Schmidt et al, 2014; Zhao et al, 2012; Inoue et al, 2009).

NICE National Institute for Health and Care Excellence

Results

Table 1 summarises that there was only 1 patient with advanced NSCLC in the UK that received docetaxel in 2L treatment. This patient received platinum-based chemotherapy with an anti-PD-1/L1i in 1L. Table 2 summarises 12 metastatic NSCLC patients receiving docetaxel in 2L in the UK. The majority 83% (n=10/12) had a PD-L1 expression of 1-49%. Nearly all patients (92%; n=11/12) received platinum-based chemotherapy with anti-PD-1/L1i in 1L before receiving docetaxel in second line.

Table 1: Prior systemic anti-cancer therapy received in 1L before 2L treatment with docetaxel in **Advanced NSCLC** patients in UK

	Advanced NSCLC							
	Overall				**EGFR/ALK/ROS-1 negative**			
	Overall		PD-L1 1-49%		Overall		PD-L1 1-49%	
	n	%	n	%	n	%	n	%
United Kingdom								
***N,* no. of patients in 2L receiving docetaxel**	**1**		**1**		**1**		**1**	
Types of prior systemic anticancer therapy in 1L								
Any anti-PD-1/L1i (monotherapy or combination)	1	100%	1	100%	1	100%	1	100%
anti-PD-1/L1i monotherapy								
Platinum-based chemotherapy WITH anti-PD-1/L1i	1	100%	1	100%	1	100%	1	100%
Platinum-based chemotherapy WITHOUT anti-PD-1/L1i								
Other								

Table 2: Prior systemic anti-cancer therapy received in 1L before 2L treatment with docetaxel in **Metastatic NSCLC** patients in UK

NICE National Institute for Health and Care Excellence

Metastatic NSCLC

	Overall						EGFR/ALK/ROS-1 negative					
	Overall		PD-L1 1-49%		PD-L1 >=50%		Overall		PD-L1 1-49%		PD-L1 >=50%	
	n	%	n	%	n	%	n	%	n	%	n	%
United Kingdom												
***N,* no. of patients in 2L receiving docetaxel**	**12**		**10**		**2**		**7**		**6**		**1**	
Types of prior systemic anticancer therapy in 1L												
Any anti-PD-1/L1i (monotherapy or combination)	11	92%	9	90%	2	100%	7	100%	6	100%	1	100%
anti-PD-1/L1i monotherapy												
Platinum-based chemotherapy WITH anti-PD-1/L1i	11	92%	9	90%	2	100%	7	100%	6	100%	1	100%
Platinum-based chemotherapy WITHOUT anti-PD-1/L1i	1	8%	1	10%								
Other												

References

Canta FB, Mitrofan L, Karanis Y, Mistry H, Anger CJaOO. 17P deletion and TP53 gene mutation (17P/TP53) testing behaviour and treatment patterns for chronic lymphocytic leukemia (CLL) patients in France, Germany, Italy, Spain and UK (EU5). 2016;27:vi317.

Inoue Y, Toiyama Y, Tanaka K, Miki C, Kusunoki MJJJOCO. A comprehensive comparative study on the characteristics of colorectal cancer chemotherapy. 2009;39:367-375.

Kafatos G, Banks V, Burdon P, et al. Biomarker testing and mutation prevalence in metastatic colorectal cancer patients in five European countries using a large oncology database. 2021a;17:1483-1494.

Kafatos G, Banks V, Burdon P, et al. Impact of biomarkers and primary tumor location on the metastatic colorectal cancer first-line treatment landscape in five European countries. 2021b;17:1495-1505.

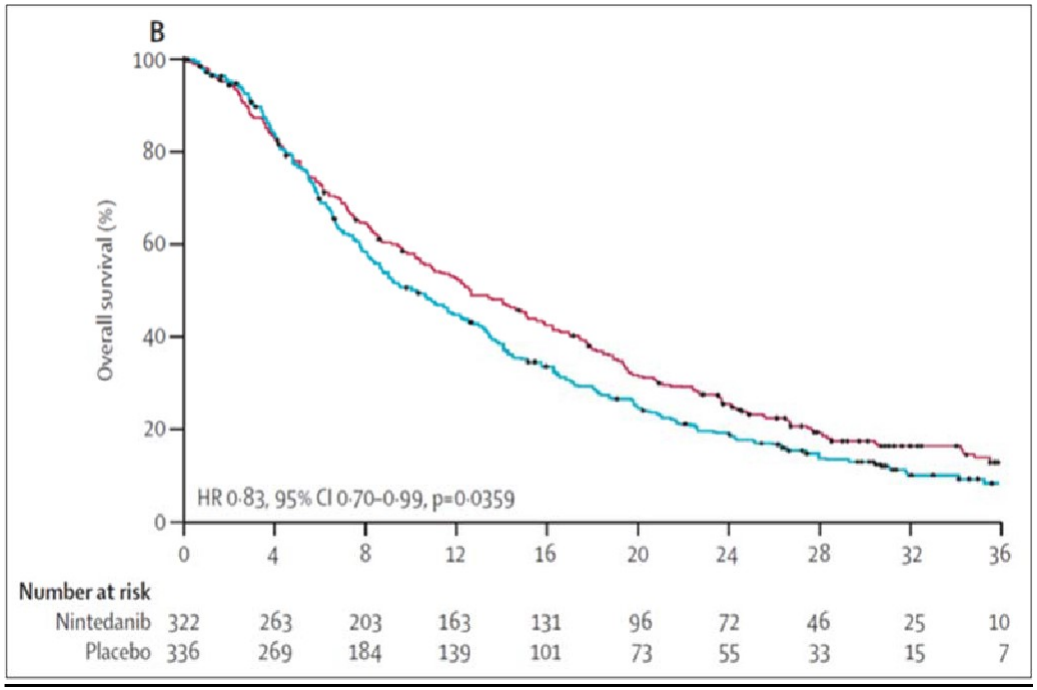
Marchetti P, Maass N, Gligorov J, et al. Patient database analysis of fulvestrant 500 mg in the treatment of metastatic breast cancer: A European perspective. 2017;32:247-255.

Maroun R, Mitrofan L, Benjamin L, et al. Real life patterns of care and progression free survival in metastatic renal cell carcinoma patients: retrospective analysis of cross-sectional data. 2018;18:1-7.

NICE National Institute for Health and Care Excellence

Schmidt N, Kostev K, Jockwig A, Kyvernitakis I, Albert U-S, Hadji PJIJCPT. Treatment persistence evaluation of tamoxifen and aromatase inhibitors in breast cancer patients in early and late stage disease. 2014;52:933-939.
Zhao Z, Pelletier E, Barber B, et al. Major surgery in patients with metastatic colorectal cancer in Western Europe. 2012;43:456-461.

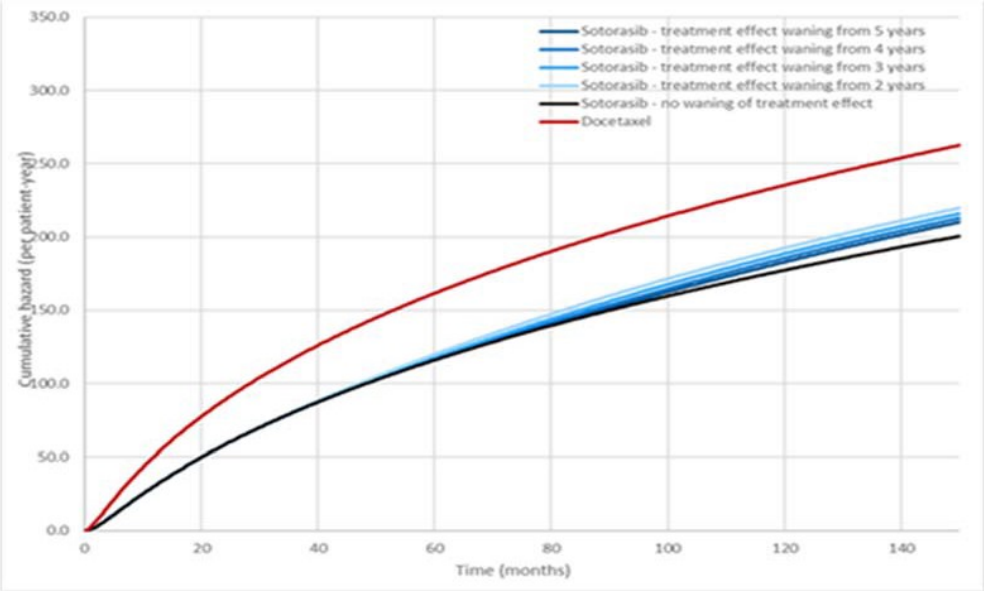
Appendix C: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line)



NICE National Institute for Health and Care Excellence

Appendix D: TEW scenario graphics (OS cumulative hazard plots)

The cumulative hazard plot (per patient-year) for the following scenarios are presented below: docetaxel base-case OS, Sotorasib base-case OS (no TEW) and gradual 5-year TEW applied to the Sotorasib OS curve from 2, 3, 4 and 5 years. The hazard plot is as expected, with the cumulative hazards of death becoming more and more equal depending on when waning begins.



NICE National Institute for Health and Care Excellence

Appendix E: results for minor comparator nintedanib+docetaxel

The table below reflects the updated confidential PAS for Sotorasib and explores the spread of deterministic ICERs (vs. nintedanib+docetaxel). Despite unresolved issues that impact this comparison (TEW, utilities and RDI), for brevity the ERG preferred base-case assumptions are set with only variation related to issue 8. **It should be noted that the probabilistic ICER for this minor comparator (vs. nintedanib+docetaxel) is consistently lower by around £1800 and this should be considered in decision making.**

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case ICER
ERG preferred settings: health state utilities with no PFS differential, RDI assumed equalised, TEW from 2 years	Original submitted base-case used time-to-death method for utilities (with IV decrement), RDI from trials (i.e. not equalised) and no TEW. In addition, a piecewise approach with three fitted HRs to obtain a treatment effect for add-on nintedanib (vs. docetaxel).	ERG base-case: three period piecewise method with first period HR=1	Change from original base-case ICER: +£8,514 ICER (base-case MAIC): **£42,142** ICER (PSWA analysis): **£22,997**
		Two period piecewise method with first period HR=1 (New fitted HR for period 6m+ is 0.7904)	Change from original base-case ICER: +£9,004 ICER (base-case MAIC): **£42,632** ICER (PSWA analysis): **£23,150**
		Two period piecewise method (all fitted)	Change from original base-case ICER: +£668 ICER (base-case MAIC): **£32,960** ICER (PSWA analysis): **£19,267**

I only have responses to 2 issues raised by NICE:

Issue 3 High number of serious adverse events
This is typical of trials in patients advanced NSCLC because these parents have multiple disease-related symptoms and complications. It is important to distinguish these from SAEs that are recorded by investigators as treatment-related

Issue 6 RCT pragmatism and trial norm
This issue relates to exclusion of PS2 patients from CODEBREAK100. This is usual in Phase I trials because of the safety concerns with first-in-human use, and the physical demands of very regular clinic review and additional investigations associated with such a protocol

Warm regards

[redacted]

NICE National Institute for Health and Care Excellence

Technical engagement response form

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments by **5pm** on **22 October 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. If you would like to comment on issues in the ERG report 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 appraisal, 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.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. 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.

- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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 replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

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.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Boehringer Ingelheim**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Population narrower than NICE scope	**No**	The population in scope for the NICE submission is adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC which would include both squamous and non-squamous histology patients. However data from CodeBreak 100 are primarily representative of the non-squamous histology with adenocarcinoma making up the vast majority of enrolled patients (95.2%) and only 1 patient (0.8%) with squamous disease, so assertion of benefit in the squamous population and indeed non-adenocarcinoma (n=5 patients only) is not supported or justified. This also applies to patients with prior TKI which are included in the NICE scope but represented only 7.1% of the entire CodeBreak 100 population.
Key issue 2: Generalisability / lack of UK participants	**No**	We agree with the assessment provided in the ERG report
Key issue 3: High risk of bias of CodeBreaK100	**Yes**	As a result of the significant change in standard of care for NSCLC patients over the past few years and the dominance of immune-oncology agents +/- chemo in the first line setting, the patient population tested in LUME-Lung 1 is different, at least in part, by prior therapy to the CodeBreak100 population. This significant difference limits the utility and increases the bias of the indirect cross-trial comparisons on which the economic and clinical benefit modelling was built (Section B.2.9.) as also recognised by Amgen (B 2.9.5. and elsewhere). No comparison has been made to the recent real world data study VARGADO with

		similar patient numbers demonstrating real world outcomes for the use of Nintedanib and Docetaxel with similar treatment history to those in CodeBreak100. VARGADO is being conducted in a patient population with prior chemo and immune checkpoint inhibitor treatment (81% of CodeBreak 100 patients received prior chemo + immune checkpoint inhibitor). In VARGADO, Cohort B (3rd line post chemo + immune checkpoint inhibitor) ORR to Nintedanib + Docetaxel was 50% compared to 37.1% for Sotorasib in CodeBreak100. In Cohort B (3rd line setting after chemotherapy and achieved mPFS 6.5 mo (n=47), mOS 12.4 mo (n=55, ESMO 2020), which is comparable to mPFS seen with sotorasib in CodeBreak 100 (6.8 mo, OS 12.5 mo). Furthermore, data from patients who received the combination in the 2nd line setting after 1st line chemo + immune checkpoint inhibitor (Cohort C) confirm activity in this setting with mPFS 4.77 months, ORR 37.5% OS is not mature (ESMO 2021)
Key issue 4: High number of serious adverse events observed in CodeBreaK100	**Yes**	Although Amgen claims that "*safety appears to be superior to that seen with docetaxel or nintedanib plus docetaxel*" (p. 28), there were 15.9% (20 patients) fatal AEs; no specification of whether any of these AEs was related to disease progression is provided. In comparison, no fatal AEs were reported in the real world study of nintedanib + docetaxel in second line or greater NSCLC VARGADO (post chemo and checkpoint inhibitor) Cohort B or C and fatal AEs possibly unrelated to disease progression in LUME-lung 1 were 5.4% for the nintedanib + docetaxel group and 3.8% for the docetaxel group (Reck et al, The Lancet Oncology, 2014), both of these rates are significantly lower than seen with sotorasib in CodeBreak 100
Key issue 5: Validity of ITC without a common comparator	**No**	Modelling comparator was Docetaxel alone, whereas we believe current standard of care comes from Docetaxel + Nintedanib in combination. The clinical data in LUME Lung 1 demonstrates clear survival benefit of docetaxel plus Nintedanib

		with mOS 12.6mths vs 10.3 mths with docetaxel alone in patients with adenocarcinoma of the lung (HR0.83, p=0.0359).
Key issue 6: Partitioned Survival Model structure not validated or justified	**No**	We agree with the assessment provided in the ERG report
Key issue 7: Exclusion of platinum-based chemotherapy as a comparator in 2nd line	**No**	We agree with the assessment provided in the ERG report
Key issue 8: Docetaxel plus nintedanib modelling approach leading to worse survival	**Yes**	In the clinical effectiveness section, median PFS and OS estimate for sotorasib as indicated in the primary comparison analysis (B.2.9.4.1) are different to those reported for this drug in the secondary comparison analysis (B.2.9.4.2): mPFS 6.3 mo vs 9.2 mo and mOS >12.5 mo vs 23.5 mo. Amgen's analysis suggests that the gain in OS is more significant against nintedanib + docetaxel (6.5 mo) compared to the gain against docetaxel alone (4.6 mo) implying addition of nintedanib worsens survival. Indeed, on p.201 the model concludes that "*adding nintedanib lowers OS at 1 year relative to docetaxel but increases it later on*"; this is at odds with clinical evidence and puts into doubt the validity of the indirect comparisons on which the modelling was based. The LUME Lung 1 data showing survival benefit are summarised above as are the VARGADO real world evidence showing patient outcomes with Nintedanib plus Docetaxel.
Key issue 9: No waning of treatment effect	**No**	We agree with the assessment provided in the ERG report
Key issue 10: TTD modelling approach inconsistent with OS and PFS modelling	**No**	We agree with the assessment provided in the ERG report
Key issue 11: Time-to-death utilities do not seem well-informed	**No**	We agree with the assessment provided in the ERG report

Key issue 12: Disutility for IV administration not well justified	**No**	We agree with the assessment provided in the ERG report
Key issue 13: Relative dose intensity and wastage assumption not justified	**No**	We agree with the assessment provided in the ERG report

Additional issues

Please use the table below to respond to additional issues in the ERG report 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 appraisal (e.g. at the clarification stage).

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response

Additional issue 1: Indirect comparisons do not take into account real-world evidence showing PFS and OS benefit for nintedanib + docetaxel in a more relevant patient population	Section 3.3, 3.4, 4.2.6	**Yes**	Although Amgen has included in its economic modelling the Flatiron Health real-world evidence for docetaxel, it has failed to take into consideration real-world data now available for nintedanib + docetaxel from the ongoing VARGADO study. VARGADO is being conducted in a patient population with prior chemo and immune checkpoint inhibitor treatment which is a more relevant comparator to the CodeBreak100 population (81% of CodeBreak 100 patients received prior chemo + immune checkpoint inhibitor). In VARGADO, nintedanib + docetaxel used post platinum-chemo and immune checkpoint inhibitor (3rd line setting, Cohort B) achieved ORR of 50%, mPFS 6.5 mo (n=47), and mOS 12.4 mo (n=55, ESMO 2020), which is supercedes ORR seen with Sotorasib (37.1%) and comparable to mPFS seen with sotorasib in CodeBreak 100 (6.8 mo, OS 12.5 mo,). Furthermore, data from patients who received the combination in the 2nd line setting after 1st line chemo + immune checkpoint inhibitor (Cohort C) confirm activity in this setting with ORR 37.5% , mPFS 4.77 months, OS is not mature (ESMO 2021)

Additional issue 2: Indirect comparisons do not take into account real-world evidence for PFS benefit shown by nintedanib + docetaxel in *KRAS*-mutant patients	Section 3.3, 3.4, 4.2.6	**Yes**	Nintedanib + docetaxel has delivered similar efficacy to sotorasib with mPFS 4.77 in KRAS mutant patients (mutation sub-type not known) in the VARGADO study (Cohort C), similar to the mPFS seen in the overall VARGADO Cohort C population. Amgen's cost effectiveness modelling should be revised to take into account these recent data since they stated in their submission (p.12) that "*there are no data for docetaxel in combination with nintedanib specifically in KRAS-mutated NSCLC patients*"
Additional issue 3: Clinical efficacy comparisons do not take into account real-world evidence for ORR benefit shown by nintedanib + docetaxel	Section 3.3, 3.4, 4.2.6	**Yes**	ORR considered to be a key parameter for efficacy in Amgen's submission (p.28, p.38) (37.1% in second or subsequent line therapy) is comparable to ORR seen with docetaxel + nintedanib in clinical practice based on real-world evidence from VARGADO Cohort B (ORR 50%) and VARGADO Cohort C (ORR 37.5%) – note substantial ORR uplift in real clinical practice (in a patient population relevant to that investigated in CodeBreak 100) compared to registrational LUME-Lung 1 study on which Amgen's modelling was built (37.5% and 50%% vs 4.7%).

Additional issue 4: Subsequent therapy options not reflected in cost-effectiveness estimate	4.2.9.4	**Yes**	In estimating the costs of subsequent therapy and the effect of that on incremental cost effectiveness ratio (ICER), Amgen affirms that "*there is no evidence either based on clinical opinion or trial data* [CodeBreak 100, n=44] *that suggests patients will receive nintedanib+docetaxel subsequent to sotorasib*", p.208. While the comment relating to clinical trial evidence is consistent with CodeBreak 100 not including any UK sites, the clinical opinion presented is not reflective of current UK clinical practice. [redacted] Not taking the widespread use of nintedanib + docetaxel into account suggests that the cost effectiveness presented by Amgen has been overestimated

Additional issue 5: Biased sotorasib utility modelling	4.2.8	**Yes**	High variability in sotorasib's PK and its short half-life of 5.5h (Hong et al, NEJM 2020) mandate the drug (8 tablets) to be taken at the same time every 24h for effective target inhibition. This places an additional burden on the patient which is not built into the utility modelling
Additional issue 6: Proposed dose is not demonstrated to be the optimal dose	4.2.9	**Yes**	The proposed sotorasib dose of 960 mg per day is consistent with the dosing regimen in CodeBreaK100, however this has not been shown to be the optimal dose for optimal benefit/risk profile putting into question the possible efficacy that the compound would achieve if the dose was significantly reduced. Indeed, Amgen has an ongoing study to evaluate the lower dose of 240 mg QD vs 960 mg QD (FDA post-marketing requirement) but this is not discussed or disclosed in the current submission

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case ICER

Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER
..	..	..	[INSERT / DELETE ROWS AS REQUIRED]
Company's preferred base case following technical engagement	**Incremental QALYs: [QQQ]**	**Incremental costs: [£££]**	**Please provide the revised company base-case ICER resulting from combining the changes described, and the change from the company's original base-case ICER**

NICE National Institute for Health and Care Excellence

Technical engagement response form

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments by **5pm** on **22 October 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's summary of key issues, critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- Please ensure your response clearly identifies the issue numbers that have been used in the executive summary of the ERG report. If you would like to comment on issues in the ERG report 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 appraisal, 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.
- Do not use abbreviations.

NICE National Institute for Health and Care Excellence

- Do not include attachments such as journal articles, letters or leaflets. 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.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- 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 replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

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.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Amgen Limited**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response	ERG comment
Key issue 1: Population narrower than NICE scope	NO	As agreed during the technical engagement call, the initial NICE scope population can sometimes differ from the final license population, but the final NICE recommendation cannot be wider than the license indication. **This issue is considered resolved following the technical engagement call.**	No further comment. ERG report highlighted this issue for the committee.
Key issue 2: Generalisability / lack of UK participants	NO	As agreed in the technical engagement call it is not unusual to have small or no numbers of UK patients in trials of targeted NSCLC treatments. The CodeBreak 100 trial is broadly generalisable to the relevant population in this appraisal. Amgen has consulted clinicians, who practice NSCLC in large centres in England, and they have	No further comment. ERG report highlighted this issue for the committee.

NICE National Institute for Health and Care Excellence

		reported that the population demographics in CodeBreak 100, including ethnicity, is largely representative of patients treated in their clinics. This issue is considered resolved following the technical engagement call.	
Key issue 3: High risk of bias of CodeBreaK100	NO	As agreed in the technical engagement call, the risk of bias associated with CodeBreak100 is broadly in line with other pivotal 1-arm trials in NSCLC that have been the basis of previous NICE appraisals. Issues related to concealment/blinding and confounding are inherent in 1-arm trials of this nature and hence the need for statistical methods such as MAIC and PSWA. **This issue is considered resolved following the technical engagement call.**	No further comment. ERG report highlighted this issue for the committee.
Key issue 4: High number of serious adverse events observed in CodeBreaK100	NO	A serious adverse event in CodeBreak100 was defined as an adverse event that meets at least 1 of the following criteria: fatal, life threatening (places the subject at immediate risk of death), requires in patient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, congenital	No further comment. ERG report highlighted this issue for the committee.

		anomaly/birth defect, or other medically important serious event. As agreed in the technical engagement call it is important to note that serious adverse events do not necessarily need to be related to treatment, and hence treatment-related adverse events are considered more relevant here. Serious adverse reactions occurred in 50% of patients treated with sotorasib. Serious adverse reactions in ≥ 2% of patients were pneumonia (8%), hepatotoxicity (3.4%), and diarrhoea (2%). Fatal adverse reactions occurred in 3.4% of patients who received sotorasib due to respiratory failure (0.8%), pneumonitis (0.4%), cardiac arrest (0.4%), cardiac failure (0.4%), gastric ulcer (0.4%), and pneumonia (0.4%). Treatment-related adverse events are those adverse events of any grade that were considered by the investigators to be related to treatment. In CodeBreak 100 a total of 88 patients (69.8%) reported adverse events of any grade that were considered by the investigators to be	

		related to treatment. (treatment-related adverse events). The worst grade of treatment-related adverse event was grade 3 in 25 patients (19.8%) and grade 4 in 1 patient (0.8%; pneumonitis and dyspnoea); no treatment-related adverse events of grade 5 (deaths) were reported. The most frequent treatment-related adverse events were diarrhoea (in 40 patients [31.7%]), nausea (in 24 [19.0%]), increase in the alanine aminotransferase level (in 19 [15.1%]), increase in the aspartate aminotransferase level (in 19 [15.1%]), and fatigue (in 14 [11.1%]). This issue is considered resolved following the technical engagement call.	
Key issue 5: Validity of ITC without a common comparator	NO	Base-case MAIC for comparison with docetaxel (using SELECT-1 RCT) Amgen conducted a MAIC that weighted CodeBreak100 patients based on BL characteristics to match the docetaxel arm of the SELECT-1 trial. The variables for inclusion were selected based on literature review and 6 individual interviews with	As recommended in Key issue 5, the company stated that they performed methods other than IPW for the PSWA, but no results have been presented. The company have also not redone the PSWA limiting to the docetaxel only data of the Flatiron study. The ERG does accept that weighting the MAIC by mutation status is infeasible, although selection of KRAS G12C

		clinicians and the resulting list is in line with previously presented MAICs in NSCLC. As agreed in the technical engagement call, there was exclusion of some variables from the MAIC which could potentially be treatment effect modifiers (brain metastases, KRAS G12C mutation status). However, as agreed given the data available it was not possible to include these in the MAIC: • SELECT-1 did not report the proportion of (inactive) brain metastases at baseline (trial reports, publications and appendices, HTA submissions) and so it was not possible to include as a matching variable in the MAIC. ○ All trials exclude *active* brain metastases, which clinicians suggest is more likely to be a modifier. ○ The proportion of patients with brain metastases was higher	mutation SELECT-1 data could potentially have been performed.

NICE National Institute for Health and Care Excellence

		in CodeBreaK100 (21%) than in LUME-Lung 1 (8%). If SELECT-1 had a similar proportion of inactive brain metastases to CodeBreak100, any bias would favour the docetaxel arm and make cost-effectiveness results conservative. • It is not feasible to match on KRAS G12C in the base-case MAIC. 100% of CodeBreak100 are KRAS G12C mutation positive (42% in SELECT-1 with the remaining having KRAS mutations other than G12C) and so a MAIC would "weight away" the sample of CodeBreak100. Supplementary analysis – propensity score weighting analysis (PSWA) using Amgen Flatiron RWE database This analysis aimed to compare OS and PFS with sotorasib against standard of care chemotherapy observed in a cohort of patients with previously treated KRAS-	

		mutated advanced/metastatic NSCLC in the Flatiron database. The PSWA used a flat-iron RWE dataset that is a basket of chemotherapies, of which around 40% received singlet docetaxel or a doublet containing docetaxel. Some of the concerns raised by the ERG were clarified in the technical engagement call: • It is not the case that only 4th line patients were included in the PSWA analysis dataset. Patients were selected based on their last line of treatment (2nd, 3rd, 4th line) and if data was available in the 5th line or later only the patient's data up to the 4th line was selected (i.e. broadly in line with inclusion criteria of CodeBreak100). • The ERG suggested exploring alternative methods for calculating a treatment effect from PSWA: ATE (average treatment effect) instead of the presented ATT (average effect of the treatment on the treated). Taking the ERGs advice Amgen can report that switching to the former made little	

		difference in relative effectiveness and so little difference to the PSWA scenario analysis ICER. Comparison of base-case MAIC and Flatiron PSWA Amgen agrees with the ERG assessment that it is difficult to assess which analysis is more robust (or less biased). There are some subtle trade-offs that make it unclear and may even favour the Flatiron PSWA: • Both data sources are fundamentally observational (i.e. unrandomised): a single arm of a controlled trial (SELECT-1) vs. an uncontrolled (but larger sample) historical control cohort (Flatiron). • Neither source has any UK patients: CodeBreak100 is a multinational Ph2 trial and Flatiron is based on an American health record database. • The MAIC analysis must weight CodeBreak100 patients to match SELECT-1 and the assumption is made that this treatment effect translates to the CodeBreak100 population, whereas the PSWA	

		weights the Flatiron data to match the CodeBreak100 population. • In the real-world disease progression is derived from physician notes in a clinical practice setting and may be informed by RECIST criteria in conjunction with other signs of progression. • Given data availability and richness of Flatiron data, the PSWA allowed consideration of and final inclusion of more weighting covariates and is therefore more heavily adjusted. o The same clinician elicitation exercise as for the MAIC informed selection of “very important” and “somewhat important” covariates but now the data available allowed all “somewhat important” to be included in the covariate selection algorithm (i.e. could potentially be included in the final model). o Therefore, the final analysis included several covariates	

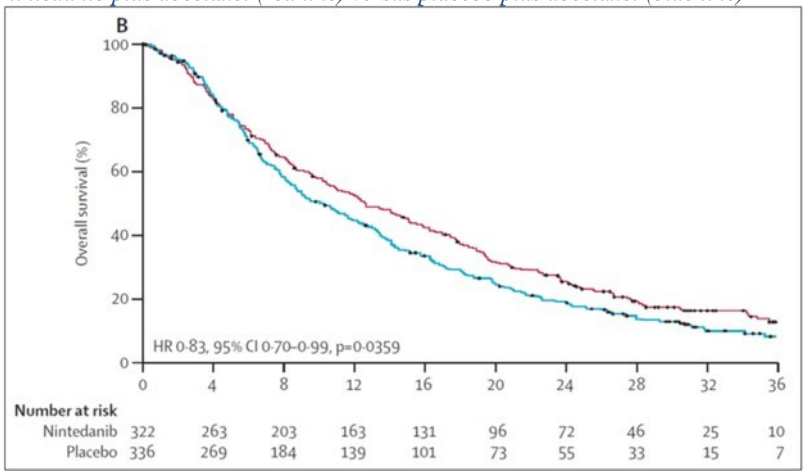
NICE National Institute for Health and Care Excellence

		not in the MAIC analysis (brain metastases, presence of non-KRAS mutations, prior lines of treatment, type of prior treatments, Albumin). **For these reasons, the PSWA is also presented alongside the base-case MAIC below for the committee's consideration.**	
Key issue 6: Partitioned Survival Model structure not validated or justified	NO	As agreed in the technical engagement call, it is not always feasible (and not conventional) to present two fundamentally different modelling methodologies for validation. As argued previously, the fundamental problems of the Partitioned Survival Model are unlikely to be resolved by a State Transition Model and the data requirements of such a model are harder to meet. **This issue is considered resolved following the technical engagement call.**	The ERG agrees that to ask for additional STM modelling may not be strictly necessary, but emphasizes that an STM could have contributed to verifying plausibility of extrapolations and exploring clinical uncertainties, reducing structural uncertainty around model results.
Key issue 7: Exclusion of platinum-based chemotherapy	YES	There is a reasonably broad consensus among consulted clinicians and NHSE that the optimal initial therapy for patients in NSCLC without current actionable	The ERG was informed by the NHSE clinical expert on this issue, and so it remains a matter of judgment.

as a comparator in 2nd line		mutations is anti-PD-1/PD-L1 immunotherapy (IO) in combination with platinum-based chemotherapy (i.e. platinum doublet chemotherapy or “PDC”). This is also in-line with the NICE pathway and historical NICE recommendations. The group of patients who are eligible for Sotorasib that will have received both an IO and PDC previously is growing (and those that have only received an IO shrinking). Therefore, for most patients Sotorasib will displace docetaxel in the pathway. Nevertheless, if PDC were considered a minor comparator an unanchored MAIC would not be possible. As agreed in the technical engagement call, Amgen can confirm that the SLR did not identify any KRAS population trials with a PDC arm (see appendix A below). The results of the SLR also support the proposition that patients in this population have been pre-treated with PDC (as does the baseline characteristics of the pivotal trial CodeBreak100 with 90% of patients having been pre-treated with PDC).	

		As additional evidence, in Appendix B a retrospective analysis using data from Oncology Dynamics ™ confirms that most patients who received docetaxel recently in the UK are likely to have received IO and PDC previously. The PSWA could be considered a reasonable proxy for a comparison with PDC, given that the most common regimen in the chemotherapy basket was platinum-based chemotherapy (but still under 1/3 of patients). This issue is considered resolved following the technical engagement call.	
Key issue 8: Docetaxel plus nintedanib modelling approach leading to worse survival	NO	A piecewise approach to generating a nintedanib treatment effect (vs. docetaxel alone) by fitting HRs to 3 periods was undertaken (via Cox PH models) in the base-case model. This is because curve diagnostics suggested that the LUME-Lung 1 OS curves did not satisfy the proportional hazards assumption and so a single fitted	The point the ERG was making in this issue was not only that the 26 month landmark was not clearly identifiable, but also that the resulting curves did not seem to reflect trial results.

NICE National Institute for Health and Care Excellence

		HR was deemed inappropriate (KM curves shown in Appendix C). Amgen agree with the ERG that it is less clear that the proportional hazards assumption is violated at 26 months (compared with 6 months) and so a scenario with piecewise HRs for only 2 periods is worth exploring (i.e. 0-6 and 6+ months). The ERG proposes setting the HR in the first period (0-6m) to 1 and so assuming equal survival between docetaxel and add-on nintedanib. However, Amgen find it highly irregular to invalidate measured trial data from a published 2-arm phase 3 trial. Trial data points are usually considered more valid than intuitive assumptions in the hierarchy of evidence. Sometimes the impact of a treatment is nuanced and Kaplan-Meier (KM) curves can cross, but this should be reflected in any fitted treatment effects. This is particularly so when the crossing occurs with many patients at risk (i.e. the sample size is higher and the curves more reliable at the beginning of a KM).	See the figures below to illustrate this (taken from ERG report, page 76, figure 4.2 and 4.3). *Figure 0.1: Modelled OS curves taken from the economic model*  *Figure 0.2: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line)* 

		This issue only impacts the minor comparison with nintedanib plus docetaxel. This issue is considered unresolved following the technical engagement call and the ICER ranges presented below will reflect this.	The Kaplan Meier shows a very slight benefit of docetaxel compared to docetaxel plus nintedanib in the first 4 months (figure 4.3), which is then transformed into a >1 yr survival benefit for docetaxel in the modelled OS curves (figure 4.2, please note time scale is deviant from figure 4.3), and so the fitted curves do not seem to reflect the K-M data. No expert opinion or any other validation was provided to justify this. The company nevertheless used the fitted curves to inform the comparison between sotorasib versus docetaxel + nintedanib. The ERG does believe that trial results should be used at all times, but in this case setting the HR to 1 would not invalidate trial results more than the seemingly implausible curves the company used.
Key issue 9: No waning of treatment effect	NO	Amgen believe that although treatment effect waning (TEW) can be useful to explore model sensitivities, it is a relatively blunt tool and should be applied considering the particulars of each case (as requested Appendix D below presents the hazard plot for context). As argued previously, there are several reasons why TEW should be limited and applied carefully in this case: • To a large extent, the impact of discontinuation on OS and PFS has already been “baked” into the hazard function (and so projected	The ERG re-iterates that data are immature and any assumptions on a sustained treatment effect are highly uncertain. Starting treatment waning at the 2 yr time point and having it decrease over a 5-yr period (as in the ERG base-case) could already be considered rather optimistic given available evidence. No new evidence was brought to the table and so the ERG did not change their preferred assumptions.

NICE National Institute for Health and Care Excellence

		survival estimates) because within the trial period (in which parametric curves have been fitted) a significant number of patients have discontinued over the period (>80%). • TEW is more intuitive and more easily defendable when the two treatments are comparable and have a similar mechanism of action and so a reasonable assumption can be made that the relationship between being on treatment and benefiting longer term are similar (e.g. two EGFR targeting TKI therapies). However, sotorasib and docetaxel are very different medicines with different actions and therefore such an assumption is more uncertain. • A case can be made that the mix of patient at the point of the March data cut (around 15 months of follow-up) in the sotorasib arm is in a better average "health state" than the docetaxel patients and so the hazards of survival will continue to	

NICE National Institute for Health and Care Excellence

		be better in the former for some time. ○ Before the extrapolated portion (i.e. within trial period), 1/2 of Sotorasib patients have yet to progress, but for docetaxel it is only <1/6. ▪ In the docetaxel arm of the model at around 15 months (i.e. the trial period of SELECT-1 and not an extrapolated portion), of the <30% alive only 4% points of patients are progression free (the remaining progressed). In contrast, by this point around half of the patients in Codebreak100 who are alive (i.e. around 40%) have yet to progress (20% points of patients).	

NICE National Institute for Health and Care Excellence

		• According to Appendix E of the company submission and the related publication, by the time of the March 2021 data cut-off around 80% (81.7%) of patients have discontinued treatment, around 40% remain alive and around 20% have yet to progress. Therefore, half of patients who are alive will have remained on sotorasib treatment at this point. o Sotorasib is given in CodeBreak100 until progression or the development of unacceptable AEs and so it is inappropriate to apply TEW early when a significant proportion of those alive are still benefiting from treatment. o Applying TEW in too blunt a fashion would bias cost-effectiveness results in that sotorasib arm patients continue to accrue the costs of treatment but not the	

NICE National Institute for Health and Care Excellence

		relative benefits of the treatment. **This issue is considered unresolved and the ICER ranges presented below will reflect this.**	
Key issue 10: TTD modelling approach inconsistent with OS and PFS modelling	NO	As described in the response to CQs, the base-case approach of connecting TTD to PFS by applying a fitted HR is reasonable and consistent with the clinical use of Sotorasib: treatment is continued until progression (or development of unacceptable AEs) with the majority discontinuations in CodeBreak100 being due to progression. In a similar way and for similar reasons, applying a HR to PFS, assumptions such as TTD=PFS or adding a mean number of cycles of treatment to newly progressed patients have been accepted in previous NSCLC appraisals. TTD is also very mature with >80% of patients having discontinued by the March 2021 data cut. This relative certainty means that applying parametric curves has a limited impact on the ICER and so the ICER ranges below reflect the ERG base-	The ERG agrees that the matter is resolved.

		case selection (fitted TTD parametric curves). The issue is considered resolved following the technical engagement call and the ICER ranges below reflect this.	
Key issue 11: Time-to-death utilities do not seem well-informed	NO	Amgen believe that both a time-to-death approach and a health state approach with a PFS utility differential (see issue 12 below) are plausible. As argued in the response to CQs, two independent sets of interviews were conducted with clinicians to validate visually the time-to-death and health state approaches and clinicians tended to favour time-to-death as more of a driver than health state (based on RECIST defined progression). The ERG had some concerns over the utility analyses: • There are limitations to all datasets, but for context: of the 123 patients in the safety dataset, 122 completed at least one eq-5d questionnaire (AN01), and of these 86 completed at least two including at baseline (AN02). Therefore, AN01 includes	The ERG re-iterates their critique on the use of the TTD approach as mentioned in the ERG report, which essentially is that TTD utility estimates were based on small sample size (in particular for those categories closer to death) and that insufficient information was provided on to assess reliability of the estimates and the superiority of the TTD approach over the health state approach. The ERG therefore will maintain their preferred assumption of using the health state approach for utilities.

		AN02 as a subset and was used as the basis of utility analyses to maximise sample size. • The ERG is correct that including baseline utility as a covariate in the MMRM models would mean we would be using the AN02 dataset, but although this was a *statistically* significant covariate it did not have a significant impact on results when excluded (e.g. table 33 in the company submission shows it had minimal impact on estimated disutility on progression) and so the larger AN01 dataset was deemed more appropriate. ○ Although some trade-off exists (sample size vs variable inclusion), all MMRM models already include a patient level random effect that takes account of correlations between observations of the same patient. This in a way has already adjusted for patients with baseline utility and may explain why	

NICE National Institute for Health and Care Excellence

		inclusion of the covariate had little impact. **This issue is considered unresolved following the technical engagement call. A time-to-death approach is plausible and this is reflected in the ICER ranges below.**	
Key issue 12: Disutility for IV administration not well justified	NO	Amgen agrees that this alone seems an arbitrary reason for a utility decrement, but as explained in the technical engagement call this should be seen in the context of an overall conservative HRQoL base-case. The company submission assumed equal on treatment PFS utilities for a targeted therapy (Sotorasib) vs. chemotherapy (docetaxel) but a differential is often seen in trials and accepted by committees (e.g. NICE TA628, TA416, TA406 and TA422). • This is the norm in other appraisals for targeted therapies in NSCLC. For example, a differential of 0.02 to 0.08 has been seen for ALK targeted therapies compared with chemotherapy.	In the clarification phase, the ERG asked for justification of the use of the disutility specifically in light of the fact that sotorasib is administered daily as 8 tablets, and docetaxel is administered once every three weeks. The company did not provide information on the potential disutility associated with the sotorasib dosing and frequency, and therefore the ERG considers the proposed disutility to be without sufficient justification. Observational data on HRQoL in a comparative setting would be required to resolve this issue. The ERG does not see any reason to adjust their preferred assumptions concerning this issue, although, as stated before, the ERG is in itself not opposed to the notion of a treatment related disutility for IV-administration.

NICE National Institute for Health and Care Excellence

		○ For example, see TA628: "This was found in PROFILE 1007, where utilities for the ALK TKI crizotinib (0.82, 95% CI: 0.79−0.85) were significantly greater ($p<0.05$) than for PDC (0.73, 95% CI: 0.70−0.79)…within the HRQoL SLR…for four of these studies, a comparison between ALK TKIs and chemotherapy was available and, in all instances, a utility decrement was applied for patients on chemotherapy compared to those receiving treatment with an ALK TKI (0.02–0.08)." ○ Applying AE decrements in only the first cycle of the model is not usually considered double counting. These are not expected to make a	

NICE National Institute for Health and Care Excellence

		difference (compared with differential on treatment PFS utilities) – as expected, removing these decrements from the base-case model has virtually no impact. The HRQoL SLR did not identify any KRAS specific utility data. PFS specific utilities from SELECT-1 are not available but there is a PFS utility available from LUME-Lung 1 (0.687 which is used in TA347 and TA416). When this is applied to the PFS base-case utility it implies a decrement of 0.047 (= 0.734 – 0.687). Given this, Amgen believe that scenarios with a health state utility approach and a 0.025 or 0.04 PFS (on treatment) utility differential between arms are reasonable compromises to explore. **This issue is considered unresolved following the technical engagement call. A time-to-death approach is plausible, but a reasonable compromise may be a health state approach with PFS utility differential (scenarios with 0.025 and**	

NICE National Institute for Health and Care Excellence

		0.04) and these are reflected in the ICER ranges below.	
Key issue 13: Relative dose intensity and wastage assumption not justified	NO	Amgen does not believe it is appropriate that relative dose intensity's (RDIs) should be equalised, given that this invalidates measured trial data. Trial data is usually considered more valid than intuitive assumptions in the hierarchy of evidence. **The issue of wastage is considered resolved following the technical engagement call (and inclusion of wastage is reflected in the ICER ranges below). The RDI assumption is not resolved and the ICER ranges presented below reflect this (small impact).**	In the response to clarification, the company stated that there was no reason to assume that RDI would be different between comparators. Given the impact on treatment costs, and immaturity of trial data, the ERG still prefers the conservative approach taken in the ERG base-case. For wastage, the ERG agrees that the issue is resolved.

Additional issues

Please use the table below to respond to additional issues in the ERG report 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 appraisal (e.g. at the clarification stage).

NICE National Institute for Health and Care Excellence

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: N/A	N/A	N/A	N/A

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

The table below reflects the updated confidential PAS for Sotorasib [redacted] and explores the spread of deterministic ICERs implied by the remaining unresolved issues that are relevant to the main comparator docetaxel. The 4 mix of settings at the bottom of the table reflect the following unresolved issues:

- *Issue 5: the alternative PSWA ICER is shown alongside the base-case MAIC ICER (last column) and this should be considered in decision making*
- *Issue 9 (TEW): the ICER ranges reflect TEW from years 2 (ERG preference), 3 and 4*
- *Issue 11 and 12 (utilities): the ICER ranges reflect health state (HS) utilities with no differential (ERG preference), HS utilities with a 0.025 PFS differential between arms, HS utilities with a 0.04 PFS differential between arms and time-to-death utilities (no decrement or differential)*
- *Issue 13 (RDI): the ICER ranges reflect assumed equal RDI (ERG preference) or RDI from trials (i.e. not equalised)*

NICE National Institute for Health and Care Excellence

Results for the minor comparator nintedanib+docetaxel are shown in Appendix E. Codebreak200 data is currently unavailable, but some PFS data (and potentially interim OS) may be available in 2022 depending on accrual of events (and thus unblinding) – estimated final completion is 2026 (up to 5 years of follow-up).

It should be noted that the probabilistic ICER (vs. docetaxel) is consistently lower by around £800 and this should be considered in decision making. Amgen is confident that this new PAS and the implied range of plausible ICER will allow the Committee to consider a positive recommendation (via baseline commissioning).

NICE National Institute for Health and Care Excellence

Key issue(s) in the ERG report that the change relates to	**Company's base case before technical engagement**	Change(s) made in response to technical engagement	**Impact on the company's base-case ICER**
Original submitted company base-case (with updated Sotorasib PAS)			ICER (base-case MAIC): **£38,715** ICER (PSWA analysis): **£33,811**
Updated base-case reflecting resolved issues (with updated Sotorasib PAS)	Original submitted base-case used HR applied to PFS to calculate TTD and assumed no wastage.	TTD Parametric curves and wastage included (ERG preferences)	Change from original base-case ICER: +£3,155 ICER (base-case MAIC): **£41,870** ICER (PSWA analysis): **£33,570**
1. ERG preferred settings: health state utilities with no PFS differential, RDI assumed equalised, while varying TEW	Original submitted base-case used time-to-death method for utilities (with IV decrement), RDI from trials (i.e. not equalised) and no TEW.	TEW from 2 years (ERG preferred base-case)	Change from original base-case ICER: +£13,054 ICER (base-case MAIC): **£51,769** ICER (PSWA analysis): **£43,412**
		TEW from 3 years	Change from original base-case ICER: +£11,337 ICER (base-case MAIC): **£50,052** ICER (PSWA analysis): **£41,791**

NICE National Institute for Health and Care Excellence

		TEW from 4 years	Change from original base-case ICER: +£10,332 ICER (base-case MAIC): **£49,047** ICER (PSWA analysis): **£40,802**
2. Health state utilities with lower PFS differential (0.025), RDI not equalised, while varying TEW		TEW from 2 years	Change from original base-case ICER: +£11,317 ICER (base-case MAIC): **£50,032** ICER (PSWA analysis): **£41,680**
		TEW from 3 years	Change from original base-case ICER: +£9,690 ICER (base-case MAIC): **£48,405** ICER (PSWA analysis): **£40,164**
		TEW from 4 years	Change from original base-case ICER: +£8,737 ICER (base-case MAIC): **£47,452** ICER (PSWA analysis): **£39,238**
3. Health state utilities with higher PFS differential (0.04), RDI equalised, while varying TEW		TEW from 2 years	Change from original base-case ICER: +£11,554 ICER (base-case MAIC): **£50,269** ICER (PSWA analysis): **£41,757**
		TEW from 3 years	Change from original base-case ICER: +£9,935 ICER (base-case MAIC): **£48,650** ICER (PSWA analysis): **£40,258**

NICE National Institute for Health and Care Excellence

		TEW from 4 years	Change from original base-case ICER: +£8,987 ICER (base-case MAIC): **£47,702** ICER (PSWA analysis): **£39,343**
4. Time-to-death utilities (no IV decrement), RDI equalised while varying TEW		TEW from 2 years	Change from original base-case ICER: +£9,387 ICER (base-case MAIC): **£48,102** ICER (PSWA analysis): **£39,528**
		TEW from 3 years	Change from original base-case ICER: +£7,697 ICER (base-case MAIC): **£46,409** ICER (PSWA analysis): **£37,992**
		TEW from 4 years	Change from original base-case ICER: +£6,707 ICER (base-case MAIC): **£45,422** ICER (PSWA analysis): **£37,059**

Appendix A: summary of clinical SLR

Neither SLR (RCTs and single-arm trials) identified trials that have a PDC arm in KRASm patients. Please see below the summary of interventions and comparators assessed as well as the inclusion criteria for patients. The Majority of identified RCTs required patients to have prior treatment with platinum-based chemotherapy. Several single-arm trials also specified that patients should have had prior treatment with platinum-containing chemotherapy regimen(s), supporting the argument that most patients have had previous PDC.

1) RCTs

Studies reporting on an exclusively KRASm study population (n=7)

- All but one of the studies assessed kinase inhibitors. Rulli et al 2015 compared docetaxel monotherapy to erlotinib monotherapy.
- Rulli et al 2015 and Carter et al 2016 required failure following a platinum-containing doublet chemotherapy regimen
  - Rulli et al 2015 reported 90% of patients had received first-line platinum-based therapy with the remaining patients having received chemotherapy in the adjuvant therapy
  - Carter et al 2016 reported that 40–55% and 45–60% of patients had received one or two prior regimens, respectively (none had more than two).

Studies reporting data for a KRASm subgroup

Chemotherapy as a comparator (n=6)

- Treatments assessed were nivolumab, atezolizumab, ganetespib, pelareorep and erlotinib, administered either as monotherapy or in combination with chemotherapy. Docetaxel or pemetrexed monotherapy was the control treatment in all six studies.
- The study by Ramalingam et al 2015 only required patients to have progressed following first-line therapy, while the remaining studies all specified prior treatment with platinum-based chemotherapy doublets (except for Bradbury et al 2018 where doublet therapy was not required in patients aged >70 years).

Erlotinib as a comparator (n=3)

- Spigel et al 2017 assessed onartuzumab (a MET-inhibiting humanized monoclonal antibody) in combination with erlotinib; Scagliotti et al 2015 assessed tivantinib in combination with erlotinib and Karampeazis et al 2013 assessed pemetrexed monotherapy.
- All three studies required one or two prior lines of platinum doublet chemotherapy (except in patients ≥65 years in Karampeazis et al 2013).

Other interventions (n=1)

- Ciuleanu et al 2017 study assessed linsitinib (a dual IGF-1R and IR inhibitor) given as maintenance therapy in conjunction with erlotinib (versus placebo + erlotinib) in patients with stages IIIB or IV NSCLC (with ECOG performance status 0–1) and stable disease or better following four cycles of first-line platinum-based chemotherapy.

2) Single arm trials

- 4 broadly defined drug classes were assessed across identified trials: inhibitors of KRASG12C, inhibitors of the EGFR/MAPK signaling pathway, immune checkpoint inhibitors, and modulators of protein stability.
- Sotorasib trial (CodeBreak 100)
- Antroquinonol trial (NCT02047344) included patients with disease progression after 2 prior LoTs (at least 1 platinum-based) or patients who refused treatment with approved treatments
- Gerber et al. 2020 (assessing Defactinib), NCT026420 (assessing Docetaxel and trametinib) and NCT02258607 (assessing Momelotinib and trametinib) required patients to have at least 1 prior platinum-based CHT
- Gulley et al. 2017 (assessing Avelumab) included patients who had progression after platinum-based doublet CHT for metastatic disease
- Pujol et al. 2020 (assessing Abemaciclib and pembrolizumab) included CHT-naïve patients with ≥ 1% TC PD-L1 staining in Cohort A and patients with ≤ 1 prior platinum-containing CHT regimen in Cohort B

NICE National Institute for Health and Care Excellence

Appendix B: Oncology Dynamics RWE survey to validate NSCLC pathway in UK

Overall Research Design

The aim of this analysis was to ascertain for the period of the last year what proportion of patients with locally advanced and/or metastatic NSCLC receiving 2nd Line docetaxel have had previous immunotherapies or platinum-based chemotherapy (monotherapy or in combination).

This was a retrospective analysis using readily available data from Oncology Dynamics TM (IQIVIA Ltd., London, UK). The inclusion criteria were as follows: adult patients aged ≥18 years diagnosed with advanced/metastatic NSCLC (and the subset who were confirmed EGFR/ALK/ROS-1 negative) and who received treatment between Q3 2020 to Q2 2021 in Oncology Dynamics.

Data Source

Oncology Dynamics is a physician-based cross-sectional survey that collects anonymised patient level data information on drug-treated cancer patients from ten countries including France, Germany, Italy, Spain and the UK in Europe. Data are collected on cancer-diagnosed patients regardless of cancer type, stage and/or treatment modality. It is designed as repeated quarterly cross-sectional cohorts, contains more than 167,000 cancer cases per year and over 35 cancer indications and covers demographic, diagnostic and treatment information recorded at time of diagnosis, as well as treatment information at ‘current’ and ‘previous’ line of therapy (as defined by the treating physician) (Kafatos et al, 2021a). The sampling plan for the database is set up in order to be representative of physician specialties treating cancer in each country with a sample size that adequately determines patterns of cancer management. The database therefore provides a representative sample of clinical practice among currently treated cancer populations. The validity of these data has been previously demonstrated for a wide range of oncology indications (Kafatos et al, 2021b; Maroun et al, 2018; Marchetti et al, 2017; Canta et al, 2016; Schmidt et al, 2014; Zhao et al, 2012; Inoue et al, 2009).

NICE National Institute for Health and Care Excellence

Results

Table 1 summarises that there was only 1 patient with advanced NSCLC in the UK that received docetaxel in 2L treatment. This patient received platinum-based chemotherapy with an anti-PD-1/L1i in 1L. Table 2 summarises 12 metastatic NSCLC patients receiving docetaxel in 2L in the UK. The majority 83% (n=10/12) had a PD-L1 expression of 1-49%. Nearly all patients (92%; n=11/12) received platinum-based chemotherapy with anti-PD-1/L1i in 1L before receiving docetaxel in second line.

Table 1: Prior systemic anti-cancer therapy received in 1L before 2L treatment with docetaxel in **Advanced NSCLC** patients in UK

	Advanced NSCLC							
	Overall				**EGFR/ALK/ROS-1 negative**			
	Overall		PD-L1 1-49%		Overall		PD-L1 1-49%	
	n	%	n	%	n	%	n	%
United Kingdom								
***N,* no. of patients in 2L receiving docetaxel**	**1**		**1**		**1**		**1**	
Types of prior systemic anticancer therapy in 1L								
Any anti-PD-1/L1i (monotherapy or combination)	1	100%	1	100%	1	100%	1	100%
anti-PD-1/L1i monotherapy								
Platinum-based chemotherapy WITH anti-PD-1/L1i	1	100%	1	100%	1	100%	1	100%
Platinum-based chemotherapy WITHOUT anti-PD-1/L1i								
Other								

NICE National Institute for Health and Care Excellence

Table 2: Prior systemic anti-cancer therapy received in 1L before 2L treatment with docetaxel in **Metastatic NSCLC** patients in UK

	Metastatic NSCLC											
	Overall						**EGFR/ALK/ROS-1 negative**					
	Overall		PD-L1 1-49%		PD-L1 >=50%		Overall		PD-L1 1-49%		PD-L1 >=50%	
	n	%	n	%	n	%	n	%	n	%	n	%
United Kingdom												
***N,* no. of patients in 2L receiving docetaxel**	**12**		**10**		**2**		**7**		**6**		**1**	
Types of prior systemic anticancer therapy in 1L												
Any anti-PD-1/L1i (monotherapy or combination)	11	92%	9	90%	2	100%	7	100%	6	100%	1	100%
anti-PD-1/L1i monotherapy												
Platinum-based chemotherapy WITH anti-PD-1/L1i	11	92%	9	90%	2	100%	7	100%	6	100%	1	100%
Platinum-based chemotherapy WITHOUT anti-PD-1/L1i	1	8%	1	10%								
Other												

References

Canta FB, Mitrofan L, Karanis Y, Mistry H, Anger CJaOO. 17P deletion and TP53 gene mutation (17P/TP53) testing behaviour and treatment patterns for chronic lymphocytic leukemia (CLL) patients in France, Germany, Italy, Spain and UK (EU5). 2016;27:vi317.

Inoue Y, Toiyama Y, Tanaka K, Miki C, Kusunoki MJJJOCO. A comprehensive comparative study on the characteristics of colorectal cancer chemotherapy. 2009;39:367-375.

Kafatos G, Banks V, Burdon P, et al. Biomarker testing and mutation prevalence in metastatic colorectal cancer patients in five European countries using a large oncology database. 2021a;17:1483-1494.

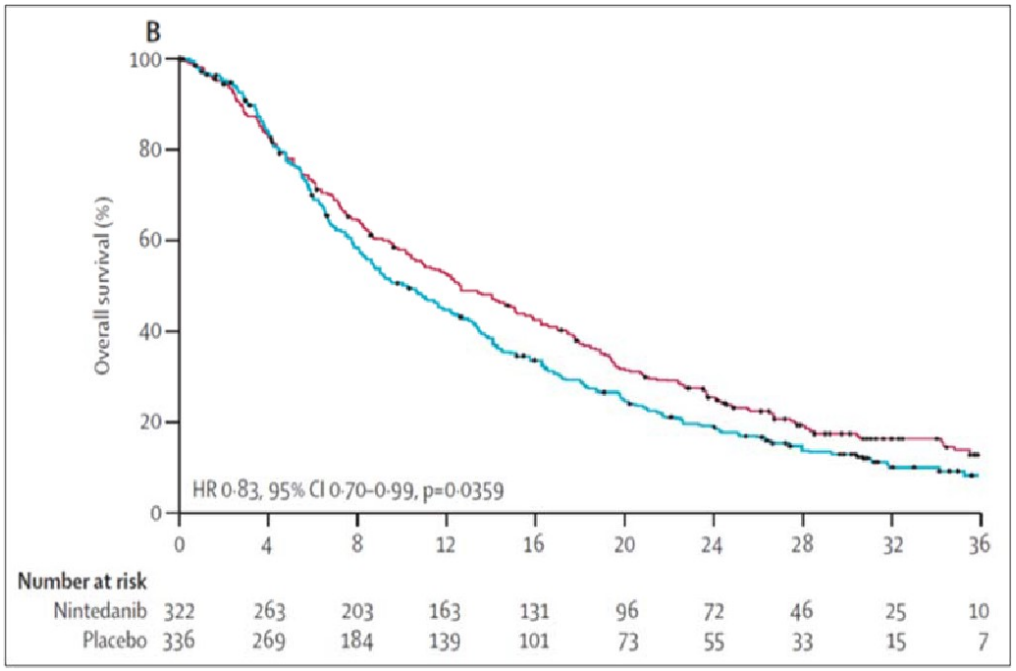
Kafatos G, Banks V, Burdon P, et al. Impact of biomarkers and primary tumor location on the metastatic colorectal cancer first-line treatment landscape in five European countries. 2021b;17:1495-1505.

Marchetti P, Maass N, Gligorov J, et al. Patient database analysis of fulvestrant 500 mg in the treatment of metastatic breast cancer: A European perspective. 2017;32:247-255.

Maroun R, Mitrofan L, Benjamin L, et al. Real life patterns of care and progression free survival in metastatic renal cell carcinoma patients: retrospective analysis of cross-sectional data. 2018;18:1-7.
Schmidt N, Kostev K, Jockwig A, Kyvernitakis I, Albert U-S, Hadji PJIJCPT. Treatment persistence evaluation of tamoxifen and aromatase inhibitors in breast cancer patients in early and late stage disease. 2014;52:933-939.
Zhao Z, Pelletier E, Barber B, et al. Major surgery in patients with metastatic colorectal cancer in Western Europe. 2012;43:456-461.

NICE National Institute for Health and Care Excellence

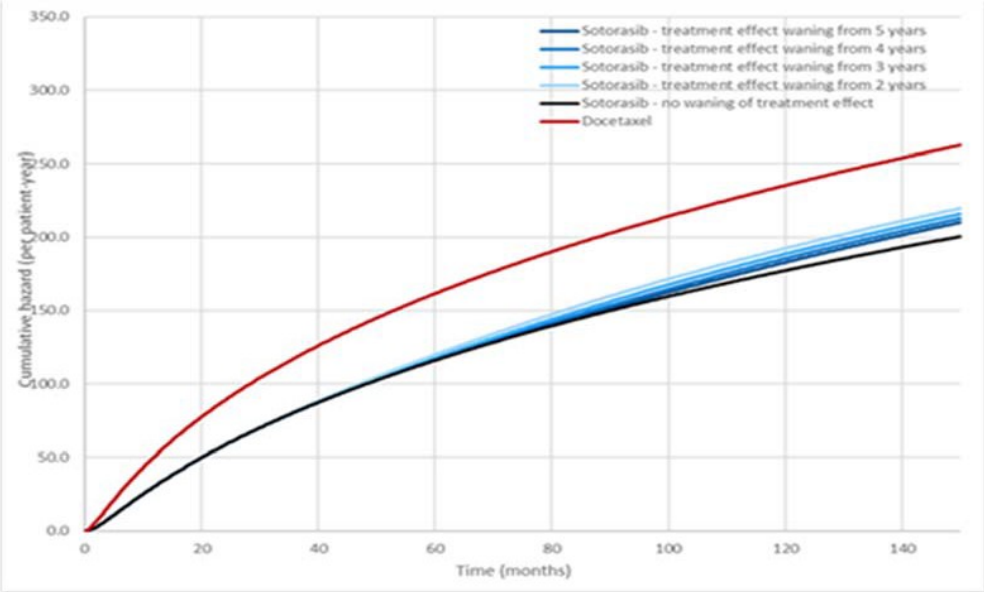
Appendix C: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line)



NICE National Institute for Health and Care Excellence

Appendix D: TEW scenario graphics (OS cumulative hazard plots)

The cumulative hazard plot (per patient-year) for the following scenarios are presented below: docetaxel base-case OS, Sotorasib base-case OS (no TEW) and gradual 5-year TEW applied to the Sotorasib OS curve from 2, 3, 4 and 5 years. The hazard plot is as expected, with the cumulative hazards of death becoming more and more equal depending on when waning begins.



NICE National Institute for Health and Care Excellence

Appendix E: results for minor comparator nintedanib+docetaxel

The table below reflects the updated confidential PAS for Sotorasib and explores the spread of deterministic ICERs (vs. nintedanib+docetaxel). Despite unresolved issues that impact this comparison (TEW, utilities and RDI), for brevity the ERG preferred base-case assumptions are set with only variation related to issue 8. **It should be noted that the probabilistic ICER for this minor comparator (vs. nintedanib+docetaxel) is consistently lower by around £1800 and this should be considered in decision making.**

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case ICER
ERG preferred settings: health state utilities with no PFS differential, RDI assumed equalised, TEW from 2 years	Original submitted base-case used time-to-death method for utilities (with IV decrement), RDI from trials (i.e. not equalised) and no TEW. In addition, a piecewise approach with three fitted HRs to obtain a treatment effect for add-on nintedanib (vs. docetaxel).	ERG base-case: three period piecewise method with first period HR=1	Change from original base-case ICER: +£8,514 ICER (base-case MAIC): **£42,142** ICER (PSWA analysis): **£22,997**
		Two period piecewise method with first period HR=1 (New fitted HR for period 6m+ is 0.7904)	Change from original base-case ICER: +£9,004 ICER (base-case MAIC): **£42,632** ICER (PSWA analysis): **£23,150**
		Two period piecewise method (all fitted)	Change from original base-case ICER: +£668 ICER (base-case MAIC): **£32,960** ICER (PSWA analysis): **£19,267**

in collaboration with:

university of groningen

University Medical Center Groningen

Addendum to:

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Produced by	Kleijnen Systematic Reviews (KSR) Ltd in collaboration with University Medical Center (UMC) Groningen, the Netherlands
Authors	Robert Wolff, Managing Director, KSR Ltd, United Kingdom (UK) Thea van Asselt, Health Economist, UMC Groningen, the Netherlands Sajad Emamipour, Health Economist, UMC Groningen, the Netherlands Simon van der Pol, Health Economist, UMC Groningen, the Netherlands Maarten Postma, Health Economist, UMC Groningen, the Netherlands Annette Chalker, Systematic Reviewer, KSR Ltd, UK Pawel Posadzki, Reviews Manager, KSR Ltd, UK Charlotte Ahmadu, Health Economist, KSR Ltd, UK Nigel Armstrong, Health Economics Manager, KSR Ltd, UK Sean Harrison, Statistician, KSR Ltd, UK Shelley de Kock, Information Specialist, KSR Ltd, UK Jos Kleijnen, Director, KSR Ltd, UK
Correspondence to	Robert Wolff, Kleijnen Systematic Reviews Ltd Unit 6, Escrick Business Park Riccall Road, Escrick York YO19 6FD United Kingdom
Date completed	06/12/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number 13/51/34.

Declared competing interests of the authors:
None.

Acknowledgements

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.

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.

Addendum to ERG report

This addendum presents deterministic and probabilistic ICERs of additional company base-case and scenarios as well as additional ERG scenarios.

Table 1 and Table 2 contain the company scenarios for the docetaxel and docetaxel + nintedanib comparators, respectively. Table 3 and 4 present further ERG scenarios for the docetaxel and docetaxel + nintedanib comparators, respectively.

Table 01: Company base-case and scenarios (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
Company base-case with updated PAS [redacted]							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	38,642	75.9%
Updated company base-case with TTD and wastage assumptions adjusted (no TEW included)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	41,120	70.1%
Updated company base-case PSWA (no TEW included)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	32,950	89.4%
Updated company base-case ERG settings apart from TEW: starting at 3 yrs + 5 yrs gradual waning							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	49,275	50.3%
Updated company base-case with TEW at 3+5 yrs and absolute PFS utility difference of 0.025*							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	49,164	58.8%
Updated company base-case with TEW at 3+5 yrs and absolute PFS utility difference of 0.04#							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	45,701	60.3%
Updated company base-case with TEW at 3+5 yrs, RDI equalized, TTD utilities, no IV disutility							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	45,709	59.0%

Source: CS updated model

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
CS = company submission; ERG = Evidence Review Group; FV = fixing violation; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year, RDI = relative dose intensity; TTD = time to treatment discontinuation * For this scenario the ERG could not exactly reproduce the ICER the company reported (£48,405) and so the ERG ran the probabilistic sensitivity analysis according to the settings they believed to best fit the description of the scenario by the company. # For this scenario the ERG could not exactly reproduce the ICER the company reported (£48,650) and so the ERG ran the probabilistic sensitivity analysis according to the settings they believed to best fit the description of the scenario by the company.							

Table 02: Company base-case and scenarios (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
Company base-case with updated PAS [redacted]							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27.117	80.9%
Updated company base-case with TTD and wastage assumptions adjusted (no TEW included)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	27,975	80.2%
Updated company base-case PSWA (no TEW included)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	15,205	94.6%
Updated company base-case with ERG preferences three period piecewise HR (with HR =1 in first period)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,663	63.9%
Source: CS updated model							

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
CS = company submission; ERG = Evidence Review Group; FV = fixing violation; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year, RDI = relative dose intensity; TTD = time to treatment discontinuation							

Table 3: ERG base-case and additional scenarios (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
ERG base-case with updated PAS ([redacted]) TEW set to start at 2 years from start of treatment, plus 5 yrs gradual decrease for HR to return to 1							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	50,789	49.6%
ERG base-case 3 yrs TEW (TEW starts at 3 yrs from start of treatment, HR set to 1, no gradual decrease)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	53,791	38.7%
ERG base-case 5 yrs TEW (TEW starts at 5 yrs from start of treatment, HR set to 1, no gradual decrease)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	49,303	50.1%
ERG base-case 3 yrs TEW plus scenarios (gen gamma for PFS, treatment emergent AEs, disutility of 0.05 for AEs that had zero disutility in CS)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,869	33.1%

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
ERG base-case 5 yrs TEW plus scenarios (gen gamma for PFS, treatment emergent AEs, disutility of 0.05 for AEs that had zero disutility in CS)							
Docetaxel	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	51,192	45.3%
Source: CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violation; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year, RDI = relative dose intensity; TTD = time to treatment discontinuation							

Table 04: ERG base-case and additional scenarios (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
ERG base-case with updated PAS ([redacted]) TEW set to start at 2 years from start of treatment, plus 5 yrs gradual decrease for HR to return to 1							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,663	63.9%
ERG base-case 3 yrs TEW (TEW starts at 3 yrs from start of treatment, HR set to 1, no gradual decrease)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,970	64.0%

Technologies	Total costs (£)	Total QALYs	Deterministic Incremental costs (£)	Deterministic Incremental QALYs	Deterministic ICER (£/QALY)	Probabilistic ICER (£/QALY)	Probability cost-effective at WTP £50,000
ERG base-case 5 yrs TEW (TEW starts at 5 yrs from start of treatment, HR set to 1, no gradual decrease)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,723	62.9%
ERG base-case 3 yrs TEW plus scenarios (gen gamma for PFS, treatment emergent AEs, disutility of 0.05 for AEs that had zero disutility in CS)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,030	64.1%
ERG base-case 5 yrs TEW plus scenarios (gen gamma for PFS, treatment emergent AEs, disutility of 0.05 for AEs that had zero disutility in CS)							
Docetaxel + nintedanib	[redacted]	[redacted]					
Sotorasib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	39,790	62.9%

Source: CS updated model

CS = company submission; ERG = Evidence Review Group; FV = fixing violation; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year, RDI = relative dose intensity; TTD = time to treatment discontinuation