Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Contents:

The following documents are made available to stakeholders:

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

1. **Company submission from Sanofi:**
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**
   a. Initial clarification response
   b. Additional clarification response

3. **Patient group, professional group, and NHS organisation submissions** from:
   a. Anthony Nolan
   b. NHS England, Specialised Commissioning

4. **External Assessment Report** prepared by BMJ Technology Assessment Group

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

6. **Technical engagement response from company, Sanofi**

7. **Technical engagement responses and statements from experts:**
   a. Professor Adrian Bloor – Consultant Haematologist, Clinical Expert nominated by NHS England
   b. Further response from Professor Adrian Bloor
   c. Daniel Yeates – Patient Expert nominated by Anthony Nolan

8. **Technical engagement responses from stakeholders:**
   a. Novartis

9. **External Assessment Group critique of company response to technical engagement** prepared by BMJ TAG
   a. EAG critique
   b. Updated EAG base case post TE

10. **Additional External Assessment Group scenarios** prepared by BMJ TAG
    a. Additional scenarios
    b. Additional scenarios part 2

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

Belumosudil for treating chronic graft-versus-host disease after 2 or more lines of systemic therapy [ID4021]

Document B
Company evidence submission

March 2023

File name	Version	Contains confidential information	Date
ID4021_Belumosudil_Document_B_20032023_ACIC_fully redacted	Final	Yes Academic and Commercial in confidence redacted	20 March 2023

Contents

Tables and figures

List of tables

List of figures

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

B.1.1. Decision problem

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

- Belumosudil is indicated for the treatment of patients aged 12 years and older with chronic graft-versus-host disease (chronic GVHD) who have received at least two prior lines of systemic therapy.(1)

The decision problem addressed in this submission, compared with that defined in the final scope issued by NICE, 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	People aged 12 years and over with chronic GVHD after 2 or more lines of systemic therapy.	As per final scope	N/A
Intervention	Belumosudil with established clinical management.	As per final scope	N/A
Comparator(s)	Established clinical management without belumosudil, including: • ECP • Imatinib • Rituximab • Sirolimus • MMF • Tacrolimus • Cyclosporine	As per final scope, excluding CNIs (i.e. tacrolimus and cyclosporine)	CNIs are generally used in the 1st or 2nd line as steroid sparing agents. They are not recommended in later treatment lines and would not be used as a standalone treatment. Therefore, CNIs are not considered as comparators to belumosudil in the 3rd line setting. See Section B.1.3.2 for further details.
Outcomes	The outcome measures to be considered include: • Response to treatment (including complete response and overall response) • Immunosuppressant sparing • Mortality • Treatment AEs • FFS • HRQoL	As per final scope	N/A
Economic analysis	The reference case stipulates that: • The cost-effectiveness of treatments should be expressed in terms of incremental cost per QALY. • 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.	As per final scope	N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	• 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 availability of any managed access arrangement for the intervention will be taken into account. • The availability and cost of biosimilar and generic products should be taken into account.		
Subgroups to be considered	If the evidence allows the following subgroups will be considered. These include: • Different organs or tissues affected by chronic GVHD • Number and type of previous treatments	As per final scope	N/A

AEs = adverse effects; CNIs = calcineurin inhibitors; ECP = extracorporeal photopheresis; FFS = failure-free survival; GVHD = graft-versus-host disease; HRQoL = health-related quality of life; MMF = Mycophenolate mofetil; N/A = not applicable; NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year

B.1.2. Description of the technology being evaluated

A description of belumosudil, the technology being appraised, has been summarised in Table 2. The summary of product characteristics and UK public assessment report are included in Appendix C.

Table 2. Technology being evaluated

UK approved name and brand name	Belumosudil (Rezurock®)
Mechanism of action	Belumosudil is a first-in-class potent and selective ROCK2 inhibitor which targets both immune response dysfunction and downregulates fibrotic processes associated with chronic GVHD.(1, 2)
Marketing authorisation/CE mark status	Belumosudil received marketing authorisation in Great Britain on 7 July 2022 and received orphan designation at the point of licensure.(1, 3) The marketing authorisation review was conducted under Project Orbis, a programme to review and approve promising (cancer) medicines and help patients access treatments faster.(4) The Project Orbis procedure relates to Great Britain only and does not include Northern Ireland.(4) Belumosudil was also granted an innovation passport by the MHRA in April 2021 (ILAP reference number ILAP/IP/21/53904/01).
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Belumosudil is indicated for the treatment of patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy.(1) There are no other licensed indications for belumosudil.
Method of administration and dosage	• The recommended dose of belumosudil is 200 mg administered orally once daily at approximately the same time with a meal.(1) • Treatment should continue until disease progression or unacceptable toxicity.(1) • The dose of belumosudil should be increased to 200 mg twice daily when co-administered with strong CYP3A inducers or proton pump inhibitors.(1) • Belumosudil treatment should be initiated and supervised by physicians experienced in the management of chronic GVHD.(1)
Additional tests or investigations	A complete blood cell count and liver function test must be performed before initiating therapy with belumosudil.(1) Perform liver function tests at least monthly throughout treatment.(1)
List price and average cost of a course of treatment	List price: £6,708.00 per box of 30 x 200 mg tablets Average cost of treatment course*: £67,326.62
Patient access scheme (if applicable)	Simple PAS (percentage discount)

CYP3A = Cytochrome P450, family 3, subfamily A; GVHD = graft-versus-host disease; ILAP = Innovative Licensing and Access Pathway; MHRA = Medicines and Healthcare products Regulatory Agency; PAS = patient access scheme; ROCK2 = Rho-associated, coiled-coil containing protein kinase-2
*Based on median treatment duration of 9.2 months for belumosudil once daily and 11.2 months for belumosudil twice daily. Assumes 95% once daily and 5% twice daily dosing according to expected clinical practice.

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. Introduction to chronic GVHD

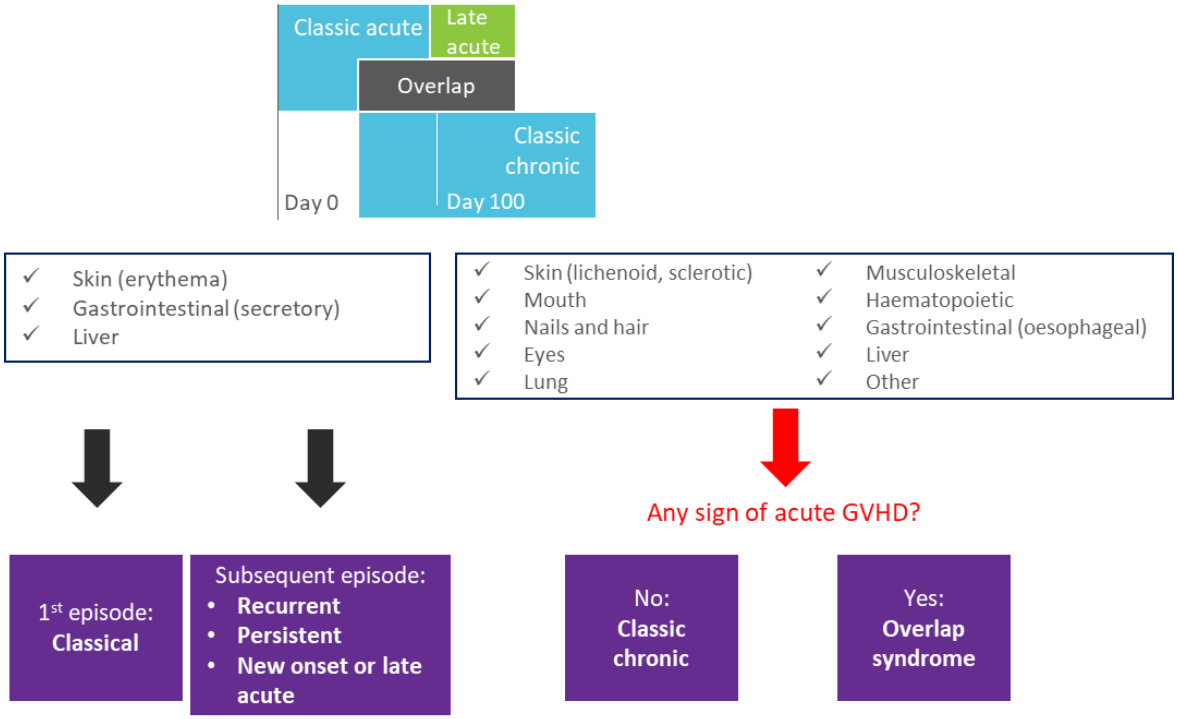
Graft-versus-host disease (GVHD) is an immune-mediated inflammatory and fibrotic disorder that is characterised by tissue damage (inflammation and fibrosis) and multi-system organ involvement.(5, 6) It is a serious complication of allogeneic haematopoietic stem cell transplant (alloHSCT), an intervention used to treat life-threatening conditions such as haematological diseases, solid tumours and immune disorders, when donor cells target tissues in the recipient of the transplant.(7)

GVHD can be classified as acute or chronic. Traditional diagnosis of acute vs. chronic GVHD relied on the timing of the manifestations, with GVHD arising within 100 days of transplant diagnosed as acute GVHD, and later-arising GVHD diagnosed as chronic GVHD.(7) Acute and chronic GVHD are now defined based on their differing pathophysiology and clinical manifestations. Acute GVHD is primarily characterised by inflammation and typically affects the skin, gastrointestinal (GI) tract or liver.(5, 7) In contrast, chronic GVHD is primarily characterised by fibrosis and affects a wider range of tissues (Figure 1).(5, 7) Inflammation is also present in chronic GVHD; however, it presents differently compared with acute GVHD, both in terms of pathophysiology and organs affected.(5, 7)

The different forms of GVHD may evolve and overlap. When chronic GVHD presents without signs of acute GVHD, it is defined as classic chronic GVHD; when signs of acute GVHD are also present, the patient is classified as having overlap disease (a subcategory of chronic GVHD).(6, 7)

Additional details on the manifestations of chronic GVHD and the associated symptomatic and HRQoL burden are described later in Section B.1.3.1.4.

Figure 1. Classification of acute and chronic GVHD



GVHD = graft-versus-host disease
SOURCE: Adapted from Mawardi 2019(7)

Clinical patterns of chronic GVHD onset include de novo (i.e., without prior acute GVHD), progressive (i.e., progressing directly from acute GVHD) or quiescent (i.e., following complete resolution of acute GVHD).(7) The majority of chronic GVHD patients develop manifestations within the first year following alloHSCT.(6) A retrospective study conducted in adult recipients of alloHSCT in England between 2008 and 2012 found that the median time to onset in patients without prior history of GVHD was 191.5 days, and 168.5 days in patients with a prior history of GVHD.(8) Chronic GVHD therefore places a heavy burden on patients who have only very recently undergone treatment for and are in remission from a life-threatening illness (e.g., acute myeloid leukaemia, myelodysplastic syndrome, acute lymphoblastic leukaemia, anaemia due to bone marrow failure).(9)

The severity of chronic GVHD is graded as mild, moderate or severe, based on the 2014 National Institutes of Health (NIH) Consensus Criteria, and provides standardisation for diagnosis and assessment of chronic GVHD severity across clinical trials (Table 3).(6)

Table 3. Chronic GVHD assessment by the 2014 NIH consensus criteria

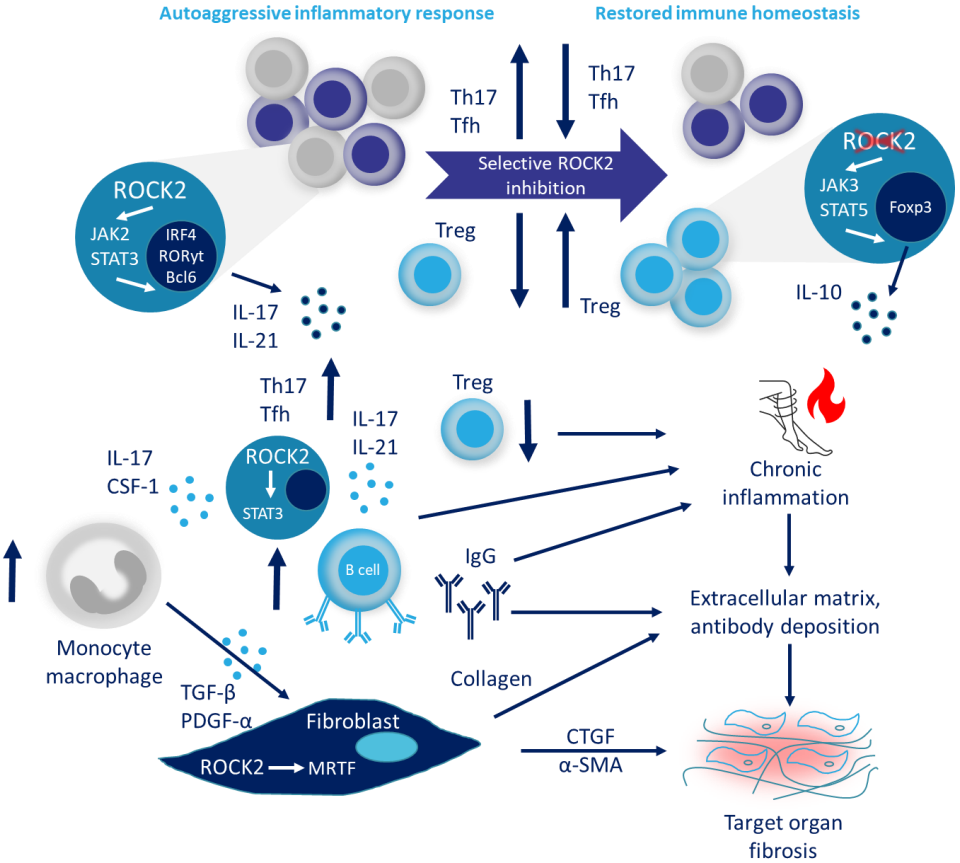
Category	Organs involved, n	Maximum severity
Mild	≤2	1 (0 for lung)
Moderate (a)	≥3	1 (0 for lung)
Moderate (b)	Any	2 (1 for lung)
Severe	Any	3 (2 for lung)

GVHD = graft-versus-host disease; NIH = National Institutes of Health
Maximum severity scale, 0: No clinical manifestations/symptoms; 1: Clinical manifestations with no more than mild disability; 2: Clinical manifestations with moderate disability; 3: Clinical manifestations with severe disability
Source: Adapted from Jagasia 2015(6)

The pathology of chronic GVHD is driven by T-cells, B-cells, macrophages, dendritic cells and neutrophils, resulting in inflammation, immune dysregulation and fibrosis.(10) Rho-associated coiled-coil containing protein kinase 2 (ROCK2) has been shown to play an integral role in the cascade of cytokines and differentiation of cell types that lead to chronic GVHD.(2) ROCK2 modifies a key transcription factor responsible for the production of pro-inflammatory cytokines IL-17 and IL-21, resulting in downstream signalling that stimulates the differentiation of Th17 cells into Tfh cells, which in turn promote the production of self-reactive mature B-cells.(2) Furthermore, ROCK2 regulates the expression of genes associated with fibrosis.(2) Thus, decreasing ROCK2 signalling offers the potential to restore immune homeostasis, modulating, rather than suppressing, immune function and avoiding abnormal, fibrotic tissue repair.(2)

As described in Section B.1.2, belumosudil is a ROCK2 inhibitor capable of targeting the fibrotic and inflammatory processes of chronic GVHD.

Figure 2. Role of the ROCK2 pathway in the pathology of chronic GVHD



CSF = colony-stimulating factor; CTGF = connective tissue growth factor; Foxp3 = forkhead box P3; GVHD = graft-versus-host disease; IgG = immunoglobulin G; IL = interleukin; IRF = interferon regulatory factor; JAK = Janus-associated kinase; MRTF = myocardin-related transcription factor; PDGF = platelet-derived growth factor receptor; ROCK2 = rho-associated coiled-coil containing protein kinase 2; ROR = retinoic-acid-receptor-related orphan nuclear receptor; STAT = signal transducer and activator of transcription; Tfh = follicular helper T-cell; TGF = transforming growth factor; Treg = regulatory T-cell
SOURCE: Adapted from Zanin-Zhorov, 2021(2)

B.1.3.1.2. Epidemiology of chronic GVHD

There is a lack of recent incidence data for chronic GVHD in England. In order to fill this gap, we consulted English clinical experts in an advisory board conducted in January 2023. The clinical experts estimated that approximately 150 patients each year with chronic GVHD in England require treatment with a third systemic line of therapy.(11) This is in line with the scope provided by NICE which estimated that 142 patients per year would require treatment (however, this estimate only considered patients with extensive chronic GVHD).(11)

Our estimate of 150 patients each year is also in line with available data from the British Society of Blood and Marrow Transplantation and Cellular Therapy (BSBMTCT), collected prior to the COVID-19 pandemic:

- In 2020, the BSBMTCT recorded a total of 1,476 allogeneic stem cell transplants in the UK and Republic of Ireland.(9)
- Of the alloHSCTs conducted in the UK and Republic of Ireland between 2009 and 2014, 90.8% took place in England.(12) Applying this rate to the UK total for 2020 results in 1,340 expected alloHSCTs in England (90.8% x 1,476=1,340).(12)
- Between 2009 and 2014, the incidence of chronic GVHD was 30% across all types of alloHSCT.(12) Applying this rate to the expected number of alloHSCTs in England results in 402 cases of chronic GVHD (30% x 1,340=402).
- It is estimated that approximately 35% to 40% of patients with chronic GVHD are expected to need third-line treatment during the course of their disease.

Prevalence figures for chronic GVHD in England are not currently available. However, given the incidence and high mortality (Section B.1.3.1.3) associated with chronic GVHD, the prevalence is not expected to be higher than 1 in 50,000. Of these, only a small proportion will have received two or more prior systemic therapies and be eligible for treatment with belumosudil which received orphan designation from the MHRA at licensure.

B.1.3.1.3. Morbidity and mortality of chronic GVHD

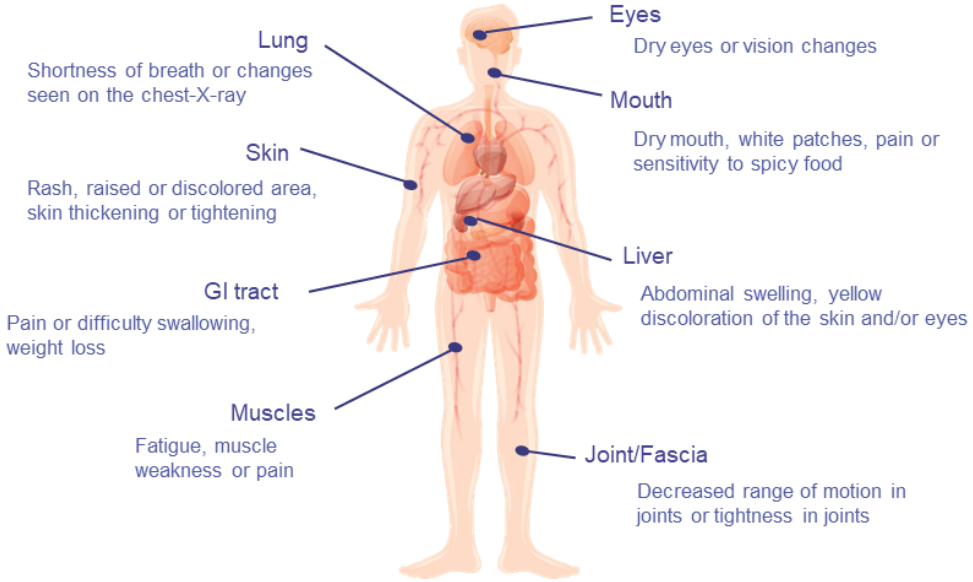
Chronic GVHD is a leading cause of morbidity and non-relapse mortality following alloHSCT and may affect one or multiple organs (Figure 3).(5, 6)

Chronic GVHD can affect almost any organ, but most commonly affects the skin, mouth and eyes, resulting in oral lesions, dry or gritty eyes that can cause great discomfort, hair loss and severe skin erythema (rash).(6, 13) Inflammatory and fibrotic manifestations of chronic GVHD include fasciitis (pain and swelling, stiffness and restriction of range of motion), dry eye syndrome, scleroderma (hardening/tightening of the skin) and bronchiolitis obliterans syndrome (irreversible obstruction of the bronchioles due to inflammation).(6, 7, 14) Both bronchiolitis obliterans and widespread sclerotic skin manifestations caused by inflammation and fibrosis are associated with poor survival.(15) Patients may also experience GI manifestations (including wasting syndrome, pancreatic atrophy and exocrine insufficiencies leading to malabsorption of nutrients) and liver manifestations that may lead to jaundice, acute hepatitis or progressive cholestatic features.(6) Pericardial or pleural effusion,

nephrotic syndrome, Raynaud's phenomenon, myasthenia gravis, peripheral neuropathy and cardiac involvement can also occur.(6)

The majority of chronic GVHD patients have ≥3 organs involved at the time of diagnosis.(14)

Figure 3. Clinical presentations of chronic GVHD



GI = gastrointestinal; GVHD = graft-versus-host disease
SOURCE: Adapted from Jagasia 2015(6), Mawardi 2019(7) and Salhotra 2020(14)

In patient interviews we have conducted in the US, a patient stated that "*chronic GVHD has impacted every facet of my life. I can't go out in the sun. 85% of my skin is affected by chronic GVHD. I can't be a mom to my children or do outdoor activities. My mobility is limited by joint stiffness and pain. I am fatigued, and I sleep a lot*."(16) In the same interview series, a patient with fibrosis stated: "*Because of fibrosis, there are definitely things I can't do now that I used to be able to do. For example, fully lifting my arms, exercising outside, running or lifting heavy items. It's like a rubber band that's been tightened*."(16)

Chronic GVHD is a leading cause of mortality in patients following alloHSCT.(17) In the UK, GVHD was responsible for 27% of non-relapse deaths between 2009 and 2014 in patients who received alloHSCT.(12)

In an advisory board conducted with nine English clinical experts during January 2023, one clinician stated that "*These patients have wide ranging and profound medical, psychological and social needs. They use an amount of resource that is undefined, and certainly grossly under-appreciated... [there are also] indirect losses to society with the morbidity these patients suffer and the care they require... this is a multi-system disease that results in pan-system effects*."(11)

B.1.3.1.4. Humanistic burden of chronic GVHD

Impact of the disease on patients

Patients who develop chronic GVHD face a multifaceted burden driven by disease severity and the involvement of multiple organs,(18, 19) which exacerbates the distress of already being diagnosed with a life-threatening illness that requires a stem cell transplant.(20) In patient interviews conducted in the US, a patient stated: "*If they told me, 'By the way, after this, you're going to experience another nightmare,' I don't know if I would have wanted to know. I don't know if I would have had the strength to fight the first battle if I knew another battle was right behind it*".(21)

Chronic GVHD can lead to debilitating consequences such as loss of sight, joint contractures and end-stage lung disease which can severely limit activities of daily living and may result in permanent disability and, in some cases, death.(22) Patients with chronic GVHD experience significant impairments in their HRQoL compared with the general population, including deficits in their physical and social functioning.(23) When asked to describe chronic GVHD in three words, a US patient stated: "Fatigue. Constant discomfort. Pain."(24)

A systematic review of evidence published between 2007 and 2017 concluded that the most important factors impacting HRQoL in patients with chronic GVHD were disease severity and type of organ involvement (with skin, GI, lung and joint/fascia manifestations having the greatest negative impacts on HRQoL).(18) A systematic literature review (SLR) conducted by us in December 2022 (described in Appendix H), noted that the following factors reduced the HRQoL of patients with chronic GVHD: disease severity, fatigue, depression, anxiety, financial burden and malnourishment.(25) When asked to describe the impact of chronic GVHD on quality of life, a US patient stated: "*One symptom that I find harder to deal with than others is the pain. There were sores because the T-cells were attacking the inside of my mouth. So pain with eating and discomfort. That is a low-grade constant pain, but then when you're eating, it's sharp. But eating is an important part of getting well and recovery. When you're growing a new immune system, it takes a huge number of calories*."(26)

In addition to functional impairments, patients with chronic GVHD experience significant psychological distress, including depression and anxiety symptoms.(20) When describing the impact of chronic GVHD after surviving the initial disease that lead to alloHSCT, a US patient stated: "*When I have to be hospitalised for anything regarding GVHD, it's traumatic. It brings up a lot of posttraumatic stress disorder —the beeps of the machines, the IV lines, the nurses and the smells. When you've gone through something like I have, and the primary disease, thank goodness, has not returned, and you're in the hospital for the secondary disease, it's a mindset. It messes with your mind. It's like, wait, I fought this major disease, and now I'm in a hospital for another disease that's related to the disease. It's absolutely terrible on the mind. It's terrible*."(26)

The involvement of skin, eye and hair manifestations can also result in chronic GVHD being a highly visible disease and can cause distress due to its visual effects on the patients' appearance (e.g., through skin pigmentation [hypo- and hyperpigmentation], skin ulcers, poor wound healing, skin lesions, premature hair greying, hair loss, thinning and brittleness, as well as conjunctivitis and other visible changes to the eyes).(6)

An interview study with eight patients with acute or chronic GVHD in England found that patients often feel restricted in what activities they can do, anxiety regarding the unpredictable manifestations of GVHD and inability to plan for the future, inability to go out in public or see family due to the risk of infection, and difficulty adapting to life as a "sick person" with multiple medications and frequent appointments, all of which can lead to depression.(20) When describing the process of getting ready for daily activities, a US patient stated: "*It used to take me about maybe 30 minutes to prepare to go out or to go to work. Now it takes me about 3 hours. I have to make sure that I have all of the medications, creams, eye drops, and mouthwash done, or else I will pay for it if I did not*."(26)

Published quotes from English patients reflect similar experiences: "*GVHD completely wrecked [my quality of life]. When I was suffering particularly badly I was literally unable to move much… I was unable to move between my bedroom, my office... I was unable to drive, to do, like I say, basic things around the house, even… it really impacted me in terms of my energy levels and my focus levels*."(20)

Patient experts invited to participate in a recent Sanofi-sponsored NICE Scientific Advice meeting reported that skin manifestations of chronic GVHD can cause the sensation of "debilitating cold" and affect their ability to spend time outdoors.(27) The patients further stated that manifestation in the gut can affect a patient's ability to eat, and that fatigue is a common symptom that affects their ability to work and function normally.(27) Additionally, if chronic GVHD affects the nervous system, it can cause permanent disability and impact the mental health and emotional well-being of a patient.(27) A US patient stated that "*It [chronic GVHD] can really make you feel like you just can't relax. It's there in your life, and you can't always go about doing what you want."(24) Another patient stated that chronic GVHD "has made a difference in every aspect of my life – what I do, who I see, who I spend time with and so many things*."(16)

A multi-national observational study of patients (n=371), caregivers (n=157) and physicians (n=107) was conducted in the UK, France, Germany, Italy, Spain and Canada to explore the current clinical practices and impact of chronic GVHD, using scientifically validated Disease Specific Programme (DSP) methodology.(28, 29) The study included UK physician-reported patient records (n=40), chronic GVHD patient or caregiver self-completion forms (n=2) and surveys with haematologists or oncologists (n=14).(28) Patients with chronic GVHD face considerable economic pressures due to lost wages and employment changes resulting from the difficulties associated with chronic GVHD.(28, 30) Of the patients with chronic GVHD who progressed after two prior lines of therapy included in the multi-national observational study, 39% were on long-term sick leave, retired or unemployed as a result of chronic GVHD.(31) The most common GVHD symptoms prompting people to stop working

were muscle weakness (89%), fatigue (78%), skin problems (78%), shortness of breath (67%) and eye or vision difficulties (67%).(31) In the UK, 61% of patients reported that they had to be hospitalised in the past 12 months, with infection listed as the predominant reason for hospitalisation (82%).(28) In a time trade-off study conducted among members of the general population in England (n=207) that aimed to compare pre- and post-transplant health states, respondents associated chronic GVHD with a significant disutility (-0.42, p<0.0001) compared with a transfusion-independent health state post-transplant.(32)

To gain a better understanding of the public perceptions of the impact on HRQoL among patients with chronic GVHD who have received two prior therapies, we conducted a non-interventional HRQoL elicitation exercise.(33) Four health state vignettes were developed based on previously published studies and validated through semi-structured interviews with five clinical experts.(33) The validated vignettes were valued using EQ-5D-5L and EQ-5D visual analogue scale (VAS) through an online survey of the UK general public (n=300).(33) Increasing severity of chronic GVHD was found to be associated with significantly lower utility values almost to extent of death in the most severe health state:(33)

[redacted]

It is important to recognise that utilities elicited amongst members of the general public are generally observed to be lower than patient-based valuations. A number of hypotheses have been put forward in the literature including scale recalibration(34), the ability of patients to contextualise health state descriptions better than members of the public(35), framing effects(36) and the extent to which the public may underestimate adaptation to negative changes in health.(36, 37) Studies have also shown that patients give more importance to the functional dimensions of disease than to symptoms(38) and conversely that relative changes in moving between health states can be larger in utilities elicited in healthy people.(39)

The results of our utility elicitation exercise demonstrate that members of the general public recognise that there is a very substantial burden associated with chronic GVHD after two or more prior lines of systemic therapy.(38) However, and possibly for the reasons provided above, the valuations obtained in the exercise were very much lower and slightly more widely spread than the directly observed data from the belumosudil studies and elsewhere (Section B.2.6.3.2). NICE takes a welfarist approach and states that HRQoL or changes in HRQoL should be measured directly by patients for the purposes of economic evaluation and so we have used the trial-based data (where available) in our modelling

(Section B.3.4). Nonetheless, in a publicly funded healthcare system, it is worthwhile to understand the view of members of the general public about the impact of the disease on HRQoL.

[redacted]

. For further details of the utility elicitation exercise, please refer to Appendix N.

Impact of the disease on caregivers and family members

Carers/families of patients with chronic GVHD are also impacted by the disease as they may have to work fewer hours, leave their job or retire earlier than planned to fulfil their caregiving requirements.(30) At the time of the multi-national observational survey, 52% of patients had a non-professional caregiver, meaning that the burden of care was likely placed on family members.(28) Physicians in the EU-4 and UK reported that the average caregiver of a patient with chronic GVHD who has progressed after two prior lines of therapy spends a mean of 35.2 hours (range: 2 to 160 hours) per week providing care for the patient.(31) Partners and spouses spend even more time per week on providing care (mean: 73.5 hours, range: 10 to 160 hours).(31) The main reasons why caregivers changed or reduced their work hours was the patient's depression/anxiety or loss of capability to complete daily tasks (e.g., washing, dressing).(31) Studies in the US have shown that caregivers of people who have received alloHSCT often have worse QoL compared to the general population.(40) Depression and sleep disorders are also more likely to occur in these caregivers.(40)

English clinicians referred to this US study when queried about caregiver QoL and further noted that the burden is likely higher for caregivers of patients who are not responding to treatment.(11) Patients who have received two or more prior therapies were reported to be unable to access clinical appointments or to remain compliant with their medication if they do not have the support of a caregiver.(11)

There is a scarcity of quantifiable data describing disutilities in caregivers of patients with chronic GVHD in England. However, given the substantial time burden placed on carers of patients with chronic GVHD who have progressed after two prior lines of therapy(28) as well as the serious and potentially life-threatening nature of the condition,(5, 6) it is anticipated that caregivers will report significant disutilities. There is no published study for chronic GVHD but evidence from multiple sclerosis (MS) is considered a relevant proxy by the clinicians we spoke to at the advisory board conducted in January 2023, due to its chronic, progressive and disabling nature as well as patients' need for daily assistance from a caregiver.(11) In a cross-sectional observational study conducted in the UK, caregivers of patients with MS (n=200) reported significantly lower HRQoL, as measured by the 36-item short form health survey (SF-36), Hospital Anxiety and Depression Scale (HADS) and EQ-5D, than matched controls (n=200).(41) Estimated disutility values for caregivers ranged from -

0.002 (normal/mild disability) to -0.173 (bilateral support[i]), depending on the severity of functional impairment experienced by the patient.(41) The disutilities reported by Acaster et al. 2013,(41) where patients in the failure-free state and patients experiencing failure were associated with caregiver disutilities of -0.045 and -0.142, respectively, were considered to be relevant proxies by the experts we spoke to (Section B.3.4.6).(11)

Some patients may also lose their support system as a result of the disease burden.(21) A US patient stated: "*Chronic GVHD has affected my loved ones tremendously. My caregiver left me. It was too heavy for him. That was traumatic for me, and traumatic for him as well*".(21)

Impact of current treatments

Currently available treatments for chronic GVHD place a further burden on patients as many are associated with high rates of toxicities and may require intravenous (IV) infusion (Section B.1.3.2).

Systemic corticosteroids (CS) are the main initial treatment for chronic GVHD;(11) however, prolonged CS use is associated with both acute and chronic health risks including osteoporosis and fractures, cardiovascular disease, hypertension and dyslipidaemia, impaired immune response and wound healing, susceptibility to viral and fungal infections, weight gain, cataracts, type 2 diabetes mellitus, psychiatric disturbances and GI events.(42) Chronic immune suppression and recurrent infections with CS limit patients' ability to go out in public and undertake activities of daily living.(20, 22) When describing the impact of being hospitalised, a US patient stated: "*I do worry about being hospitalised in the future, since I've spent over 126 nights in the hospital since my diagnosis, not counting the day trips and hours I've spent for infusions. I can't even count the days and hours that it adds up to. I now have claustrophobia. And understandably so. So, for me to be hospitalised is one of my biggest fears. So I worry about that all the time. I'm very careful about not hitting my arms, because I don't want to get infections. I'm very cautious about coming around sick people. So that's always in the back of my mind. I do not want to get hospitalised again*."(26) Another patient described being hospitalised over Christmas: "*I ended up being in hospital this past Christmas. And it was a perfect storm in the sense that I went in with a high temperature. They diagnosed me with influenza and then also with pneumonia. And because I had GVHD in my lungs, they put me in the intensive care unit (ICU) for a week. That was scary. Very scary*."(26)

GI events associated with CS, such as gastritis, ulcers, dyspepsia and abdominal distension, result in many patients requiring concomitant proton pump inhibitors (PPIs) or alternatives such as H2 antagonists (e.g. famotidine).(42) Factors that increased the need for concomitant PPI therapy include upper gut involvement and chronic thrombocytopaenia.(11) There is a clear need for treatments with a steroid sparing effect.

[i] Patient requires two canes or crutches, or a walker, to be able to walk 25 feet. They may use a scooter or wheelchair for longer distances.

The need for regular topical CS treatment also places a burden on patients, with a UK survey in atopic dermatitis demonstrating that application of topical CS has a substantial negative impact on HRQoL, which worsens with increasing frequency of treatment.(43)

The HRQoL of patients with chronic GVHD is negatively affected by increasing disease severity and lack of response to CS.(44) In an international cross-sectional survey of patient records and patient-completed forms (n=143) that included UK patient data, disease severity was the driving factor in overall symptom burden and number of symptoms.(44) EuroQol-5 Dimension-5 Level questionnaire (EQ-5D-5L) results were significantly lower in patients with severe disease (0.58), than in patients with moderate (0.69) or mild (0.82) disease.(44) EQ-5D-5L VAS scores were reported to be lower in patients who were steroid-refractory or steroid-dependent compared with steroid-responsive patients.(44) Of the steroid-refractory/dependent patients, 44.0% reported poor HRQoL, compared with 32.3% of steroid-responsive patients.(44) These outcomes demonstrate a need for a novel efficacious therapy. In the same study, mean EQ-5D crosswalk scores in UK patients reduced with each line of treatment, from 0.71 for patients receiving their first therapy to 0.60 for patients receiving their third or later therapy.(31)

An English expert clinician stated: "*Any patient that can be rescued from that living hell is worth trying everything on – to spend 10 years in this [moderate - severe chronic GVHD] state is torture*".(11)

B.1.3.1.5. Economic burden of chronic GVHD

Chronic GVHD is associated with significant disease management costs (medical resource use for management of the disease excluding any treatment costs or costs of treating treatment-related adverse events [AEs]).(45) Due to the unavailability of data on long-term disease management costs from the belumosudil clinical trials or real-world clinical practice in England, we conducted a study using secondary care data from the HES database to estimate disease management costs for the submission model (Section B.3.5.3). The HES database contains information on reimbursed diagnoses and procedures from all National Health Service (NHS) inpatient admissions, outpatient appointments and emergency care (EC) attendances in England.(45)

The HES study included data on patients aged ≥12 years with an alloHSCT between 1 April 2017 and 31 December 2020.(45) HES diagnosis data are limited to four-character International Classification of Diseases 10th Revision (ICD-10) codes; however, chronic GVHD is not identified through a definitive code.(45) To identify episodes of chronic GVHD, one of the following criteria had to apply:(45)

- Marker of GVHD, defined using ICD-10 code D89.8, ≥100 days after alloHSCT, or
- Marker of GVHD, defined using ICD-10 code D89.8, <100 days after alloHSCT and a subsequent code for a feature of chronic GVHD, where the chronic GVHD feature must have occurred after the marker for GVHD, or

- Marker of GVHD, defined using ICD-10 code T86.0 (at any time following alloHSCT) and a subsequent code for a feature of chronic GVHD, where the chronic GVHD feature must have occurred after the marker for GVHD.

For comparison, patients with chronic GVHD were matched to patients who had received alloHSCT but had no evidence of GVHD following the procedure, based on age, gender, time from alloHSCT and type of malignancy.(45) In total, 3,650 episodes of alloHSCT were recorded in patients aged ≥12 years, 821 (22.5%) of which had evidence of chronic GVHD, 987 (27.0%) had evidence of GVHD but not chronic GVHD, and 1,842 (50.5%) had no evidence of GVHD.(45) Matching criteria were applied, resulting in 721 episodes belonging to 721 unique patients with chronic GVHD and 718 unique patients without GVHD, three of whom were re-used as controls.(45)

Patients with chronic GVHD experienced more inpatient, outpatient, ICU and EC activity overall, compared with patients who did not have GVHD.(45) In total, 10,321 inpatient admissions were recorded for 721 patients with chronic GVHD (compared to 7,623 admissions among 718 patients without GVHD).(45) Patients with chronic GVHD experienced more inpatient admissions (74.6% vs. 66.6%, or 10 admissions per person-year vs. 6.3 admissions per person-year) with a similar mean length of stay (11.2 days vs. 11.3 days).(45) For ICU admissions, the length of stay was longer for patients with chronic GVHD compared to patients without chronic GVHD (8.6 days vs. 4.9 days).(45) Overall, 30,024 outpatient appointments were recorded for 579 patients with chronic GVHD compared with 18,835 appointments among 604 patients without GVHD, corresponding to an average of 13.5 more specialist outpatient appointments per person-year (29.0 appointments vs. 15.5, respectively).(45) Additionally, a greater proportion of patients with chronic GVHD had an EC attendance than those without chronic GVHD (39.3% vs. 30.5%).(45)

Overall, the cost of all-cause inpatient admissions (excluding ICU costs) was [redacted] among patients with chronic GVHD than patients without GVHD [redacted].(45) Amongst patients with at least one admission, the mean cost of inpatient admissions per person-year was [redacted] for those with chronic GVHD and [redacted] for those without GVHD.(45) The total cost of outpatient appointments was [redacted] in patients with chronic GVHD, and the cost of EC was [redacted] compared to patients without GVHD.(45) Patients with chronic GVHD incurred [redacted] in ICU admission costs compared with [redacted] incurred by patients without GVHD, with a mean cost per ICU episode of [redacted] and [redacted], respectively.(45)

When stratified by reported use of high-cost treatments (ECP, pentostatin, rituximab, ruxolitinib, imatinib), HCRU and costs were higher among patients who had received one or more high-cost therapies than in those who had not (Table 4).(45)

Table 4. Annual disease management costs by health states

Original category in HES	All non-GVHD patients	Chronic GVHD patients with no high-cost therapy	Chronic GVHD patients with first high-cost therapy	Chronic GVHD patients with at least two high-cost therapies
Mean cost of inpatient attendance per person-year	[redacted]	[redacted]	[redacted]	[redacted]
Mean cost of outpatient attendance per person-year	[redacted]	[redacted]	[redacted]	[redacted]
Mean cost of A&E attendances per person-year	[redacted]	[redacted]	[redacted]	[redacted]
Mean cost of ICU attendance per person-year	[redacted]	[redacted]	[redacted]	[redacted]
Mean total cost per person-year	[redacted]	[redacted]	[redacted]	[redacted]

A&E = accident and emergency; GVHD = Graft-Versus-Host Disease, HES = Hospital Episode Statistics; ICU = intensive care unit
SOURCE: Sanofi 2022(46)

Feedback from English clinicians indicates that there are significant capacity issues for some facilities as the use of IV infusion therapies places pressures on service delivery across diseases.(11) Due to the necessary centralisation of services, English clinicians considered ECP to have significant limitations.(11) Patients who lack carers or financial resources require support to access some treatments, including ECP, which are location-specific and require transportation and day unit infusion chair time.(11) English clinicians considered chronic GVHD to be associated with a large but difficult-to-define resource usage that required a holistic multi-disciplinary approach to address the pan-system consequences that result from multi-organ involvement.(11) Experts pointed out that there is an unmet need for NICE technology appraisals, early access programmes and compassionate access programmes for novel agents to treat chronic GVHD.(11)

B.1.3.2. Description of clinical pathway of care

B.1.3.2.1. Treatment pathway for chronic GVHD in England

Guidelines on the treatment of chronic GVHD in the UK were last published by the British Society for Haematology (BSH) and BSBMTCT in 2012 and are presented in Table 5.(47) The guidelines noted that there was a paucity of evidence supporting the majority of treatment options for chronic GVHD.(47)

Table 5. UK guidelines on the management of chronic GVHD (2012)

First-line therapy
• For patients with mild disease, topical treatments and supportive agents may be sufficient • CS are recommended in the first-line treatment of chronic GVHD • An initial starting dose of 1 mg/kg prednisolone is recommended • CNIs may be helpful in the initial treatment of GVHD as a steroid-sparer
Second-line therapy/steroid-resistant or refractory chronic GVHD
• ECP may be considered as a second-line treatment in skin, oral or liver chronic GVHD • mTOR inhibitors are suggested as a second-line treatment option in refractory chronic GVHD • Pentostatin is suggested as a second-line treatment option in refractory chronic GVHD

• Rituximab is suggested as a second-line treatment option in refractory cutaneous or musculoskeletal chronic GVHD • Imatinib is suggested as a second-line treatment option in refractory pulmonary or sclerodermatous chronic GVHD
Third-line therapy
• ECP, imatinib and rituximab may be considered as third-line treatment options in chronic GVHD involving other organs The following agents are suggested as third-line treatment options in refractory chronic GVHD: • MMF • MTX • Pulsed CS
Additional recommendations
• Azathioprine is not recommended due to risk of oral malignancy

CNI = calcineurin inhibitor; CS = corticosteroids; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; MMF = mycophenolate mofetil; mTOR = mammalian (mechanistic) target of rapamycin; MTX = methotrexate; UK = United Kingdom
SOURCE: Dignan 2012(47)

A 2017 Clinical Commissioning Policy issued by NHS England was developed to define a treatment pathway for the clinical management of GVHD in England, and to outline funding arrangements in this population (Table 6); it was noted that the evidence used to inform the proposal was limited and of varying quality.(48)

Table 6. NHS England Clinical Commissioning Policy on the management of chronic GVHD (2017)

First-line (1L) therapy: • CS with an initial starting dose of 1 mg/kg prednisolone
If at risk of developing adverse effects or becoming CS dependent: • Calcineurin inhibitors (tacrolimus or cyclosporine) are indicated to reduce dose of systemic steroids
If no complete response (i.e. steroid-refractory chronic GVHD), significant adverse effects to first-line treatments or steroid-dependent: • Sirolimus
"Second-line" (2L) therapy/steroid-resistant or refractory chronic GVHD
The following treatments are proposed to be added as second-line options (by organ/indication): • Refractory chronic GVHD: Pentostatin (1.5 mg/m^2) • Skin, oral, liver and pulmonary chronic GVHD: ECP (only second-line treatment of choice for skin, oral, liver and pulmonary chronic GVHD)[a] • Refractory cutaneous or musculoskeletal chronic GVHD: Rituximab • Refractory pulmonary or sclerodermatous chronic GVHD: Imatinib
"Third-line" (3L) therapy
Where patients show incomplete response to **two different second-line options**[b] and/or have developed significant adverse effects, the following treatments are indicated third-line: • MMF • Methotrexate • Pulsed CS

CS = corticosteroids; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; MMF = mycophenolate mofetil
[a]ECP is not readily available to all transplant centres and patients are often too sick to travel.
[b]Patients at this point in their treatment journey may have received several lines of systemic therapy.
SOURCE: NHS England 2017(48)

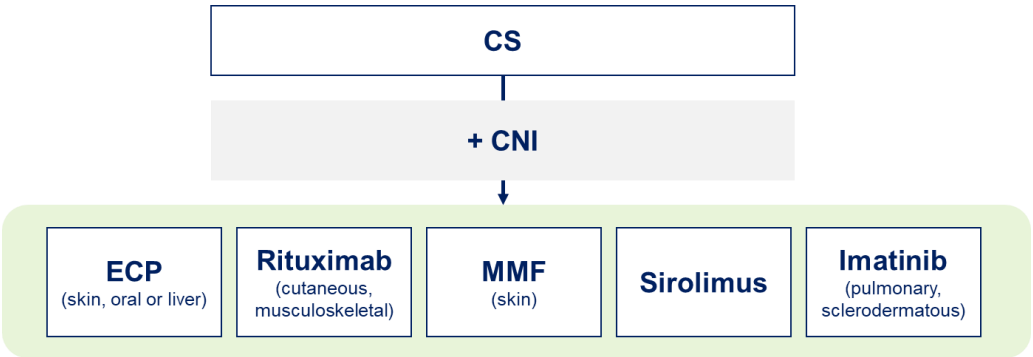
When discussing treatment options for chronic GVHD, it is critical to understand the nuances of the real-world treatment pathway and to appreciate that 'prior lines of therapy' and 'numbers of prior treatments' are not synonymous. Lines of treatment can be both individual treatments and combinations.

Both the 2012 BSH/BSBMTCT guidelines and the 2017 NHS England Commissioning Policy state that, in patients who fail one 'second-line' therapy, another 'second-line' agent should be used before moving to 'third-line' options.(47, 48) In this case, the patient would still be considered to be receiving second-line therapy, despite having received a third (or fourth) systemic treatment.(47, 48) In this sense, the 'second-line' treatment options listed by the BSH/BSBMTCT guidelines and NHS England Commissioning Policy may, therefore, be used in patients who have received two or more prior lines of systemic therapy.(47, 48) This is in line with a 2015 consensus paper published by the NIH, which defines the first-line of treatment as the beginning of systemic treatment for chronic GVHD, while subsequent lines of treatment are characterised by the introduction of any new systematic agent that was not previously used.(49) This could include, for example, the addition of CNIs to corticosteroids which would be defined as a second-line treatment. The belumosudil trials are also aligned with this definition of a line of therapy. In ROCKstar, a treatment was considered another line of therapy if it was added after 4 or more weeks.(50)

Given the lack of recent specific treatment guidelines for England, we conducted an advisory board in January 2023 with clinical and health economic experts to understand the current treatment pathway for chronic GVHD in England. Feedback from English clinical experts indicates that CS remain the preferred initial treatment for patients with chronic GVHD, with or without the addition of CNIs (e.g., tacrolimus or cyclosporin; Figure 4). After two prior lines of therapy, ECP is currently the most commonly used treatment (approximately 65% of patients), with a small number of patients receiving sirolimus, imatinib or organ-specific treatments such as MMF and rituximab (Figure 4).(11)

English clinical experts highlighted that there is currently a lack of suitable options for patients who have received two prior lines of therapy.(11) Patients often experience poor outcomes on current treatments and generally would not stay on a single treatment for a long time in this disease stage.(11) English clinicians stated that tapering patients off immunosuppressive treatment following a response was an important therapy goal, with all clinicians stating they would take responders off a chronic GVHD treatment with careful monitoring.(11)

Figure 4. Treatment pathway for patients with chronic GVHD in England



CNI = calcineurin inhibitors; CS = corticosteroids; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; MMF = mycophenolate mofetil
SOURCE: Adapted from Dignan 2012(47) and Sanofi 2023(11)

B.1.3.2.2. Current treatments for patients with chronic GVHD who have received at least two prior lines of systemic therapy

Currently available treatments are prescribed according to their organ-specific benefits; however, chronic GVHD is a complex and heterogeneous disease that affects multiple organs.(47) Due to the heterogenous nature of the disease and the prescription of different medicines according to manifestation as well as access to treatment (e.g., proximity to ECP treatment centres), it is appropriate to consider best available therapy (BAT) in the form of a basket of therapies as the overarching standard of care.

An overview of current treatments included in the BAT basket for patients with chronic GVHD in England who have received at least two prior lines of systemic therapy is presented in Table 7, with additional details provided in the following sections. There is a paucity of robust, randomised controlled trial evidence for the treatments used and reimbursed in England for the treatment of chronic GVHD, and the majority of treatments are used off-label.(47)

Table 7. Overview of current treatments for patients with chronic GVHD in England who have received at least two prior lines of systemic therapy

Treatment	Mode of administration	Indication	Position in pathway
ECP (47, 51, 52)	UVA irradiation of patient's blood via peripheral- or central-venous catheter 3-4 hours on 2 consecutive days, every two weeks, for a minimum of 3 months, requiring 2 hours of specialist nurse time per administration	Indication not strictly defined, most commonly prescribed after two prior lines of therapy in chronic GVHD	ECP may be considered as a preferred treatment for skin, oral or liver chronic GVHD following CS ± CNI
Rituximab (47, 53)	IV; administered once weekly for 4 weeks	Off-label use – Not indicated for chronic GVHD in SmPC	Rituximab may be considered as a preferred treatment for cutaneous or musculoskeletal chronic GVHD following CS ± CNI
Imatinib (47, 54)	Oral capsules	Off-label use - Not indicated for chronic GVHD in SmPC	Imatinib may be considered as a preferred treatment for refractory pulmonary or sclerodermatous chronic GVHD following CS ± CNI
Sirolimus (47, 55)	Oral	Off-label use - Not indicated for chronic GVHD in SmPC	Sirolimus may be considered as a preferred treatment for refractory chronic GVHD following CS ± CNI
Mycophenolate mofetil (47, 56)	Oral	Off-label use - Not indicated for chronic GVHD in SmPC	Mycophenolate mofetil may be considered as a preferred treatment for chronic GVHD affecting the skin following CS ± CNI

CNI = calcineurin inhibitors; CS = corticosteroid; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; IV = intravenous; SmPC = summary of product characteristics; UVA = ultraviolet A

Ruxolitinib (a Janus kinase [JAK] inhibitor) recently received approval from the European Commission for the treatment of patients aged 12 years and older with acute or chronic GVHD who have inadequate response to CS or other systemic therapies, based on the randomised open-label multicentre phase 3 REACH3 trial.(57, 58) Up until April 2022, ruxolitinib was reimbursed by NHS

England through an interim rapid commissioning policy in the context of the COVID-19 pandemic.(59) However, use in new patients is no longer permitted due to the withdrawal of the interim commissioning policy.(27) There is currently no NICE recommendation for ruxolitinib in GVHD.(27)

Extracorporeal photopheresis (ECP)

ECP is currently the most common therapy for patients who have received two or more prior lines of systemic therapy for chronic GVHD in England.(11) However, ECP is associated with a high burden for patients, caregivers and NHS England.(51)

ECP can take 3-4 hours per treatment for the patient and requires a time commitment of two consecutive days, every 2 weeks, disrupting the daily lives of patients and their caregivers.(51) It can take over six months before any improvement, and there is no mandated discontinuation timeframe.(60) In some cases treatment can last several years or treatment may be restarted in responders who subsequently relapse.

The clinical efficacy of ECP varies across different organ manifestations.(11, 61) The efficacy of ECP is primarily relevant to skin, oral and liver chronic GVHD (which account for manifestations in approximately 60% of patients(15)); response rates in other manifestations are considerably lower and more variable, with limited data to support ECP's use in these indications.(15)

ECP is typically performed via peripheral venous access catheter; however, when this is not possible, a central-venous catheter must be used. Feedback from English clinicians indicates that approximately 25% to 30% of patients receiving ECP in England require central-venous access,(11) for which line placement can take an additional 1 hour on top of the ECP procedure time. Central-venous catheters are associated with an increased risk of infection,(47) particularly in immunosuppressed chronic GVHD patients, with an estimated 20% of patients developing an infection.(11) Line infection (septicaemia) requires a hospital stay of at least 5-10 days and the administration of IV antibiotics.(11)

Access to ECP is limited to five NHS Blood and Transplant (NHSBT) Therapeutic Apheresis Units in Bristol, Oxford, Manchester, Liverpool and London,(62) and a limited number of hospital trusts providing ECP services independently.(11) According to the 2017 commissioning policy, patients eligible for ECP are often too sick to travel to the centres which provide it.(48) Patients who do not have a caregiver to help them travel to one of the centres may not be able to access treatment.(11) An additional disadvantage for caregivers of patients who reside far from the nearest ECP centre is the necessity to take 2 consecutive days off work to travel with the patient. The extended travel time places an additional burden on the patient and carer and may prevent patients from accepting ECP treatment.(11) Socioeconomic challenges may also limit treatment access as patients living too close to the nearest centre to qualify for overnight accommodation reimbursement may not be able to continuously afford the transportation and accommodation, or may have to decline this form of treatment altogether.(11) While ECP centres offer hospital transport, many patients will not take up

the offer and describe the experience as ‘horrendous’ and choose to make their own way to the centres at their own expense.(11) Use of public transport is also a concern for many patients due to the heightened risk of contracting infections.(11)

The disutility associated with ECP for each patient is expected to be notable. In the absence of utility assessments for chronic GVHD therapies, a study (Matza 2013) assessing utilities of injection vs. infusion treatments for patients with bone metastases can be considered a relevant proxy.(63) In Matza 2013, a 2-hour infusion treatment was associated with a disutility of −0.037 (SD: 0.106) for each treatment modality.(63) This study was previously used by NICE to approximate the impact of infusion therapies in the evaluation report for migalastat for treating Fabry disease.(64)

ECP is also an expensive procedure (£1,585 per session) as it requires quality-assured sites with validated machines (e.g., THERAKOS™ CELLEX™ Photopheresis System), specially trained staff and 2 hours of specialist nurse time (£110 per administration).(47, 52, 65) Additionally, overnight accommodation for patients receiving ECP are reimbursed by the NHS at £150 per night.(66) Based on feedback from expert clinicians in England, it is estimated that 50% of patients receiving ECP require an overnight stay.(11)

Other treatment options

Other treatments that are used in patients with chronic GVHD who have received two prior systemic therapies in England and will be included in the BAT basket for this submission are rituximab, imatinib, sirolimus and MMF.

Rituximab places a substantial burden on patients and secondary care due to the occurrence of IV catheter-related infections, high rates of concomitant medication use (e.g., CS, CNIs) and limited accessibility due to the required hospital stay.(48, 67) Additionally, rituximab is only recommended by clinical guidelines for the treatment of refractory cutaneous or musculoskeletal manifestations of chronic GVHD.(47) Similarly, imatinib is only suggested as a treatment option for refractory pulmonary or sclerodermatous chronic GVHD, which leaves patients with other manifestations without a suitable treatment option.(47)

Sirolimus is intended as a combination therapy with CS and should not be used in combination with CNIs; additionally, patients have to be monitored for hyperlipidaemia.(47) The other current treatment option, MMF, is associated with significant (including life-threatening) AEs, most common of which are infections and cytopenia.(47)

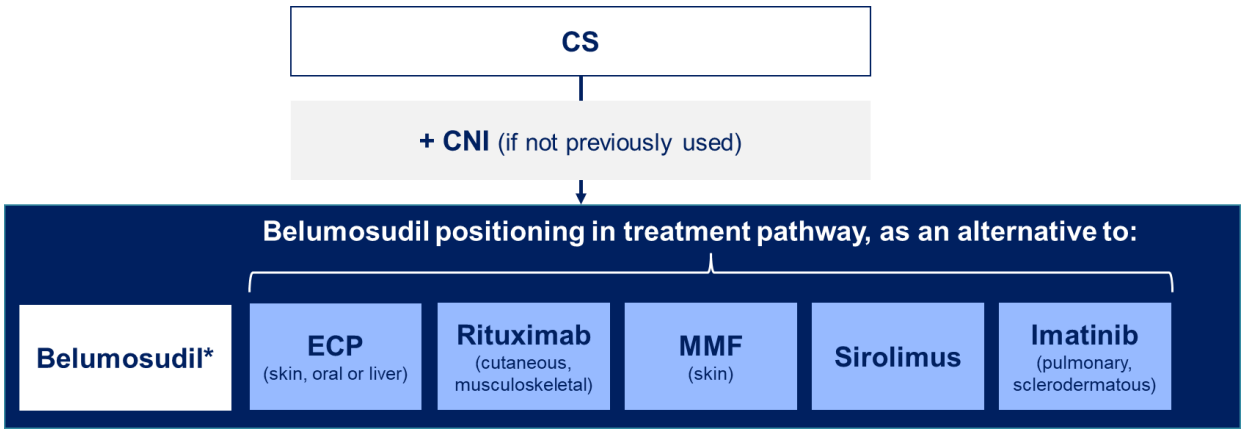
B.1.3.2.3. Unmet need

Chronic GVHD is an immune-mediated inflammatory and fibrotic multi-system disease that is associated with significant morbidity and mortality, ultimately compromising the clinical and HRQoL benefits offered by alloHSCT.(5, 6, 17, 68)

Currently available treatments are prescribed according to their organ-specific benefits and rely on immunosuppression only, without specifically targeting profibrotic and inflammatory processes, leading to rapid disease progression.(47, 69) However, GVHD is a complex, heterogenous disease affecting multiple organs. There is a need for treatments that target the underlying cause of the disease (both inflammatory and profibrotic pathways) and are not organ-specific. There is a paucity of clinical evidence supporting the treatment options currently available to patients with chronic GVHD in England, including several off-label therapies being used without strong evidence bases to demonstrate proven efficacy.(47, 70) Additionally, current treatments for patients in England are associated with significant limitations, including catheter-related infections, limited accessibility, high cost and significant (potentially life-threatening) AEs such as infections.(47, 48, 62, 67) Consequently, patients with chronic GVHD who have received at least two prior lines of systemic therapy require urgent access to a convenient and effective treatment with a positive risk-benefit profile. This treatment should target the underlying pathophysiologic drivers of the disease and be associated with a relatively low risk of toxicities and infections, allowing patients to continue their daily lives with a lower burden of disease. Belumosudil meets these requirements and provides an additional option for this small but highly burdened population of patients.

The optimal place in therapy for belumosudil, intended to be used as a monotherapy, in the chronic GVHD treatment pathway is after oral CS (with or without the addition of CNIs) and at least one other systemic therapy, such as sirolimus or the later addition of a CNI (Figure 5). This positions belumosudil, within its licensed indication, as an alternative to whichever treatment the prescribing healthcare professional selects from the full list of best available therapies discussed above.

Figure 5. Belumosudil place in therapy



CNI = calcineurin inhibitors; CS = corticosteroids; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; MMF = mycophenolate mofetil
*Only after two systemic treatments
SOURCE: Adapted from Dignan 2012(47) and Sanofi 2023(11)

B.1.4. Equality considerations

Currently, only 37% of patients from a minority ethnic background can find the best possible stem cell match from a stranger (compared to 72% of patients from White backgrounds).(71) Mismatched, unrelated donors are consistently reported as a risk factor for chronic GVHD following alloHSCT.(11) It is plausible therefore that chronic GVHD is more likely to occur in people from a minority ethnic background.(11) This submission does not discriminate against these groups, however it is worth noting that the current lack of licensed, reimbursed, effective treatments with a favourable safety profile in itself may disadvantage these populations.

Skin manifestations are some of the most common and major complications of chronic GVHD, and dermatological assessment is required for disease diagnosis and severity grading, yet current physician- and patient-reported outcome measures may not adequately capture the subtle changes in patients with non-white skin, leading to potential errors or delays in diagnosis for such patients.(72)

Geographical access to ECP services and specialist blood and marrow transplant clinics can be a barrier to people in lower socioeconomic groups who may be unable to take time off work or afford to travel to appointments.(11) Access to ECP can be particularly challenging given the need for two 3-4-hour procedures on consecutive days every fortnight or every month.(51) English clinicians consulted in our advisory board reported that frequent travel to specialised centres becomes too great of a burden for some patients.(11) Patients with lower socioeconomic status may have to decline ECP if they fall outside of the travel distance requirements that would grant them free accommodation between the two therapy days.(11) Inequity in access to specialists for organ-specific chronic GVHD (e.g., respiratory, dermatology, oral medicine, gynaecology, ophthalmology) as well as access to psychological support for patients and families was also flagged by English clinicians.(11) Minority ethnic patients may also experience increased stigma around skin changes (hyper- or hypopigmentation, chronic skin shedding) and bowel urgency, which may lead to a greater degree of social isolation and an increased psychological impact.(11)

Having the option of an oral, at home treatment alternative could be particularly beneficial in these groups.(11) Therefore, a NICE recommendation in this therapy area could have a positive impact on people protected by the equality legislation.

B.2. Clinical effectiveness

B.2.1. Identification and selection of relevant studies

An SLR was conducted with a cut-off date of January 2023 to identify studies reporting on the clinical efficacy and safety of treatment options for adult patients with chronic GVHD after alloHSCT who have failed at least one prior line of therapy (Appendix D). The included population scope for the SLR (patients failing at least one prior line of therapy) was broader than the population of interest for the submission (patients who had received at least two prior lines of therapy) to ensure that no studies reporting data on the population of interest were missed as the definition of prior therapies and how it is reported can be variable in the scientific literature.

A total of 670 records were identified through the database searches, 563 of which were unique titles, and 3 conference proceedings were identified from the grey literature search for a total of 38 publications reporting on 26 unique trials. Of the 26 trials, two investigated belumosudil (ROCKstar and Jagasia 2021 [Phase 2a]) and are described in detail below.

B.2.2. List of relevant clinical effectiveness evidence

Two single-arm clinical trials, the randomised, open-label Phase 2 ROCKstar trial (KD025-213; NCT03640481) and the dose-finding, open-label Phase 2a KD025-208 trial (NCT02841995), support the use of belumosudil for the treatment of patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy.(73, 74)

A summary of the ROCKstar and Phase 2a trials is provided in Table 8.

Table 8. Clinical effectiveness evidence

Study	**Phase 2a (KD025-208; NCT02841995)(74)**	**ROCKstar (KD025-213; NCT03640481)(73)**
Study design	Open-label, dose-escalation, multicentre study	Open-label, randomised, multicentre study
Population	Patients ≥18 years who had received allogeneic bone marrow transplant or alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 1 to 3 prior lines of systemic treatment	Patients ≥12 years who had received alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 2 to 5 prior lines of systemic treatment
Intervention(s)	Belumosudil 200 mg once daily, 200 mg twice daily or 400 mg once daily	Belumosudil 200 mg once daily or 200 mg twice daily
Comparator(s)	None	None
Indicate if study supports application for marketing authorisation	Yes	Yes

Study	Phase 2a (KD025-208; NCT02841995)(74)	ROCKstar (KD025-213; NCT03640481)(73)
Indicate if study used in the economic model	Yes	Yes
Rationale if study not used in model	N/A	N/A
Reported outcomes specified in the decision problem	Primary endpoint: • **Best ORR at any time (proportion of patients who achieved CR or PR)** Secondary endpoints: • **Failure-free survival** • Number and percentage of patients with steroid-dependent chronic GVHD with a best response of PR or CR • **Duration of response** • Response by organ system • Changes in symptom burden/bother using the LSS score • Changes in corticosteroid dose • **Overall survival** • Safety and tolerability of belumosudil	Primary endpoint: • **Best ORR at any time (proportion of patients who achieved CR or PR)** Secondary endpoints: • **Failure-free survival** • **Duration of response** • **Time to response*** • Response by organ • Changes in symptom burden/bother using the LSS score • **Overall survival** • Changes in corticosteroid dose • Changes in calcineurin inhibitor dose • Safety and tolerability of belumosudil Exploratory endpoint: • Changes in the PROMIS Global Health sub-scores for physical and mental functioning • Overall Response Rate using Kadmon algorithmic response assessment (KARA)
All other reported outcomes	Secondary endpoints: • Time to next treatment • Pharmacokinetics of belumosudil Exploratory endpoints: Pharmacodynamics of belumosudil	Secondary endpoints: • Time to next treatment • Changes in NIH chronic GVHD global severity rating • Changes in symptom activity using the chronic GVHD Activity Assessment Patient Self-Report • Pharmacokinetics of belumosudil Exploratory endpoints: • Changes in relevant biomarkers after belumosudil administration

AlloHSCT = allogeneic haematopoietic stem cell transplant; CR = complete response; GVHD = graft-versus-host disease; LSS = Lee Symptom Scale; N/A = not applicable; NIH = National Institute of Health; ORR = objective response rate; PR = partial response; PROMIS = Patient-Reported Outcomes Measurement Information System; TTD = time to treatment discontinuation; TTR = time to response
*While TTR and TTD were not trial endpoints they were derived from the pooled Phase 2 belumosudil studies for the purpose of the economic analysis.
SOURCES: Adapted from Cutler et al. 2021,(73) Jagasia et al. 2021,(74)

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

B.2.3.1. Study methodology

A summary of the study designs and methodology of the ROCKstar (KD025-213; NCT03640481) and Phase 2a (KD025-208; NCT02841995) trials is presented in Table 9.

Table 9. Comparative summary of trial methodology

Trial number (acronym)	Phase 2a (KD025-208; NCT02841995)(74, 75)	ROCKstar (KD025-213; NCT03640481)(73)
Location	US	US
Trial design	Phase 2a, dose-finding, open-label study	Phase 2, randomised, multicentre study; primary analysis at 6 months and follow-up analysis at 12 months
Eligibility criteria for participants	Patients ≥18 years who had received allogeneic bone marrow transplant or alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 1 to 3 prior lines of treatment **Key inclusion criteria:** • Had undergone alloHSCT and had persistent active chronic GVHD, as defined by the 2014 NIH Consensus Development Project on Criteria for Clinical Trials in chronic GVHD, after at least 2 months of steroid therapy • Were receiving glucocorticoid therapy and calcineurin therapy or glucocorticoid therapy alone for chronic GVHD at study entry. Other therapies not considered to be immunosuppressive, such as ECP, were allowed on a case-by-case basis • Had received no more than 3 prior lines of treatment for chronic GVHD • Karnofsky Performance Scale >40 • Had adequate safety laboratory values including total bilirubin ≤1.5 × ULN, ALT and AST ≤3 × ULN, and GFR ≥30mL/min/1.73m^2 **Key exclusion criteria:** • Acute GVHD • On investigational GVHD treatment within 28 days of study entry and having acute GVHD • History of other severe illness including alcoholism, HIV, active HCV or HBV, relapse of underlying cancer or diagnosis with other malignancy within 3 years of enrolment (with the exception of basal cell and squamous cell carcinoma of the skin, resected in situ cervical malignancy, resected breast ductal carcinoma in situ and low-risk prostate cancer after curative resection)	Patients ≥12 years who had received alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 2 to 5 prior lines of treatment **Key inclusion criteria:** • Have undergone alloHSCT • Have previously received ≥2 and ≤5 lines of systemic therapy for chronic GVHD • Have received glucocorticoid therapy with a stable dose over the 2 weeks before screening • Karnofsky or Lansky Performance Status Scale score ≥60 • Have persistent chronic GVHD manifestations requiring systemic therapy • Have adequate safety laboratory values including total bilirubin ≤1.5 × ULN, ALT and AST ≤3 × ULN, and GFR ≥30mL/min/1.73m^2 **Key exclusion criteria:** • Not being on a stable dose/regimen of systemic chronic GVHD treatments for at least 2 weeks before screening (investigational GVHD treatments excluded) • Histological relapse of underlying cancer or post-transplant lymphoproliferative disease at the time of screening • FEV_1≤ 39% (or lung score of 3) • Receiving treatment with ibrutinib within 28 days before randomisation
Settings and locations where the data were collected	7 secondary care centres across the US	33 secondary care centres across the US
Trial drugs	Subjects were enrolled (1:1:1)† into 3 sequential cohorts: • Belumosudil 200 mg orally once daily (n=17) • Belumosudil 200 mg orally twice daily (n=16)	Subjects were randomly assigned (1:1) to 1 of 2 treatment arms: • Belumosudil 200 mg orally once daily (n=66) • Belumosudil 200 mg orally twice daily (n=66)

Trial number (acronym)	Phase 2a (KD025-208; NCT02841995)(74, 75)	ROCKstar (KD025-213; NCT03640481)(73)
	• Belumosudil 400 mg orally once daily (n=21)	
Permitted and disallowed concomitant medication	The use of CYP3A4 inhibitor/inducers and/or drugs known to prolong the QT/QTc interval was prohibited. Taking immunosuppressant drugs for GVHD, including mTOR inhibitors, was also prohibited.	Concomitant corticosteroids, calcineurin inhibitors, sirolimus, mycophenolate mofetil, methotrexate, rituximab and extracorporeal photopheresis were allowed. Systemic investigational GVHD treatments were not permitted.
Primary outcomes (including scoring methods and timings of assessments)	Best ORR at any time, defined as the proportion of subjects who achieved CR or PR according to the 2014 NIH Consensus Criteria as assessed by investigators	Best ORR at any time, defined as the proportion of subjects who achieved CR or PR according to the 2014 NIH Consensus Criteria as assessed by investigators
Other outcomes used in the economic model/specified in the scope	FFS, number and the percentage of patients with steroid-dependent chronic GVHD who had a best response of PR or CR, DOR, response by organ system, changes in LSS summary score, CS dose reductions, OS	FFS, DOR, TTR, response by organ, changes in LSS summary score, CS and CNI dose reductions, OS, safety
Pre-planned subgroups	None	Severe chronic GVHD at screening (yes/no) Duration of chronic GVHD prior to enrolment (>50th percentile; ≤50th percentile) Number of organs involved at baseline (≥4; <4) Number of prior systemic LOTs (≥4; <4) Prior ibrutinib (yes/no) Receiving concomitant PPI* on C1D1 (yes/no)
Post-hoc subgroups	Belumosudil dose (200 mg once daily, 200 mg twice a day, 400 mg once daily) Refractory to prior line (yes; no) Prior lines (≥ 2; 1) Chronic GVHD severity at baseline (severe, non-severe) Number of organs involved at baseline (≥ 4; <4)	Best response to last systemic LOT Prior ruxolitinib (yes/no)

AlloHSCT = allogenic haematopoietic stem cell transplant; ALT = alanine aminotransferase; AST = aspartate aminotransferase; C1D1 = cycle one, day one; CI = confidence interval; CNI = calcineurin inhibitor; CR = complete response; CS = corticosteroid; DOR = duration of response; ECP = extracorporeal photopheresis; FEV_1 = forced expiratory volume in 1 second; FFS = failure-free survival, GFR = glomerular filtration rate; GVHD = graft-vs-host disease; HBV = Hepatitis B Virus; HCV = Hepatitis C Virus; HIV = Human Immunodeficiency Virus; LOT = line of therapy; LSS = Lee Symptom Scale; n/a = not available; mTOR = mammalian target of rapamycin; NIH = National Institutes of Health; ORR = overall response rate; OS = overall survival; PPI = proton pump inhibitor; PR = partial response; TTR = time to treatment response, ULN = upper limit of normal; US = United States

SOURCES: Adapted from Cutler et al. 2021,(73) Jagasia et al. 2021,(74)

* The subgroup analysis of patients receiving PPI on C1D1 was added in a protocol amendment as a result of a pharmacokinetic bioavailability study. The PK study showed that the assumed bioavailability of 1 in a fasted healthy individual was reduced by 48% for healthy individuals or patients with chronic GVHD who received concomitant PPIs and that absorption was delayed.(76) Further analyses showed that the maximal concentration (C_{max}) and area under the curve (AUC) were reduced by 87% and 80%, respectively, for patients treated with a strong PPI and reduced by 68% and 47%, respectively for patients treated with a weaker PPI.(76) However, due to the flat exposure-efficacy relationship between belumosudil and PPIs across the evaluated exposure range, no dose adjustment was needed for the administration of belumosudil in the study.(76)

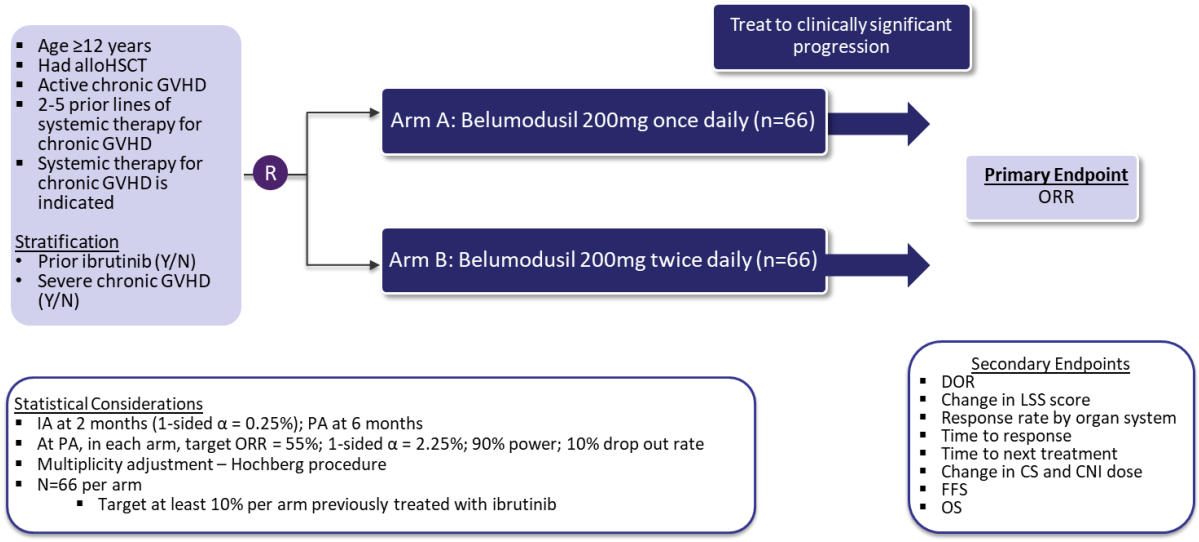
†16 patients were planned for each of the cohorts

B.2.3.1.1. ROCKstar (KD025-213; NCT03640481)

The ROCKstar trial enrolled patients aged ≥12 years who had received alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 2 to 5 prior lines of treatment.(73) The ROCKstar population was representative of patients with advanced disease who may not have many treatment options left available to them in the current treatment pathway in England.(73) English clinicians consulted at our advisory board confirmed that the ROCKstar trial population is generalisable to patients with chronic GVHD in England, with the distinction that prior therapies in the trial are not fully aligned with standard treatment practices in England as patients have access to a more limited selection of treatment options.(11) These differences are discussed further in Section B.2.5.2.

The study design is illustrated in Figure 6, with additional detail provided in Appendix D.2.

Figure 6. Study design for ROCKstar (KD025-213; NCT03640481)



alloHSCT = allogeneic haematopoietic stem cell transplant; CNI = calcineurin inhibitor; CS = corticosteroid; cGVHD = chronic graft-versus-host disease; DOR = duration of response; GVHD = graft-versus-host disease; FFS = failure-free survival; IA = interim analysis; LSS = Lee Symptom Scale; ORR = overall response rate; OS = overall survival; PA = primary analysis; R = randomisation; Y/N = yes/no
SOURCE: Adapted from Cutler et al. 2021(73)

It is critical to note that in the ROCKstar trial, a line of therapy (LOT) was defined as a regimen of systemic therapies indicated for the treatment of chronic GVHD (post diagnosis of chronic GVHD).(77) A LOT may have included more than one drug if started at the same time or within 4 weeks (for example systemic corticosteroids [CS] and calcineurin inhibitors [CNI]).(77) In general, when a new systemic therapy was added to the treatment that was not by itself effective, it was defined as a new line, this includes the addition of CNI to CS after 4 weeks on CS monotherapy.(69, 77) Topical treatments were not considered systemic chronic GVHD therapies and therefore did not contribute to prior lines of systemic therapy in ROCKstar.(77)

B.2.3.1.2. Phase 2a study (KD025-208; NCT02841995)

The Phase 2a trial enrolled patients aged ≥18 years who had received alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 1 to 3 prior lines of treatment. LOTs were defined in the same way as the ROCKstar study. Additional detail on the study design is provided in Appendix D.2.

B.2.3.2. Baseline characteristics

B.2.3.2.1. ROCKstar (KD025-213; NCT03640481)

The baseline characteristics for ROCKstar are presented in Table 10. Information about patient disposition is provided in Appendix D.2.

Table 10. Baseline characteristics for ROCKstar

Baseline characteristic	200 mg once daily (n=66)	200 mg twice daily (n=66)	Total (N=132)
Median age (range), years	53 (21—77)	57 (21—77)	56 (21—77)
Males, n (%)	42 (64%)	33 (50%)	75 (57%)
HLA matching of donor/recipient, n (%)			
Matched	57 (86%)	62 (94%)	119 (90%)
Partially matched	8 (12%)	3 (5%)	11 (8%)
Unknown	0	1 (2%)	1 (1%)
Missing	1 (2%)	0	1 (1%)
Time from chronic GVHD diagnosis to enrolment, median (range), months	25 (2—162)	30 (4—144)	29 (2—162)
NIH chronic GVHD severity[a] n (%)			
Severe	46 (70%)	43 (65%)	89 (67%)
Moderate	18 (27%)	23 (35%)	41 (31%)
Mild	2 (3%)	0	2 (2%)
Organ involvement, n (%)			
No. of organs involved, median (range)	4 (0—7)	4 (2—7)	4 (0—7)
≥4 organs involved	33 (50%)	35 (53%)	68 (52%)
Skin	55 (83%)	55 (83%)	110 (83%)
Joints/fascia	51 (77%)	49 (74%)	100 (76%)
Eyes	48 (73%)	49 (74%)	97 (74%)
Mouth	30 (46%)	42 (64%)	72 (55%)
Lungs	24 (36%)	23 (35%)	47 (36%)
Oesophagus	19 (29%)	12 (18%)	31 (24%)
Upper GI	13 (20%)	10 (15%)	23 (17%)
Lower GI	6 (9%)	7 (11%)	13 (10%)
Liver	9 (14%)	4 (6%)	13 (10%)
Prior therapy characteristics, n (%)			
Median prior LOTs, n	3	4	3
≥4 prior LOTs	30 (45%)	35 (53%)	65 (49%)
≥6 prior LOTs	1 (2%)	2 (3%)	3 (2%)
Refractory to prior LOT	44 (79%)	35 (65%)	79 (72%)
Concomitant systemic chronic GVHD therapies[b], n (%)			
CS	65 (99%)	66 (100%)	131 (99%)

Baseline characteristic	200 mg once daily (n=66)	200 mg twice daily (n=66)	Total (N=132)
CNI	24 (36%)	25 (38%)	49 (37%)
ECP	17 (26%)	22 (33%)	39 (30%)
Sirolimus	17 (26%)	18 (27%)	35 (27%)
MMF	11 (17%)	2 (3%)	13 (10%)
Imatinib	1 (2%)	1 (2%)	2 (2%)
Rituximab	1 (2%)	0	1 (1%)
Ruxolitinib	1 (2%)	0	1 (1%)
Other	9 (14%)	13 (20%)	22 (17%)

[a] Severity was determined using the NIH Global Severity of chronic GVHD scoring
[b] Classified as concomitant systemic chronic GVHD medications on Cycle 1 Day 1 (at baseline)
CNI = calcineurin inhibitor; CS = corticosteroids; ECP = extracorporeal photopheresis; GI = gastrointestinal tract; GVHD = graft-versus-host disease; HLA = human leukocyte antigen; LOT = line of therapy; MMF = mycophenolate mofetil; NIH = National Institutes of Health
SOURCE: Adapted from Cutler et al. 2021(73)

B.2.3.2.2. Phase 2a study (KD025-208; NCT02841995)

The baseline characteristics for the Phase 2a study are presented in Table 11.

Table 11. Baseline characteristics for Phase 2a

Baseline characteristic	200 mg once daily (n=17)	200 mg twice daily (n=16)	400 mg once daily (n=21)
Median age (range), years	50 (20—63)	55 (30—75)	46 (25—75)
Males, n (%)	13 (77%)	9 (56%)	12 (57%)
HLA matching of donor/recipient, n (%)			
Matched	14 (82%)	13 (81%)	18 (86%)
Partially matched	3 (18%)	3 (19%)	2 (10%)
Unknown	0	0	1 (5%)
Missing	NR	NR	NR
Time from chronic GVHD diagnosis to enrolment, median (range), months	26.4 (0.0—130.7)	18.0 (1.0—69.9)	16.0 (1.0—161.9)
NIH chronic GVHD severity,[a] n (%)			
Severe	12 (71%)	14 (88%)	16 (76%)
Moderate	5 (29%)	2 (13%)	4 (19%)
Mild	0	0	1 (5%)
Organ involvement, n (%)			
No. of organs involved, median (range)	3 (2—6)	4 (1—7)	3 (2—7)
≥4 organs involved	8 (47%)	10 (63%)	9 (43%)
Eyes	14 (82%)	11 (69%)	17 (81%)
Skin	13 (77%)	12 (75%)	15 (71%)
Mouth	13 (77%)	11 (69%)	11 (52%)
Joints/fascia	11 (65%)	11 (69%)	12 (57%)
Lungs	4 (24%)	3 (19%)	10 (48%)
Upper GI	2 (12%)	4 (25%)	2 (10%)
Oesophagus	2 (12%)	0	4 (19%)
Lower GI	1 (6%)	2 (13%)	1 (5%)
Liver	0	2 (13%)	0

Baseline characteristic	200 mg once daily (n=17)	200 mg twice daily (n=16)	400 mg once daily (n=21)
Prior therapy characteristics, n (%)			
Median prior LOTs, n	3	2	2
≥2 prior LOTs	15 (88%)	9 (56%)	14 (67%)
≥4 prior LOTs	NR	NR	NR
≥6 prior LOTs	NR	NR	NR
Refractory to prior LOT[b]	11/15 (73%)	9/13 (69%)	15/20 (75%)
Concomitant systemic chronic GVHD therapies, n (%)			
CS	17 (100%)	16 (100%)	21 (100%)
CNI	7 (41%)	6 (38%)	12 (57%)
ECP	4 (24%)	4 (25%)	4 (19%)

[a] Severity was determined using the NIH Global Severity of chronic GVHD scoring
[b] For Phase 2a, denominator excludes patients with unknown status (6 total)
CNI = calcineurin inhibitors; CS = corticosteroids; ECP = extracorporeal photopheresis; GI = gastrointestinal; GVHD = graft-versus-host disease; HLA = human leukocyte antigen; LOT = line of therapy; NIH = National Institutes of Health; NR = not reported
SOURCE: Adapted from Jagasia et al. 2021(74)

B.2.3.3. Expert elicitation

Formal expert elicitation was not undertaken. However, an advisory board was held in January 2023 with 9 experienced clinical experts (including haematology consultants, a specialist nurse and a specialist pharmacist) and one health economist. Key objectives included validation of the assumptions used in the economic model and understanding the value proposition for belumosudil. Experts were selected based on their experience in the therapy area, and to represent a range of treatment centres. Questions and answers were shared by all attendees via an online platform, with three one-hour virtual meetings taking place over the course of 8 days. Individual virtual calls were arranged with some of the attendees to clarify specific points after the advisory board was completed. Transcripts from the online platform and virtual calls were synthesised into a summary report.(11)

B.2.3.4. Real-world evidence

No real-world studies of belumosudil effectiveness have been completed to date. Sanofi are planning real-world evidence studies to gather additional insights into the treatments and outcomes of patients with chronic GVHD after two prior therapies.

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

A summary of the statistical analyses performed in the ROCKstar and Phase 2a trials is provided in Table 12, with further information presented in the following sections. Details of participant flow in each trial are provided in Appendix D.2.

Table 12. Summary of statistical analyses

Trial number (acronym)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
ROCKstar (KD025-213; NCT03640481)	The primary efficacy endpoint was best ORR at any time.(73) The null hypothesis was that ORR was ≤30%.(77)	The Hochberg procedure was used for multiplicity adjustment for the primary endpoint.(73) The primary analysis was conducted using the mITT population.(73) Descriptive statistics, without multiplicity adjustment, were provided for all secondary and exploratory endpoints.(77)	The sample size was based on the primary efficacy endpoint, with one planned interim analysis and a target ORR of 55%.(73) Approximately 63 participants per treatment arm were required to provide 90% power to yield a 95% CI of ORR that excluded 30% as the lower bound.(77)	Reason for discontinuation or withdrawal was documented in the eCRF according to treatment group.(77)
Phase 2a (KD025-208; NCT02841995)	The primary efficacy endpoint was best ORR at any time.(75) A null hypothesis was not stipulated.	The study was not powered to show significant differences between dose groups with respect to efficacy, AEs or PD analyses.(75) The primary analysis was conducted using the safety population.(75)	A sample size of 16 subjects per dose group was planned to provide >90% chance of ≥1 subject experiencing an AE with an underlying rate of ≥14%.(75) Assuming a best ORR of 25%, the study was expected to have approximately 90% probability to show a response in ≥2 patients per dose group.(75)	Reason for discontinuation or withdrawal was documented in the eCRF according to dose group.(75)

AE = adverse event; CI = confidence interval; eCRF = electronic case report form; mITT = modified intent-to-treat; PD = pharmacodynamics; ORR = objective response rate

B.2.4.1. ROCKstar (KD025-213; NCT03640481)

B.2.4.1.1. Study population and sample size

The ROCKstar trial enrolled patients aged ≥12 years who had received alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 2 to 5 prior lines of treatment.(73) The sample size was based on the primary efficacy endpoint of best ORR and took into account the following considerations:(73, 77)

- The primary efficacy objective was to demonstrate an ORR >30% (i.e. the lower bound of the CI of ORR is greater than 30%)
- One planned interim analysis was planned with 0.0025 1-sided alpha spend
- A target ORR of 55% and dropout rate of 10% was assumed

For a single-arm, with a power of 90% and 2-sided alpha of 0.045 to demonstrate ORR >30%, the sample size was calculated to be 63.(73, 77)

In the event of any of the following safety findings occurring in either treatment arm (after at least 10 subjects had been enrolled), enrolment would have been paused for assessment of safety:

- Secondary graft failure in >10% of subjects
- Histological recurrence of underlying malignancy within 6 months of randomisation in >20% of subjects
- Withdrawal due to related AEs in >20% of subjects

B.2.4.1.2. Patient populations analysed

The primary analysis was conducted using the modified intent-to-treat (mITT) population, which included all randomised participants who received at least one dose of study drug.(73) The safety population in ROCKstar was equivalent to the mITT population.(77)

The responder and non-responder populations were used for some subgroup analyses:(77)

- The responder population was defined as subjects in the mITT population that achieved a PR or CR at any post-baseline response assessment
- The non-responder population was defined as any subject in the mITT population that was not a responder

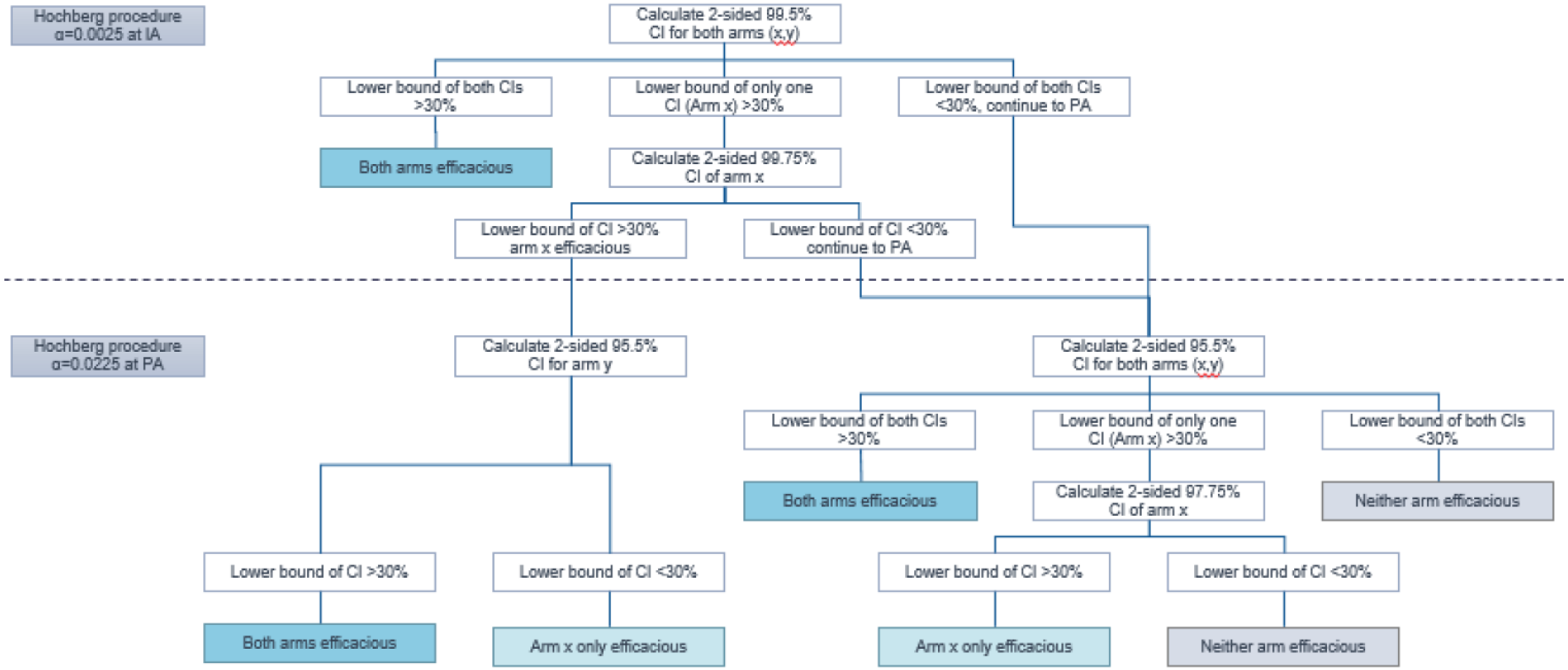
B.2.4.1.3. Statistical analyses

For the primary endpoint of best ORR, point estimates, confidence intervals (CIs) (Clopper-Pearson [exact] method), and unadjusted and Hochberg adjusted p-values corresponding to the null hypothesis of ORR ≤30% versus the alternative hypothesis of ORR >30% by treatment arms were reported.(77) Alpha was only allocated to the primary endpoint.(73)

As ROCKstar had two belumosudil treatment arms, the Hochberg procedure was used for multiplicity adjustment of the primary endpoint (Figure 7).(77)

Only descriptive statistics, without multiplicity adjustment, were provided for all secondary and exploratory endpoints (Table 13).(77)

Figure 7. Overall response rate multiplicity adjustment in ROCKstar



CI = confidence interval; IA = interim analysis; PA = primary analysis
SOURCE: Data on file. ROCKstar clinical study report(77)

Table 13. Descriptive statistics for key secondary and exploratory endpoints in the ROCKstar trial

Endpoint	Statistics provided	Populations analysed
Duration of response	• Kaplan-Meier plots and descriptive statistics, with censoring rules applied (Table 14) • Landmark analyses: number and percentage of subjects with a response sustained for ≥12, ≥20, ≥24, ≥32, ≥36, and ≥48 weeks	Responder population
Time to response	• Descriptive statistics and plots of cumulative number and percentage of responders over time (4, 8, 12, 16, 24, 32, 40, and ≥48 weeks)	Responder population
Response by organ system	• Best response at any time (CR or PR) for the 9 individual organs (skin, eyes, mouth, oesophagus, upper GI, lower GI, liver, lungs, and joints and fascia) plus GSR	mITT and responder populations
Change in LSS score	• Descriptive statistics of absolute score and change from baseline score (summary score and domain scores) as continuous variables by treatment arm and visit • Number and percentage of subjects with a ≥7-point reduction from baseline (C1D1) • Number and percentage of subjects with a ≥7-point reduction from baseline on 2 consecutive assessments • Duration of first ≥7-point reduction	mITT, responder, and non-responder populations
FFS*	• Kaplan-Meier plots, descriptive statistics of FFS, and the landmark analyses at 6, 12, 18, and 24 months • Number of events for each of the 3 components of FFS	mITT population
Time to next treatment*	• Kaplan-Meier survival method and landmark analyses	mITT population
OS	• Kaplan-Meier plots, descriptive statistics of OS, and landmark analyses at 6, 12, 18, and 24 months	mITT population
Change in CS dose	• Systemic CS dose over time • Change and percent change from baseline (C1D1) to the greatest CS dose reduction during belumosudil treatment period • Number and percentage of subjects who reduced systemic CS dose during belumosudil treatment period • Number and percentage of subjects who ever discontinued systemic CS usage during belumosudil treatment period	mITT, responder, and non-responder populations
Change in CNI dose	• Number and percentage of subjects who reduced CNI dose during the belumosudil treatment period • Number and percentage of subjects who ever discontinued CNI during the belumosudil treatment period	mITT population
Changes in the PROMIS Global Health sub-scores	• Raw scores and change from baseline values (physical and mental domains) as continuous variables by visit • Number of subjects with a ≥4.7-point reduction from baseline (C1D1)	mITT, responder, and non-responder populations

*Censored by last response assessment or long-term follow-up assessment, whichever was the latest and available

CR = complete response; CNI = calcineurin inhibitor; CS = corticosteroid; FFS = failure-free survival; GI = gastrointestinal; GSR = global severity rating; LSS = Lee symptom scale; mITT = modified intent-to-treat population; OS = overall survival; PR = partial response; PROMIS = Patient-Reported Outcomes Measurement Information System

SOURCE: Data on file. ROCKstar clinical study report(77)

Table 14. Censoring rules for duration of response in ROCKstar

Duration of response	Events	Censoring
Primary	• Deterioration from best response • Initiation of new systemic therapy for chronic GVHD • Death	Last documented response assessment If LR or initiation of new systemic therapy happened immediately after 2 or more missed response
Secondary	• Documented LR	

Company evidence submission for belumosudil for treating chronic graft-versus-host disease after 2 or more lines of systemic therapy [ID4021]

Duration of response	Events	Censoring
	• Initiation of new systemic therapy for chronic GVHD • Death	assessments, the event date was set as 4 weeks (1 cycle) after last documented response assessment prior to this event
Tertiary	• Initiation of new systemic therapy for chronic GVHD • Death	Last response assessment or Long-Term Follow-Up assessment, whichever was the latest and available
Quaternary	• Documented LR • Initiation of new systemic therapy for chronic GVHD • Death With summation of DOR from multiple episodes	Same with censoring rule for primary and secondary

DOR = duration of response; GVHD = graft-versus-host disease; LR = lack of response
SOURCE: Data on file. ROCKstar clinical study report(77)

B.2.4.1.4. Planned analyses

Three analyses were planned:(73)

- Interim analysis at approximately 2 months after 126 subjects have been enrolled into the mITT population. A nominal 1-sided alpha of 0.0025 will be spent, but there will be no early study termination for efficacy
- Primary analysis at approximately 6 months after 126 subjects have been enrolled into the mITT population, with 1-sided alpha 0.0225 (or 0.025 if the ORRs of both arms are significant at interim)
- Follow-up analysis at approximately 12 months after 126 subjects have been enrolled into the mITT population.

B.2.4.1.5. Participant flow

Detailed information on participant flow in the ROCKstar trial is provided in Appendix D, including the consort diagram.

B.2.4.2. Phase 2a study (KD025-208; NCT02841995)

B.2.4.2.1. Study population and sample size

The Phase 2a study enrolled patients aged ≥18 years who had received allogeneic bone marrow transplant or alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 1 to 3 prior lines of systemic treatment and currently receiving CS treatment ± CNI ± concurrent ECP.(74)

The planned sample size of 16 patients per dose group provided >90% probability of one or more patients experiencing an AE that had an underlying rate of ≥14%.(74) Assuming a best ORR of 25%, each dose group of 16 patients had an approximately 90% chance of at least two patients meeting the overall response criteria.(74)

B.2.4.2.2. Patient populations analysed

The primary analysis was conducted using the safety population which was defined as enrolled patients who received at least one dose of study treatment.(74)

B.2.4.2.3. Statistical analyses

The Phase 2a study was not powered to show significant differences between dose groups with respect to efficacy, AEs or exploratory pharmacodynamic analyses.(74) No adjustments were made for the multiplicity of endpoints, and missing values were not imputed.(75)

For the primary efficacy endpoint of best ORR, the Clopper-Pearson (exact) method was used to construct the two-sided 95% CI.(74) The Kaplan-Meier method was used to calculate estimates of FFS and OS.(74)

B.2.4.2.4. Planned analyses

Patients were enrolled sequentially into the three dose groups (200 mg once daily, 200 mg twice daily and 400 mg once daily).(75) Safety data in each dose group were analysed after 8 patients reached 2 months of treatment to ensure there was no safety signal before enrolment in the next group and dose-escalation.(75)

- If ≥25% of patients in a dose group experienced a Common Terminology for Adverse Events (CTCAE) v4.03 Grade 2 liver toxicity or a CTCAE v4.03 Grade ≥3 AE in the same organ or body system, or if >25% of patients in a cohort were discontinued for toxicity that persists for 14 days, then dose-escalation to the next cohort would not occur and all subjects in that dose cohort would be dose reduced.
- If ≥25% of patients in the 200 mg once daily group experienced a CTCAE v4.03 Grade 2 liver toxicity or a CTCAE v4.03 ≥Grade 3 AE in the same organ or body system, or if >25% of patients are discontinued for toxicity that persists for 14 days, then further dosing would not occur and the study would be terminated.

Belumosudil was administered orally in 28-day cycles until disease progression or unacceptable toxicity.(75) Long-term follow-up was conducted every 8 weeks until study end.(75)

B.2.4.2.5. Participant flow

Detailed information on participant flow in the Phase 2a study is provided in Section D.2.2, including the consort diagram.

B.2.5. Critical appraisal of the relevant clinical effectiveness evidence

Study results published in a peer-reviewed journal were used as the primary source of data where available; clinical study reports (CSRs) were used as additional data sources as needed.

B.2.5.1. Quality assessment

The ROCKstar trial was assessed using the Cochrane Risk of Bias Assessment Tool 2.0 and found to have a low risk of bias across all domains (Table 15). As the Phase 2a study is a single-arm non-randomised trial, it was assessed using the Downs and Black Quality Assessment Checklist(78) and given a score of 17 (corresponding to a quality level of 'fair').(79)

A detailed overview of the quality assessments for each trial identified by the clinical SLR is provided in Appendix D.

Table 15. Cochrane Risk of Bias Tool Assessment of ROCKstar

Bias Domain	Signalling Questions	Response
Bias arising from the randomisation process	1.1 Was the allocation sequence random?	Y
	1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?	NI
	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	N
	Risk of bias judgement	**Low**
Bias due to deviations from intended interventions	2.1. Were participants aware of their assigned intervention during the trial?	Y
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y
	2.3 If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N
	2.4 If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	NA
	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	NA
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomised?	N
	Risk of bias judgement	**Low**
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Y
	3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	NA
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA

Bias Domain	Signalling Questions	Response
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA
	Risk of bias judgement	**Low**
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N
	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	N
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NA
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	N
	Risk of bias judgement	**Low**
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalised before unblinded outcome data were available for analysis?	Y
	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g., scales, definitions, time points) within the outcome domain?	N
	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	N
	Risk of bias judgement	**Low**
Overall bias	**Risk of bias judgement**	**Low**

N = no; NI = no information; PN = probably no; PY = probably yes; Y = yes

B.2.5.2. Applicability of the study results to clinical practice in England

The patients included in the ROCKstar trial and Phase 2a study were exclusively recruited at centres in the US.(73, 74) However, no physiological differences are expected between the US and English patient populations.(11)

Patients in the US and England are likely to receive the same CS and/or CNI-based initial therapies; however, after two prior therapies, patients in the US have access to a wider selection of possible treatment options and may experience a different treatment pathway compared to patients in England, where ECP, MMF, rituximab, sirolimus, imatinib and CNIs are the only widely used therapies for previously-treated patients with chronic GVHD.(69, 80) Results from the ROCKstar trial (Section B.2.6.1) showed that the use of different concomitant and prior therapies had no significant impact on outcomes with belumosudil.(77) Therefore the results of ROCKstar are expected to be applicable to patients in routine clinical practice in England.

B.2.6. Clinical effectiveness results of the relevant studies

B.2.6.1. ROCKstar (KD025-213; NCT03640481)

The clinical effectiveness data presented in this section and subsequently used in the modelling as part of the pool with the Phase 2a data is based on the data cut from August 2021. The ROCKstar study is an ongoing trial and a subsequent data cut was carried out for regulatory purposes. This became available during the late stages of preparation of this dossier. This regulatory analysis was performed on the ROCKstar data but not for the pooled data representing the licenced population that is used in the model. No additional analysis specific to the model (for example curve fits, or mapping of PROMIS) was performed on the ROCKstar data. For completeness these data are presented in Appendix O and confirm the results from the 2021 data cut.

B.2.6.1.1. Efficacy

The primary endpoint in ROCKstar, best overall response rate (ORR) at any time, was defined as the percentage of patients that had either complete response (CR) or partial response (PR), using the 2014 National Institutes of Health (NIH) Consensus Criteria as assessed by investigators.(73, 81) The NIH criteria are provided in full in Appendix M.

Failure-free survival (FFS) was included as a secondary endpoint, defined as the time from the first dose of belumosudil to the time of initiation of new systemic chronic GVHD therapy, non-relapse mortality or recurrent malignancy (whichever occurred first; Table 16).(73) FFS is a simple yet robust endpoint, that incorporates key objective measures of chronic GVHD disease progression into a single composite endpoint.(82) FFS is therefore a clinically meaningful endpoint which captures chronic GVHD disease control (prevent or delay the need for chronic GVHD treatment change), control of the underlying disease (malignancy) and survival information (Section B.3.2.2).(82)

Table 16. Summary of the response criteria and failure-free survival definition used in ROCKstar

Criteria	Summary of definition
Complete Response	Resolution of all manifestations of chronic GVHD in each organ or site
Partial Response	Improvement in at least 1 organ or site without progression in any other organ or site
Failure-Free Survival	The absence of chronic GVHD treatment change, non-relapse mortality, and recurrent malignancy

GVHD = graft-versus-host disease
SOURCE: Kadmon Pharmaceuticals 2020(77)

Treatment with belumosudil in ROCKstar led to a rapid response with 72% of patients achieving response within 6 months, and with an overall median time to response of 5 weeks.(83) The response rates for ROCKstar are presented in Table 17 and Figure 8. The data cut-off date for the main analysis was August 2020 with a second analysis conducted with a data cut-off date of August 2021. Note that results are not expected to change between data cuts for best ORR, best ORR within 6 months, and best ORR within 12 months due to the nature of these endpoints. Results are therefore

only presented for the August 2021 data cut, as these data were used to inform the belumosudil cost-effectiveness model (Section B.3.3).

Table 17. ROCKstar best overall response rates in reported time periods (mITT)

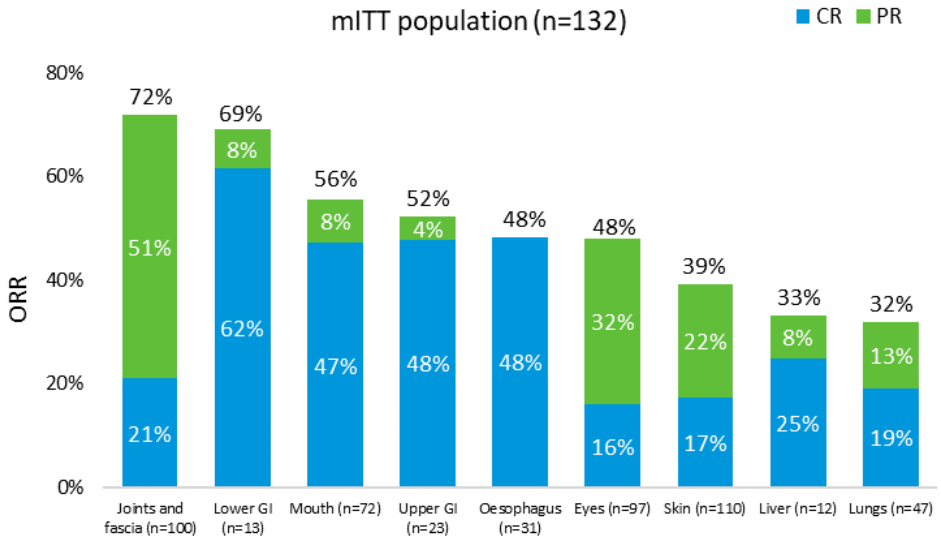
	August 2021 data cut		
	200 mg once daily (n=66)	**200 mg twice daily (n=66)**	**Total (N=132)**
Median time to response, weeks (range)	4 (4, 41)	5 (4, 66)	5 (4, 66)
Best ORR,[a] n (%)	49 (74%)	51 (77%)	100 (76%)
[95% CI]	[62, 84]	[65, 87]	[68, 83]
Best ORR within 6 months of treatment,[a] n (%)	47 (71.2%)	48 (72.7%)	95 (72.0%)
[95% CI]	[59, 82]	[60, 83]	[64, 79]
CR, n (%)	2 (3.0%)	1 (1.5%)	3 (2.3%)
PR, n (%)	45 (68.2%)	47 (71.2%)	92 (69.7%)
Best ORR within 12 months of treatment,[a] n (%)	49 (74.2%)	50 (75.8%)	99 (75.0%)
[95% CI]	[62, 84]	[64, 86]	[67, 82]
CR, n (%)	3 (4.5%)	1 (1.5%)	4 (3.0%)
PR, n (%)	46 (69.7%)	49 (74.2%)	95 (72.0%)
Best ORR by organ system,[a] n/N (%)			
Joints and fascia	37/51 (72.5%)	35/49 (71.4%)	72/100 (72.0%)
Lower GI	4/6 (66.7%)	5/7 (71.4%)	9/13 (69.2%)
Mouth	16/30 (53.3%)	24/42 (57.1%)	40/72 (55.6%)
Upper GI	8/13 (61.5%)	4/10 (40.0%)	12/23 (52.2%)
Oesophagus	10/19 (52.6%)	5/12 (41.7%)	15/31 (48.4%)
Eyes	19/48 (39.6%)	27/49 (55.1%)	46/97 (47.4%)
Liver	3/9 (33.3%)	1/3 (33.3%)	4/12 (33.3%)
Skin	17/55 (30.9%)	26/55 (47.3%)	43/110 (39.1%)
Lungs	9/24 (37.5%)	6/23 (26.1%)	15/47 (31.9%)

[a] Best ORR at any time was defined as the percentage of patients that had either CR or PR, using the 2014 NIH Consensus Criteria

CI = confidence interval; CR = complete response; GI = gastrointestinal; mITT = modified intent-to-treat population; NIH = National Institutes of Health; ORR = overall response rate; PR = partial response

SOURCE: Kadmon Pharmaceuticals 2022(83)

Figure 8. ROCKstar best overall response rates by organ system among responders (mITT)



August 2021 data cut.
Note, numbers may not sum to totals due to rounding.
CR = complete response; GI = gastrointestinal; mITT = modified intent-to-treat population; ORR = overall response rate; PR = partial response
SOURCE: Kadmon Pharmaceuticals 2022(83)

High ORRs were observed in all subgroups, including patients with severe chronic GVHD, patients with ≥4 organs involved at baseline and patients with prior ibrutinib or ruxolitinib treatment (Appendix E).(83) Neither ibrutinib nor ruxolitinib are used in clinical practice in England. However, the ROCKstar results indicate that prior treatment with these agents (as with all subgroups tested) does not significantly influence efficacy outcomes with belumosudil, therefore the results of ROCKstar are expected to be applicable to patients in routine clinical practice in England.(83) Additionally, ORR outcomes did not vary significantly by concomitant therapy, highlighting that the improvements observed in ROCKstar are most likely attributable to the addition of belumosudil.(77)

The majority of patients in ROCKstar (68%) reduced their CS use during the study.(83) The mean reduction in CS dose from baseline was 50% in the mITT population, and many patients discontinued CS and other concomitant medications.(83) The rates of discontinuation and measures of reduction are presented in Table 18.

Table 18. ROCKstar: Discontinuations and reductions of concomitant medications (mITT)

	August 2021 data cut		
	200 mg once daily (n=66)	**200 mg twice daily (n=66)**	**Total (N=132)**
CS reduction, n (%)	42 (64%)	48 (73%)	90 (68%)
Mean change in CS dose from baseline, %	−48%	−52%	−50%
CS discontinuation, n (%)	19 (29%)	16 (24%)	35 (27%)
CNI discontinuation, %	21%	33%	27%

CNI = calcineurin inhibitor; CS = corticosteroid; mITT = modified intent-to-treat population; NR = not reported
SOURCE: Kadmon Pharmaceuticals 2022(83)

Patients treated with belumosudil also experienced a sustained response in ROCKstar.(83) The median duration of response (DOR) for patients who achieved response exceeded one year (83.1 weeks).(83) A summary of the sustained response measures is presented in Table 19.

The Kaplan-Meier plots for FFS in the total mITT population are shown in Figure 9. Median FFS was 14.3 months (95% CI: 10.2, 23.8).(83) The Kaplan-Meier estimate of FFS was 75% at 6 months, 56% at 12 months and 41% at 2 years.(83) The most common failure event at 12 months was initiation of new systemic therapy for chronic GVHD (43 [32.6%] subjects).(83) In total, 9 (6.8%) subjects had a failure event of non-relapse mortality at 12 months.(83) Overall survival at two years was 83% (Table 19).(83)

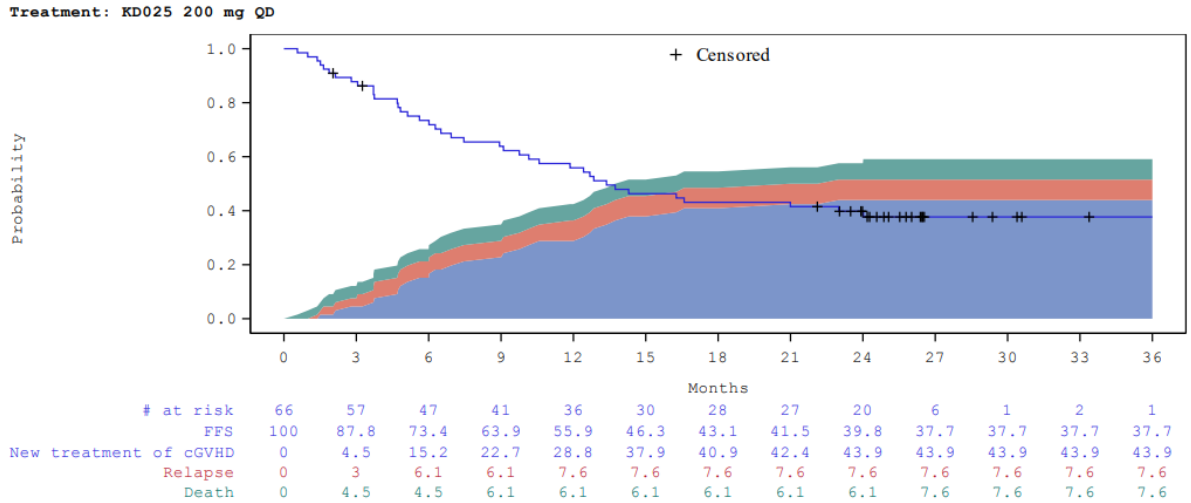
Table 19. ROCKstar: Duration of response, failure-free survival and overall survival (mITT)

	August 2021 data cut		
	200 mg once daily (n=66)	**200 mg twice daily (n=66)**	**Total (N=132)**
Median DOR[a] in responders (primary/secondary), weeks	22.1	24.1	24.1
Median DOR[a] in responders (quaternary), weeks	96.0	74.3	83.1
DOR[a] ≥20 weeks in responders, %	46.9%	51.0%	49.0%
FFS median (months) (95% CI)	13.4 (9.1, 24.0)	15.1 (9.6, NR)	14.3 (10.2, 23.8)
FFS[b] at 6 months, % (95% CI)	73% (61, 83)	76% (63, 84)	75% (66, 81)
FFS[b] within 12 months, % (95% CI)	56% (43, 67)	56% (43, 67)	56% (47, 64)
FFS[b] within 24 months, % (95% CI)	40% (28, 52)	43% (30, 54)	41% (33, 50)
2-year OS rate, % (95% CI)	84% (72, 91)	83% (71, 90)	83% (75, 89)

[a] DOR was measured from the time of first documentation of response to the time of first documented deterioration from best response (primary), to the time of first documented lack of response (secondary), or to the time of first documented lack of response with durations summed for multiple response/lack of response episodes (quaternary)
[b] FFS was defined as the interval between the start of treatment and the addition of a new chronic GVHD therapy, relapse, or non-relapse mortality
CI = confidence interval; DOR = duration of response; FFS = failure-free survival; GVHD = graft-versus-host disease; mITT = modified intent-to-treat population; NR = not reported; OS = overall survival
SOURCES: Kadmon Pharmaceuticals 2022(83)

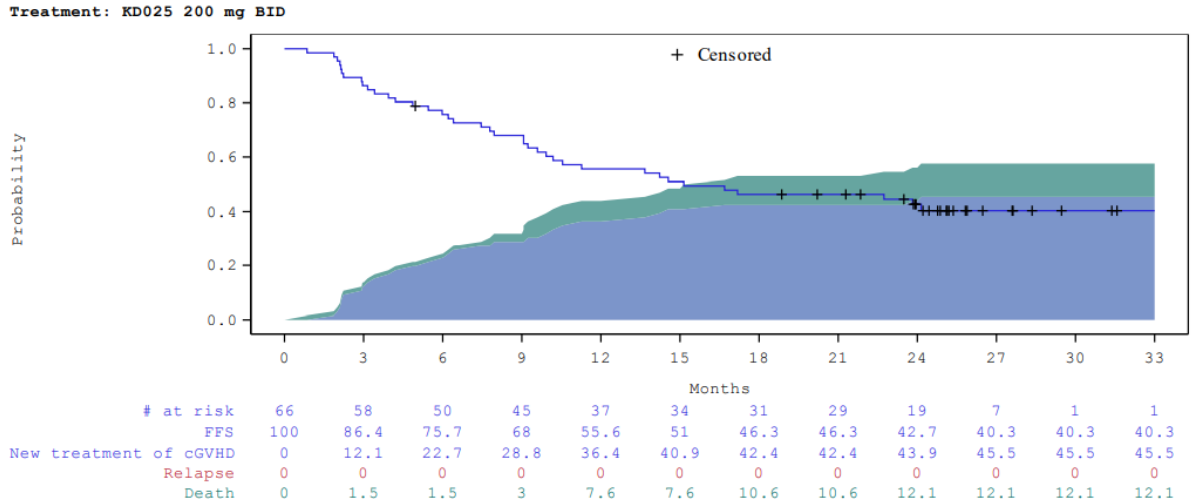
Figure 9. ROCKstar: Kaplan-Meier plot for failure-free survival (total population; mITT)

(A) Belumosudil 200 mg once daily



# at risk	66	57	47	41	36	30	28	27	20	6	1	2	1
FFS	100	87.8	73.4	63.9	55.9	46.3	43.1	41.5	39.8	37.7	37.7	37.7	37.7
New treatment of cGVHD	0	4.5	15.2	22.7	28.8	37.9	40.9	42.4	43.9	43.9	43.9	43.9	43.9
Relapse	0	3	6.1	6.1	7.6	7.6	7.6	7.6	7.6	7.6	7.6	7.6	7.6
Death	0	4.5	4.5	6.1	6.1	6.1	6.1	6.1	6.1	7.6	7.6	7.6	7.6

(B) Belumosudil 200 mg twice daily

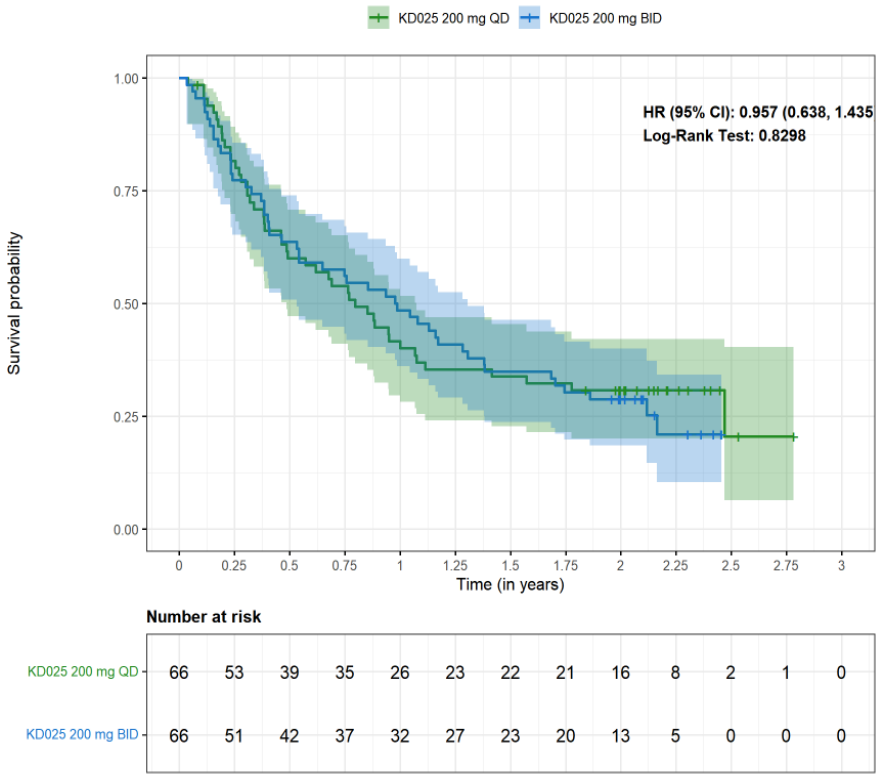


# at risk	66	58	50	45	37	34	31	29	19	7	1	1
FFS	100	86.4	75.7	68	55.6	51	46.3	46.3	42.7	40.3	40.3	40.3
New treatment of cGVHD	0	12.1	22.7	28.8	36.4	40.9	42.4	42.4	43.9	45.5	45.5	45.5
Relapse	0	0	0	0	0	0	0	0	0	0	0	0
Death	0	1.5	1.5	3	7.6	7.6	10.6	10.6	12.1	12.1	12.1	12.1

Data cut-off: August 2021
BID = twice daily; cGVHD = chronic graft-versus-host disease; FFS = failure-free survival; mITT = modified intent-to-treat population; QD = once daily
SOURCE: Kadmon Pharmaceuticals 2022(84)

The Kaplan-Meier plot for time to treatment discontinuation (TTD) is shown in Figure 10. Median TTD was 9.4 months for belumosudil 200 mg once daily and 11.8 months for 200 mg twice daily.(83)

Figure 10. ROCKstar: Kaplan-Meier plot for time to treatment discontinuation (total population; mITT)



Data cut-off: August 2021
BID = twice daily; CI = confidence interval; HR = hazard ratio; KD025 = belumosudil; mITT = modified intent-to-treat population; QB = once daily
SOURCE: Parametric fitting analyses of event (discontinuation) counts in the ROCKstar trial

B.2.6.1.2. Patient-reported outcomes

In ROCKstar, QoL and symptom bother were measured using the 7-day Lee Symptom Scale (LSS) summary score (described in Appendix M).(73) A clinically meaningful improvement from baseline, defined as a reduction of ≥7 points, was achieved in a majority of patients treated with 200 mg once daily (62.1%) and 200 mg twice daily (63.6%) in the mITT population (Table 20).(83) ROCKstar also assessed QoL with an exploratory endpoint using the PROMIS-GH questionnaire (described in Appendix M).(85) PROMIS physical and mental health scores can be mapped to the EuroQoL-5 Dimensions (EQ-5D) measure to provide utility scores for use in QoL analyses (Section B.3.4.2).(86, 87) A greater percentage of patients reported improvement of ≥4.7 points, previously identified as a clinically meaningful difference in chronic GVHD,(19) in their raw physical health scores than in their raw mental health scores (Table 20).(83)

Table 20. ROCKstar: Improvement in HRQoL scores (mITT)

	August 2021 data cut		
	200 mg once daily (n=66)	**200 mg twice daily (n=66)**	**Total (N=132)**
Improvement in LSS score ≥7 points from baseline			
Overall, n (%)	41 (62.1%)	42 (63.6%)	83 (62.9%)
Responders, n/N (%)	NR	NR	NR
Non-responders, n/N (%)	NR	NR	NR
Improvement in PROMIS raw mental health score ≥4.7 points from baseline, n (%)	31 (47.0%)	32 (48.5%)	63 (47.7%)
Improvement in PROMIS raw physical health score ≥4.7 points from baseline, n (%)	35 (53.0%)	31 (47.0%)	66 (50.0%)

HRQoL = health-related quality of life; LSS = Lee Symptom Scale; mITT = modified intent-to-treat population; NR = not reported; PROMIS = Patient-Reported Outcomes Measurement Information System
SOURCES: Kadmon Pharmaceuticals 2022(83)

B.2.6.2. Phase 2a study (KD025-208; NCT02841995)

B.2.6.2.1. Efficacy

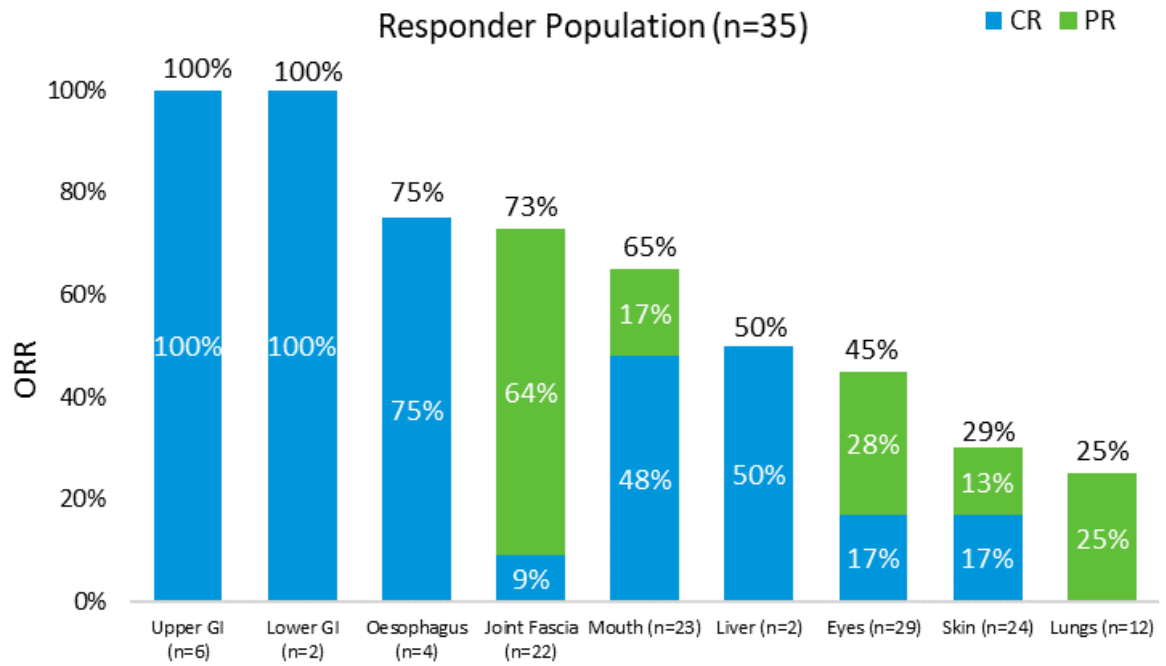
In the Phase 2a study, the majority of patients achieved a rapid response, with >75% of the patients who responded to treatment achieving a response by Week 8.(74) The response rates for the Phase 2a study are presented in Table 21. Results for best response by organ system are shown in Figure 11.

Table 21. Phase 2a: Response rates (safety population)[a]

	200 mg once daily (n=17)	**200 mg twice daily (n=16)**	**400 mg once daily (n=21)**	**Overall (N=54)**
Best ORR,[a] n (%) [95% CI]	11 (65%) [38, 86]	11 (69%) [41, 89]	13 (62%) [38, 82]	35 (65%) [51, 77]
Median follow-up, months	36	32	24	29

[a] Data cut-off: February 2020
[b] ORR at any time was defined as the percentage of patients that had either complete response or partial response, using the 2014 NIH Consensus Criteria
CI = confidence interval; mITT = modified intent-to-treat population; NIH = National Institutes of Health; NR = not reported; ORR = overall response rate
SOURCES: Adapted from Jagasia et al. 2021(74)

Figure 11. Phase 2a: Best response by organ system among responders (responder population)[a]



Note: n represents the number of patients in the responder population for global severity rating and number of specific organs involved at baseline. No partial responses were observed for upper GI, lower GI, oesophagus or liver. Numbers may not sum to totals due to rounding.
[a] Data cut-off: February 2020
CR = complete response; GI = gastrointestinal; ORR = overall response rate; PR = partial response
SOURCE: Adapted from Jagasia et al. 2021(74)

Best ORR was ≥50% in all subgroups in the Phase 2a study, including patients with ≥2 lines of prior therapy, patients with ≥4 organs involved at baseline, and patients with severe chronic GVHD (Appendix E).(74)

The majority of patients in the Phase 2a study (67%) reduced their CS use during the study, and approximately one fifth of patients were able to discontinue CS.(74) The rates of discontinuation and measures of reduction are presented in Table 22.

Table 22. Phase 2a: Discontinuations and reductions of concomitant medications (mITT)[a]

	200 mg once daily (n=17)	200 mg twice daily (n=16)	400 mg once daily (n=21)	Total (N=54)
CS reduction, n (%) [95% CI]	13 (76%) [50, 93]	9 (56%) [30, 80]	14 (67%) 43, 85]	36 (67%) [53, 79]
Mean change in CS dose from baseline, %	−50%	−36%	−47%	−45%
CS discontinuation, n (%) [95% CI]	4 (24%) [7, 50]	2 (13%) [2, 38]	4 (19%) [5, 42]	10 (19%) [9, 31]
Median time to CS discontinuation, weeks	NR	NR	NR	29

[a] Data cut-off: February 2020
CI = confidence interval; CS = corticosteroid; mITT = modified intent-to-treat; NR = not reported
SOURCE: Adapted from Jagasia et al. 2021(74)

Patients treated with belumosudil in the Phase 2a study achieved a sustained response for a median of 35 weeks, which increased to 38 weeks in patients with ≥2 prior systemic therapies.(74) A summary of the sustained response measures is presented in Table 23.

Table 23. Phase 2a: Duration of response (mITT)[a]

	Total (N=54)
Median DOR[b] in responders, weeks	35
FFS,[c] % (95% CI)	
6 months	76% (62, 85)
12 months	47% (33, 60)
24 months	33% (21, 46)
FFS[c] with response at 12 months, %	24%
OS, % (95% CI)	
12 months	91% (79, 96)
24 months	82% (69, 90)
Median TTNT, months	14

[a] Data cut-off: February 2020
[b] DOR was measured from the time of initial partial response or complete response until documented progression from best response of chronic GVHD, time from initial response to start of additional systemic chronic GVHD therapy, or death
[c] Reasons for failure included new treatment of chronic GVHD, relapse, or death
CI = confidence interval; DOR = duration of response; FFS = failure-free survival; GVHD = graft-versus-host disease; OS = overall survival; mITT = modified intent-to-treat; NR = not reported; TTNT = time to next treatment
SOURCES: Adapted from Jagasia et al. 2021(74)

B.2.6.2.2. Patient-reported outcomes

The Phase 2a study measured QoL and symptom bother using the LSS summary with the same definition for clinically meaningful improvement (a reduction of ≥7 points; see Appendix M).(74) A clinically meaningful improvement was achieved in a substantial proportion of patients treated with 200 mg once daily, 200 mg twice daily, and 400 mg once daily (Table 24).(74)

Table 24. Improvement in LSS score (mITT)[a]

	200 mg once daily (n=17)	**200 mg twice daily (n=16)**	**400 mg once daily (n=21)**	**Total (N=54)**
Improvement in LSS score ≥7 points from baseline				
Overall, n (%)	9 (53%)	7 (44%)	11 (52%)	27 (50%)
Responders, n/N (%)	8/11 (73%)	3/11 (27%)	9/13 (69%)	20/35 (57%)
Non-responders, n/N (%)	1/6 (17%)	4/5 (80%)	2/8 (25%)	7/19 (37%)

[a] Data cut-off: February 2020
LSS = Lee Symptom Scale; mITT = modified intent-to-treat
SOURCES: Adapted from Jagasia et al. 2021(74)

B.2.6.3. Pooled analysis of ROCKstar trial and Phase 2a study

A pooled analysis of the ROCKstar and Phase 2a study data was conducted. For this analysis, data from the respective groups in each trial receiving 200 mg belumosudil once daily (N=83 in total) or 200 mg belumosudil twice daily (N=82 in total) were pooled.(88) The population included all patients who received at least one dose of the study medication (safety population). Trial data were summarised by extent of exposure using six month intervals to group exposure ranges, while overall

exposure was summarised in total patient-years. Patient disposition for those discontinuing treatment and primary reason for discontinuation were summarised and presented by exposure range. Demographic information, baseline characteristics, medical history and concomitant medications were summarised. Kadmon Algorithmic Response Assessments (KARA) were used for overall response analyses. Individual organ responses were assessed using individual scores or KARA for skin, joints/fascia, eyes, mouth, lungs, oesophagus, upper GI, lower GI and liver. Safety analyses included AEs, clinical laboratory evaluations, vital sign measurements and ECG data. For missing start or end dates of AE information, the worst or most conservative judgement was used. Any patient lost to follow-up without any response assessment was counted as a non-responder in the pooled analysis.

The patient population and baseline demographics (Table 25) included in the analysis were generally consistent with the real-world patient population expected for chronic GVHD.(11, 88)

Table 25. Baseline characteristics of pooled analysis

Baseline characteristic	200 mg once daily n=83	200 mg twice daily n=82	Combined 200 mg N=165
Median age (range), years	53.0 (20—77)	57.0 (21—77)	55.0 (20—77)
Males, n (%)	55 (66.3%)	42 (51.2%)	97 (58.8%)
GVHD prophylaxis after transplant, n (%)			
None	0	1 (1.2%)	1 (0.6%)
CNI only	5 (6.0%)	6 (7.3%)	11 (6.7%)
CNI + methotrexate	33 (39.8%)	32 (39.0%)	65 (39.4%)
CNI + methotrexate + other	8 (9.6%)	7 (8.5%)	15 (9.1%)
CNI + MMF	11 (13.3%)	17 (20.7%)	28 (17.0%)
CNI + MMF + other	4 (4.8%)	3 (3.7%)	7 (4.2%)
CNI + MMF + ATG	0	1 (1.2%)	1 (0.6%)
CNI + sirolimus	8 (9.6%)	9 (11.0%)	17 (10.3%)
CNI + corticosteroids	2 (2.4%)	1 (1.2%)	3 (1.8%)
Other regimen	12 (14.5%)	5 (6.1%)	17 (10.3%)
HLA matching of donor/recipient, n (%)			
Matched	71 (85.5%)	75 (91.5%)	146 (88.5%)
Partially matched	11 (13.3%)	6 (7.3%)	17 (10.3%)
Unknown	0	1 (1.2%)	1 (0.6%)
Missing	1 (1.2%)	0	1 (0.6%)
History of acute GVHD, n (%)	55 (66.3%)	60 (73.2%)	115 (69.7%)
Time from chronic GVHD diagnosis to enrolment, median (range), months	25.26 (0.0—162.4)	28.14 (1.0—144.1)	27.20 (0.0—162.4)
NIH chronic GVHD severity[a] n (%)			
Severe	58 (69.9%)	57 (69.5%)	115 (69.7%)
Moderate	23 (27.7%)	25 (30.5%)	48 (29.1%)
Mild	2 (2.4%)	0	2 (1.2%)
Organ involvement, n (%)			
≥4 organs involved	42 (50.6%)	44 (53.7%)	86 (52.1%)
≥6 organs involved	13 (15.7%)	11 (13.4%)	24 (14.5%)
Skin	68 (81.9%)	67 (81.7%)	135 (81.8%)
Joints/fascia	62 (74.7%)	60 (73.2%)	122 (73.9%)
Eyes	62 (74.7%)	60 (73.2%)	122 (73.9%)

Baseline characteristic	200 mg once daily n=83	200 mg twice daily n=82	Combined 200 mg N=165
Mouth	43 (51.8%)	53 (64.6%)	96 (58.2%)
Lungs	29 (34.9%)	25 (30.5%)	54 (32.7%)
Oesophagus	21 (25.3%)	12 (14.6%)	33 (20.0%)
Upper GI	15 (18.1%)	14 (17.1%)	29 (17.6%)
Lower GI	7 (8.4%)	9 (11.0%)	16 (9.7%)
Liver	9 (10.8%)	5 (6.1%)	14 (8.5%)
No. of organs involved, median (range)	4.0 (0—7)	4.0 (1—7)	4.0 (0—7)
Refractory to prior LOT, n (%)	56 (78.9%)	44 (65.7%)	100 (72.5%)
Number or prior lines of therapy, n (%)			
1	2 (2.4%)	7 (8.5%)	9 (5.5%)
2	26 (31.3%)	15 (18.3%)	41 (24.8%)
3	23 (27.7%)	25 (30.5%)	48 (29.1%)
4	17 (20.5%)	14 (17.1%)	31 (18.8%)
5	14 (16.9%)	19 (23.2%)	33 (20.0%)
≥6	1 (1.2%)	2 (2.4%)	3 (1.8%)
Median	3.0	3.0	3.0
Prior systemic chronic GVHD therapies, n (%)[b]	82 (98.8%)	81 (98.8%)	163 (98.8%)
Prednisone	48 (57.8%)	50 (61.0%)	98 (59.4%)
Tacrolimus	39 (47.0%)	41 (50.0%)	80 (48.5%)
Sirolimus	36 (43.4%)	36 (43.9%)	72 (43.6%)
ECP	23 (27.7%)	24 (29.3%)	47 (28.5%)
Ibrutinib	22 (26.5%)	19 (23.2%)	41 (24.8%)
Mycophenolate mofetil	23 (27.7%)	16 (19.5%)	39 (23.6%)
Rituximab	20 (24.1%)	18 (22.0%)	38 (23.0%))
Ruxolitinib			
Concomitant systemic chronic GVHD therapies, n (%)			
Systemic hormonal preparations	82 (98.8%)	81 (98.8%)	163 (98.8%)
ECP	20 (24.1%)	26 (31.7%)	46 (27.9%)
Tacrolimus	29 (34.9%)	30 (36.6%)	59 (35.8%)
Sirolimus	17 (20.5%)	18 (22.0%)	35 (21.2%)
MMF	11 (13.3%)	2 (2.4%)	13 (7.9%)
Imatinib	1 (1.2%)	0	1 (0.6%)
Rituximab	1 (1.2%)	0	1 (0.6%)
Ruxolitinib	1 (1.2%)	0	1 (0.6%)

[a] Severity was determined using the NIH Global Severity of chronic GVHD scoring
[b] This table includes the most common therapies for chronic GVHD (≥10%), as well as ECP
ATG = antithymocyte globulin, CNI = calcineurin inhibitor; ECP = extracorporeal photopheresis, GI = gastrointestinal; GVHD = graft-versus-host disease; HLA = human leukocyte antigen; MMF = mycophenolate mofetil; NIH = National Institutes of Health
SOURCES: Kadmon Pharmaceuticals 2022(88)

B.2.6.3.1. Efficacy

The pooled analysis of data from both phase 2 studies showed the same overall pattern of results as the individual trials: high levels of response (Table 26),(83) with no meaningful differences across subgroups (Appendix E),(88, 89) and responders experiencing benefit rapidly and for a long duration (Figure 12).(85)

Table 26. Results of pooled efficacy analysis (mITT; 2 year analysis)

	August 2021 data cut		
	200 mg once daily (n=83)	**200 mg twice daily (n=82)**	**Combined 200 mg (N=165)**
Median time to response, weeks (range)	7.9 (3.7—80.1)	7.9 (3.7—40.1)	7.9 (3.7—80.1)
Best ORR,[a] n (%)	60 (72.3%)	60 (73.2%)	120 (72.7%)
Best ORR by organ system, n/N (%)			
Skin	20/68 (29.4%)	28/67 (41.8%)	48/135 (35.6%)
Eyes	25/62 (40.3%)	30/60 (50.0%)	55/122 (45.1%)
Mouth	21/43 (48.8%)	29/53 (54.7%)	50/96 (52.1%)
Oesophagus	11/21 (52.4%)	5/12 (41.7%)	16/33 (48.5%)
Upper GI	10/15 (66.7%)	8/14 (57.1%)	18/29 (62.1%)
Lower GI	4/7 (57.1%)	7/9 (77.8%)	11/16 (68.8%)
Liver	2/9 (22.2%)	2/5 (40.0%)	4/14 (28.6%)
Lungs	8/29 (27.6%)	6/25 (24.0%)	14/54 (25.9%)
Joints/fascia	42/62 (67.7%)	38/60 (63.3%)	80/122 (65.6%)
Median DOR in responders (primary/secondary),[b] weeks (95% CI)	22.1 (11.43, 44.14)	24.1 (12.14, 48.14)	24.1 (16.14, 35.14)
Median DOR in responders (quaternary), weeks (95% CI)	62.3 (28.29, 108.43)	71.7 (32.14, 95.43)	69.9 (40.00, 95.43)
Median FFS, months (95% CI)	13.7 (8.51, 24.02)	15.1 (9.59, 23.82)	14.8 (10.58, 20.57)
FFS, % (95% CI)			
FFS at 6 months	72% (60, 80)	78% (67, 85)	75% (67, 81)
FFS within 12 months	54% (42, 64)	56% (45, 66)	55% (47, 63)
FFS within 24 months	40% (29, 50)	39% (28, 49)	39% (31, 47)
Median OS (months)	NA (39.46, NA)	NA (NA, NA)	NA (NA, NA)
OS, % (95% CI)			
OS within 12 months	90% (81, 95)	91% (83, 96)	91% (85, 94)
OS within 24 months	85% (75, 91)	85% (75, 91)	85% (78, 90)
Median TTD, months (range)	9.2 (0.5—56.7)	10.6 (0.4—45.9)	10.2 (0.4—56.7)

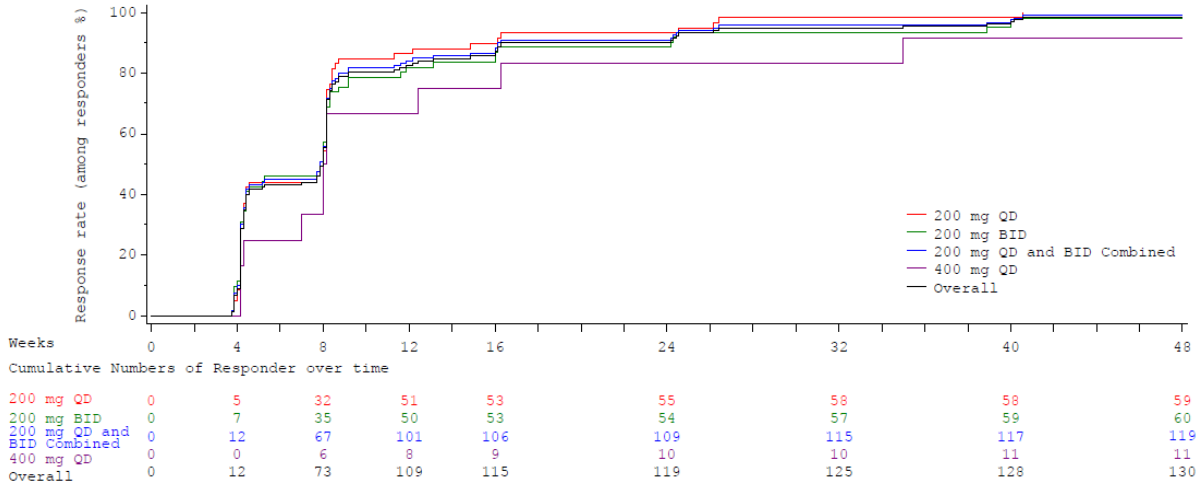
[a] Best ORR at any time was defined as the percentage of patients that had either complete response or partial response, using the 2014 NIH Consensus Criteria

[b] DOR was measured from the time of first documentation of response to the time of first documented deterioration from best response (primary), to the time of first documented lack of response (secondary), or to the time of first documented lack of response with durations summed for multiple response/lack of response episodes (quaternary)

CI = confidence interval; DOR = duration of response; FFS = failure-free survival; GI = gastrointestinal; mITT = modified intent-to-treat; NA = not available; NIH = National Institutes of Health; ORR = overall response rate; OS = overall survival; TTD = time to treatment discontinuation

SOURCES: Kadmon Pharmaceuticals 2022(88, 89)

Figure 12. Cumulative response rate (among responders; 1 year analysis)[a,b]



[a] Response assessments performed on or after initiation of new systemic therapy for chronic GVHD are excluded from the analysis
[b] Data cut-off August 19, 2020. The cumulative response rates presented in this figure do not include the latest data cut from August 2021; however, as the outcomes of the August 2021 data cut are very similar, the figures would not differ significantly.
BID = twice daily; cGVHD = chronic graft -versus-host disease; QD = once daily
SOURCE: Kadmon Pharmaceuticals 2022(85)

The majority of patients in the pooled analysis (68.5% in the combined 200 mg group) reduced their CS use during the study.(88) The rates of discontinuation and measures of reduction are presented in Table 27.

Table 27. Discontinuations and reductions of concomitant medications in pooled analysis (mITT; 2 year analysis)

	August 2021 data cut		
	200 mg once daily (n=83)	**200 mg twice daily (n=82)**	**Combined 200 mg (N=165)**
CS reduction, n (%)	56 (67.5%)	57 (69.5%)	113 (68.5%)
Mean change in CS dose from baseline, %	-49.41%	-50.11%	-49.76%
CS discontinuation, n (%)	23 (27.7%)	19 (23.2%)	42 (25.5%)
CNI discontinuation, %	NR	NR	NR

CNI = calcineurin inhibitor; CS = corticosteroid; mITT = modified intent-to-treat population; NR = not reported
SOURCES: Kadmon Pharmaceuticals 2022(88)

High levels of response (Table 28) and reductions in CS use (Table 29) were also observed in the subgroup of patients who received ≥2 prior lines of therapy from the pooled phase 2 analysis.(88) This subgroup is aligned with the licensed indication for belumosudil.

Table 28. Results of pooled efficacy analysis (≥2 prior lines of therapy; 2 year analysis)

	August 2021 data cut		
	200 mg once daily (n=81)	**200 mg twice daily (n=75)**	**Combined 200 mg (N=156)**
Median time to response, weeks (range)	7.9 (3.7—80.1)	5.3 (3.7—40.1)	7.9 (3.7—80.1)
Best ORR,[a] n (%)	59 (72.8%)	55 (73.3%)	114 (73.1%)
Best ORR by organ system, n/N (%)			
Skin	20/67 (29.9%)	27/61 (44.3%)	47/128 (36.7%)
Eyes	24/60 (40.0%)	28/56 (50.0%)	52/116 (44.8%)
Mouth	20/41 (48.8%)	25/47 (53.2%)	45/88 (51.1%)
Oesophagus	11/21 (52.4%)	5/12 (41.7%)	16/33 (48.5%)
Upper GI	10/15 (66.7%)	5/11 (45.5%)	15/26 (57.7%)
Lower GI	4/7 (57.1%)	5/7 (71.4%)	9/14 (64.3%)
Liver	2/9 (22.2%)	1/4 (25.0%)	3/13 (23.1%)
Lungs	8/29 (27.6%)	6/24 (25.0%)	14/53 (26.4%)
Joints/fascia	42/62 (67.7%)	37/55 (67.3%)	79/117 (67.5%)
Median DOR in responders (primary/secondary)[b] weeks (95% CI)	22.1 (9.43, 40.00)	24.1 (12.43, 53.14)	24.1 (16.14, 36.14)
Median DOR in responders (quaternary), weeks (95% CI)	62.3 (28.29, 108.43)	74.3 (52.29, NA)	69.9 (40.43, 95.43)
Median FFS, months (95% CI)	13.7 (9.10, 24.02)	15.1 (9.59, NA)	14.8 (10.61, 20.73)
FFS, % (95% CI)			
FFS at 6 months	72% (61, 81)	77% (66, 85)	75% (67, 81)
FFS within 12 months	54% (42, 64)	58% (46, 68)	56% (48, 63)
FFS within 24 months	39% (28, 50)	39% (28, 50)	39% (32, 47)
Median OS (months)	NA (39.46, NA)	NA (NA, NA)	NA (NA, NA)
OS, % (95% CI)			
OS within 12 months	90% (81, 95)	90% (81, 95)	90% (84, 94)
OS within 24 months	84% (74, 91)	83% (73, 90)	84% (77, 89)
Median TTD, months (range)	9.2 (0.5—56.7)	11.2 (0.4—35.8)	10.2 (0.4—56.7)

[a] Best ORR at any time was defined as the percentage of patients that had either complete response or partial response, using the 2014 NIH Consensus Criteria
[b] DOR was measured from the time of first documentation of response to the time of first documented deterioration from best response (primary), to the time of first documented lack of response (secondary), or to the time of first documented lack of response with durations summed for multiple response/lack of response episodes (quaternary)
CI = confidence interval; DOR = duration of response; FFS = failure-free survival; GI = gastrointestinal; NA = not available; NIH = National Institutes of Health; ORR = overall response rate; OS = overall survival; TTD = time to treatment discontinuation
SOURCES: Kadmon Pharmaceuticals 2022(88)

Table 29. Discontinuations and reductions of concomitant medications in pooled analysis (≥2 prior lines of therapy; 2 year analysis)

	August 2021 data cut		
	200 mg once daily (n=81)	**200 mg twice daily (n=87)**	**Combined 200 mg (N=156)**
CS reduction, n (%)	55 (67.9%)	53 (70.7%)	108 (69.2%)
Mean change in CS dose from baseline, %	-49.72%	-52.24%	-50.95%
CS discontinuation, n (%)	23 (28.4%)	18 (24.0%)	41 (26.3%)
CNI discontinuation, %	NR	NR	NR

CNI = calcineurin inhibitor; CS = corticosteroid; NR = not reported
SOURCES: Kadmon Pharmaceuticals 2022(88)

B.2.6.3.2. Patient-reported outcomes

A combined analysis of PROs from the ROCKstar trial and the Phase 2a study was conducted to assess the correlation of improvements in overall PRO scores and organ-specific PRO scores with complete or partial clinical response measures.(90) This included the NIH chronic GVHD Consensus Conference measures and the LSS.

Organ-specific PROs for the skin, joints, mouth, GI tract, eye and lung were included in the response analysis, as well as overall response.(90)

The analysis established a strong positive correlation between the complete or partial clinical overall response or clinical organ-specific response and improvements in PRO symptom scores for most organs.(90) At least one PRO measure for each organ, except for the oesophagus, lower GI tract and joints, showed a clinically meaningful change that statistically correlated with a clinical organ response ($p<0.05$, Table 30).(90) This strong correlation between clinical efficacy outcomes and PROs suggests that the measures assess a similar type of response in the organ.(90) The results demonstrate that PROs used in the belumosudil Phase 2 studies (see Appendix M) capture valuable information on patients' perspective of disease activity in chronic GVHD and the impact of treatment.(90)

The poor correlation between clinical efficacy outcomes and PROs for the impact on joints, oesophagus and the lower GI tract, which are known to be a burden to patients (Section B.1.3.1), suggests that the PROs may be measuring a different aspect of these symptoms compared with the clinical assessment, or that one of the assessments may not be sufficiently sensitive to capture symptoms or changes in their burden.(90)

Table 30. Outcome correlation between clinical efficacy measures and PRO measures in ROCKstar and the Phase 2a study

System or Organ	Response Measure (Range)	Clinically Meaningful Score Change (Point[s])	PRO (Range)	OR for 1-point PRO Improvement Predicting NIH Response (95% CI; p-value)
Eye	NIH eye score (0-3)	1	LSS eye scale (0-100)	1.03 (95% CI: 1.02 to 1.05; **p<0.0001**)
	NIH eye score (0-3)	1	Worse eye complaint (0-10)	1.17 (95% CI: 1.04 to 1.32; **p=0.009**)
GI tract	*NIH oesophagus score (0-3)*	*1*	*LSS nutrition scale (0-100)*	*1.04 (95% CI: 0.99 to 1.08; p=0.105)*
	NIH upper GI score (0-3)	1	LSS nutrition scale (0-100)	1.06 (95% CI: 1.01 to 1.12; **p=0.026**)
	NIH lower GI score (0-3)	*1*	*LSS nutrition scale (0-100)*	*1.00 (95% CI: 0.96 to 1.04; p=0.934)*
Joints	*P-ROM (4-25)*	*2*	*LSS single item joints (0-4)*	*1.16 (95% CI: 0.88 to 1.53; p=0.281)*
	NIH joint score (0-3)	*1*	*LSS single item joints (0-4)*	*1.16 (95% CI: 0.90 to 1.50; p=0.244)*
Lungs	*FEV 1%*	*10%*	*LSS lung scale (0-100)*	*1.01 (95% CI: 0.97 to 1.04;*

System or Organ	Response Measure (Range)	Clinically Meaningful Score Change (Point[s])	PRO (Range)	OR for 1-point PRO Improvement Predicting NIH Response (95% CI; p-value)
				p=0.729)
	NIH lung score (0-3)	1	LSS lung scale (0-100)	1.05 (95% CI: 1.02 to 1.07; **p=0.001**)
Mouth	OM rating scale	2	LSS mouth scale (0-100)	1.04 (95% CI: 1.02 to 1.06; **p=0.0001**)
	OM rating scale	2	Mouth sensitivity (0-10)	1.35 (95% CI: 1.12 to 1.63; **p=0.002**)
Skin	NIH skin score (0-3)	1	LSS skin scale (0-100)	1.03 (95% CI: 1.00 to 1.05; **p=0.020**)
	Sclerotic skin (0-10)	*2*	*LSS skin scale (0-100)*	*1.01 (95% CI: 0.99 to 1.03; p=0.200)*
	Sclerotic skin (0-10)	2	Skin tightening (0-10)	1.19 (95% CI: 1.04 to 1.35; **p=0.012**)
Overall	Response vs. nonresponse	-	LSS summary scale (0-100)	1.04 (95% CI: 1.01 to 1.07; **p=0.006**)
	Response vs. nonresponse	-	*Overall chronic GVHD (0-10)*	*1.15 (95% CI: 0.99 to 1.34; p=0.072)*
	Overall severity (0-10)	2	LSS summary scale (0-100)	1.05 (95% CI: 1.01 to 1.08; **p=0.006**)
	Overall severity (0-10)	2	Overall chronic GVHD (0-10)	1.27 (95% CI: 1.03 to 1.56; **p=0.026**)

CI = confidence interval; FEV 1% = forced expiratory volume in 1 second; GVHD = graft-versus-host disease; LSS = Lee Symptom Scale; NIH = National Institutes of Health; OM = oral mucositis; OR = Odds ratio; P-ROM = photographic range of motion; PRO = patient-reported outcome
Note: Outcomes that were not statistically significant were formatted in *italics;* statistically significant p-values are presented in **bold**
Source: Adapted from Lee et al. 2022(90)

B.2.7. Subgroup analysis

As described in Section B.2.3, the pre-planned and post-hoc subgroup analyses described in Table 31 were undertaken for ROCKstar and the Phase 2a study. The population of patients with chronic GVHD who have received two prior systemic therapies is highly heterogeneous in terms of disease presentation and treatment pathway. Subgroups were therefore selected to explore the impact of key disease characteristics and prior treatment. An overview of the results in these subgroups is presented in Appendix E.

As Sanofi do not anticipate a restriction to a subgroup of patients, no further analyses than those presented were undertaken and no tests for interactions are necessary that do not apply to the overall patient population with chronic GVHD.

Please note, the trials were not powered to show significance between subgroups. All subgroup analyses were exploratory with no multiplicity adjustment.

Table 31. Comparative summary of trial subgroup methodology

Trial number (acronym)	ROCKstar (KD025-213; NCT03640481)	Phase 2a (KD025-208; NCT02841995)
Pre-planned subgroups	Severe chronic GVHD at screening (yes/no) Duration of chronic GVHD prior to enrolment (>50th percentile; ≤50th percentile) Number of organs involved at baseline (≥4; <4) Number of prior systemic LOTs (≥4; <4) Prior ibrutinib (yes/no) Receiving concomitant PPI* on C1D1 (yes/no)	None
Post-hoc subgroups	Best response to last systemic LOT Prior ruxolitinib (yes/no)	Belumosudil dose (200 mg once daily, 200 mg twice a day, 400 mg once daily) Refractory to prior line (yes; no) Prior lines (≥ 2; 1) Chronic GVHD severity at baseline (severe, non-severe) Number of organs involved at baseline (≥ 4; <4)

C1D1 = cycle one, day one; FFS = failure-free survival, GFR = glomerular filtration rate; GVHD = graft-versus-host disease; LOT = line of therapy; LSS
SOURCES: Adapted from Cutler et al. 2021,(73) Jagasia et al. 2021,(74) Kadmon Pharmaceuticals 2019,(50) and Kadmon Pharmaceuticals 2016(91)
* The subgroup analysis of patients receiving PPI on C1D1 was added in a protocol amendment as a result of a pharmacokinetic bioavailability study. The PK study showed that the assumed bioavailability of 1 in a fasted healthy individual was reduced by 48% for healthy individuals or patients with chronic GVHD who received concomitant PPIs and that absorption was delayed.(76) Further analyses showed that the maximal concentration (Cmax) and area under the curve (AUC) were reduced by 87% and 80%, respectively, for patients treated with a strong PPI and reduced by 68% and 47%, respectively for patients treated with a weaker PPI.(76) However, due to the flat exposure-efficacy relationship between belumosudil and PPIs across the evaluated exposure range, no dose adjustment is needed for the administration of belumosudil.(76)
†16 patients were planned for each of the cohorts

B.2.8. Meta-analysis

Beyond the pooling of the ROCKstar trial and the Phase 2a study described in Section B.2.6.3, no other data were available to assess the clinical effectiveness of belumosudil and as a result, no meta-analysis was conducted.

B.2.9. Indirect and mixed treatment comparisons

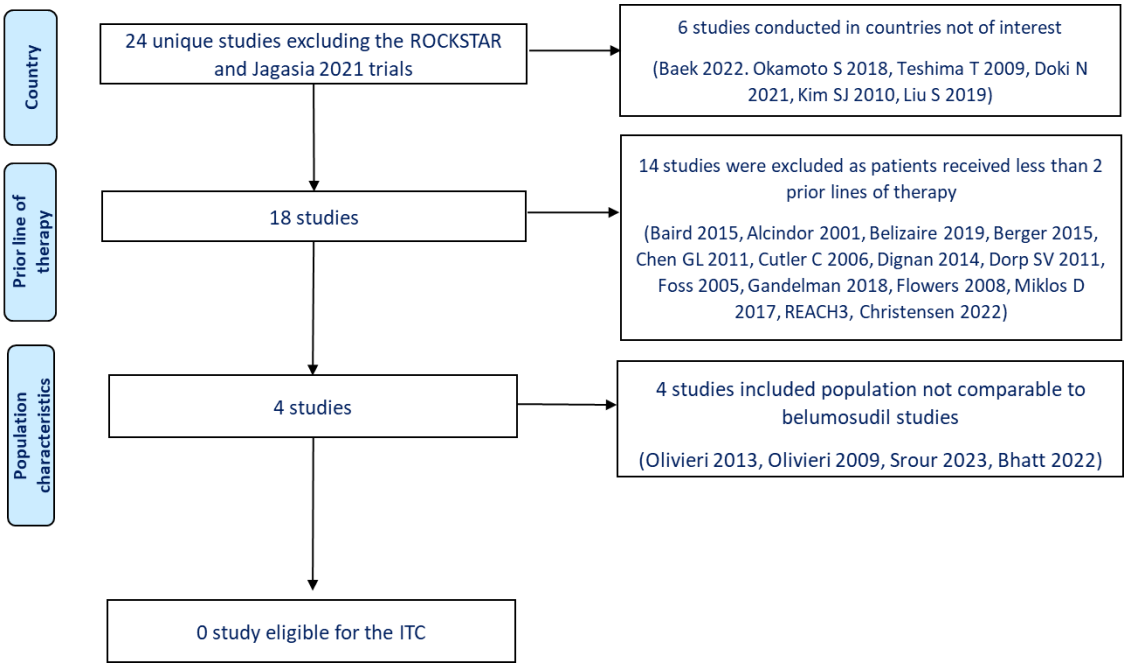
B.2.9.1. Feasibility assessment for indirect treatment comparison (ITC)

There is a paucity of robust clinical data for chronic GVHD treatments in the literature. The ROCKstar study of belumosudil provides some of the best available clinical efficacy and safety evidence for a treatment in this therapy area and is the data source from which the belumosudil marketing authorisation was granted. However, as an uncontrolled Phase 2 study in a heavily pre-treated (at least two prior systemic therapies) patient population, it does not enable direct comparison with other treatment options. Other data for treatments similarly positioned to belumosudil consist mainly of case series and small, non-controlled clinical studies (see Appendix M).

As described in Section B.2.1, we conducted an SLR in January 2023 to identify studies reporting on the clinical efficacy and safety of treatment options for adult patients with chronic GVHD after alloHSCT who have failed at least one prior line of therapy. To ensure an unbiased selection of evidence, the SLR was undertaken according to guidance from the National Institute for Health and Care Excellence (NICE) with respect to technology appraisal (TA) submissions, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and the Cochrane Handbook for Systematic Reviews of Interventions. All abstracts and full texts were examined by two independent researchers who applied a set of predefined inclusion and exclusion criteria (details provided in Appendix D). Results of the SLR are described in Section B.2.1 and Appendix D.

The criteria for which each trial was assessed and selected for inclusion in a potential indirect treatment comparison (ITC) is presented in Figure 13. Given the differences in population characteristics, outcome definitions and prior lines of therapy between the ROCKstar trial and the comparator trials, a robust statistical and methodological analysis is not possible and therefore it is not feasible to conduct an ITC for belumosudil (details provided in Appendix D).

Figure 13. Selection of trials for an indirect treatment comparison of belumosudil



ITC = indirect treatment comparison
Note, 5 studies conducted in Asian countries were excluded. Inclusion of such studies could create heterogeneity in patient populations and/or health systems. A recent article stated, there are differences by ethnicity in terms of affected organ sites, severity, and clinical outcomes for both acute GVHD and chronic GVHD. Thus, the 5 studies conducted in the Asian countries were not eligible to be included in the ITC as the ethnicity of the study populations would differ substantially from the ROCKstar included study population.

B.2.9.2. External control arm (ECA) study

[redacted]

[redacted]

B.2.10. Adverse reactions

Overall, belumosudil was well-tolerated and the AEs experienced in the treatment groups were consistent with those expected in a population of patients with chronic GVHD receiving CS and other immunosuppressants.(73, 74) The frequency of discontinuations due to possible drug-related AEs occurred in 12% of patients in ROCKstar and 5.6% of patients in the Phase 2a study.(73, 74)

In addition, to ensure all relevant safety evidence for belumosudil and potential comparator therapies was identified, systematic searches for randomised controlled trials (RCT) safety outcomes were carried out as part of the clinical SLR. Results are presented in Appendix F.

B.2.10.1. ROCKstar (KD025-213; NCT03640481)

In ROCKstar, 99% of patients experienced any AE; 60% experienced Grade ≥3 AEs, and 44% experienced SAEs.(93) AEs that occurred in ≥30% of subjects in the entire treated population were expected in this population and included fatigue (39%), diarrhoea (35%), nausea (31%) and cough (30%).(93) The overall safety profile for patients in the ROCKstar study is presented in Table 32.

Table 32. ROCKstar: Safety profile (safety population)

	August 2021 data cut		
	200 mg once daily (n=66)	**200 mg twice daily (n=66)**	**Total (N=132)**
Any AE, %	65 (99%)	66 (100%)	131 (99%)
Grade ≥3 AEs, %	41 (62%)	38 (58%)	79 (60%)
Drug-related AEs, %	50 (76%)	42 (64%)	92 (70%)
SAEs, %	30 (46%)	28 (42%)	58 (44%)
Deaths[a], n (%)	14 (21%)	14 (21%)	28 (21%)
AEs leading to deaths, n (%)	4 (6%)	5 (8%)	9 (7%)
Drug-related SAEs, %	6 (9%)	3 (5%)	9 (7%)

[a] Six patients died during long-term follow-up (>28 days after last dose)
AE = adverse event; SAE = serious adverse event
SOURCES: Adapted from Cutler et al. 2022(93) and Kadmon Pharmaceutical 2022(83)

At the one-year analysis (August 2020 cut-off), relative dose intensity (RDI), defined as the actual dose intensity divided by the planned dose intensity (where dose intensity was cumulative dose over

the duration of exposure [mg/d]) was used as a surrogate for measuring drug tolerability.(73) The median RDI for patients in ROCKstar was 99.7% overall (81% of patients had an RDI >95% at 1 year) and the median RDI was 99.5% at 2 years, indicating that patients were generally able to tolerate their planned dose of belumosudil.(73, 83)

As another measure of drug tolerability, the incidence of belumosudil dose reductions and interruptions was also reported for patients in the ROCKstar study. Two-year safety results (cut-off date: 19 August 2021) for ROCKstar demonstrated that, overall, 20% of patients experienced a dose modification and 10% of patients experienced a dose interruption on account of one or more drug-related AEs.(93)

B.2.10.2. Phase 2a study (KD025-208; NCT02841995)

In the Phase 2a study, 98% of patients experienced AEs; 61% experienced Grade ≥3 AEs and 43% experienced SAEs.(74) There were four deaths in the Phase 2a study, with none attributed to belumosudil.(74) AEs that occurred in ≥20% of subjects in the entire treated population were expected in this population and included upper respiratory infection (46%), diarrhoea (33%), nausea (33%), fatigue (33%), alanine amino transferase (ALT)/aspartate aminotransferase (AST) increased (33%), dyspnoea (30%), peripheral oedema (24%), headache (24%), cough (22%) and hypertension (20%).(74) The overall safety profile for patients in the Phase 2a study is presented in Table 33.

Table 33. Phase 2a: Safety profile (safety population)[a]

	200 mg once daily (n=17)	200 mg twice daily (n=16)	400 mg once daily (n=21)	Total (N=54)
Any AE, %	17 (100%)	16 (100%)	20 (95%)	53 (98%)
Grade ≥3 AEs, %	9 (53%)	10 (63%)	14 (67%)	33 (61%)
Drug-related AEs, %	8 (47%)	8 (50%)	14 (67%)	30 (56%)
SAEs, %	5 (29%)	6 (38%)	12 (57%)	23 (43%)
Deaths, %	0	0	2 (10%)	2 (4%)
Drug-related SAEs, %	0	0	0	0

[a] Data cut-off: February 2020
AE = adverse event; SAE = serious adverse event
SOURCES: Adapted from Jagasia et al. 2021(74)

The median RDI was 98% overall and the proportion of patients with an RDI >95% was 77% in patients treated with 200 mg once a daily, 63% in patients treated with 200 mg twice a daily, and 71% in patients treated with 400 mg once daily at 1 year.(74)

Overall, 9% of patients in the Phase 2a study had a dose reduction, with the median duration of reduction equal to 97 days (range: 21 to 859 days); 41% of patients had a dose interruption with a median duration of interruption of equal to 10 days (range: 2 to 39 days).(74)

B.2.10.3. Pooled Analysis of ROCKstar and Phase 2a

The overall safety profile from the pooled analysis of belumosudil is presented in Table 34. At two years, infections were observed in approximately two thirds of patients treated with belumosudil, with the majority being mild or moderate and nonserious.(94)

Table 34. Safety profile from pooled analysis (safety population; 2 year analysis)

	August 2021 data cut			
	200 mg once daily (n=83)	**200 mg twice daily (n=82)**	**400 mg once daily (n=21)**	**Total (N=186)**
Any AE, n (%)	82 (98.8%)	82 (100.0%)	20 (95.2%)	184 (98.9%)
Any drug-related AE, n (%)	60 (72.3%)	50 (61.0%)	14 (66.7%)	124 (66.7%)
Grade ≥3 AEs, n (%)	52 (62.7%)	48 (58.5%)	14 (66.7%)	114 (61.3%)
Drug-related Grade ≥3 AEs, n (%)	15 (18.1%)	14 (17.1%)	3 (14.3%)	32 (17.2%)
SAE, n (%)	36 (43.4%)	34 (41.5%)	13 (61.9%)	83 (44.6%)
Drug-related SAE, n (%)	6 (7.2%)	3 (3.7%)	1 (4.8%)	10 (5.4%)
Fatal AEs, n (%)	4 (4.8%)	5 (6.1%)	4 (19.0%)	13 (7.0%)
Infections and infestations (any grade), n (%) Grade ≥3, n (%)	55 (66.3%) 18 (21.7%)	55 (67.1%) 20 (24.4%)	15 (71.4%) 7 (33.3%)	125 (67.2%) 45 (24.2%)
Cytopenias[a]	15 (18.1%)	19 (23.2%)	3 (14.3%)	37 (19.9%)
Most common AEs (incidence ≥25%)				
Fatigue	37 (44.6%)	24 (29.3%)	10 (47.6%)	71 (38.2%)
Diarrhoea	33 (39.8%)	26 (31.7%)	7 (33.3%)	66 (35.5%)
Upper respiratory tract infection	26 (31.3%)	28 (34.1%)	7 (33.3%)	61 (32.8%)
Nausea	29 (34.9%)	22 (26.8%)	9 (42.9%)	60 (32.3%)
Dyspnoea	25 (30.1%)	20 (24.4%)	7 (33.3%)	52 (28.0%)
Cough	22 (6.5%)	22 (26.8%)	7 (33.3%)	51 (27.4%)
Oedema peripheral	22 (26.5%)	21 (25.6%)	6 (28.6%)	49 (26.3%)
Headache	21 (25.3%)	21 (25.6%)	6 (28.6%)	48 (25.8%)
Vomiting	21 (25.3%)	14 (17.1%)	4 (19.0%)	39 (21.0%)
Muscle spasms	14 (16.9%)	15 (18.3%)	6 (28.6%)	35 (18.8%)

[a] Anaemia, thrombocytopenia, neutropenia or leukopenia or cytopenia affecting more than 1 cell line
AE = adverse event; SAE = serious adverse event
SOURCES: Kadmon Pharmaceuticals 2022(94)

In order to inform the economic model (described in Section B.3), an additional analysis was performed to identify Grade ≥3 AEs occurring in patients who received ≥2 prior lines of therapy from the pooled analysis. Grade ≥3 AEs occurring in >5% of patients in either treatment arm are reported in Table 35.(95) A full analysis of safety data for patients who received ≥2 prior lines of therapy is not currently available.

Table 35. Grade ≥3 AEs occurring in >5% of patients in either treatment arm in the pooled analysis (≥2 prior lines of therapy; 2 year analysis)

	August 2021 data cut		
	200 mg once daily (n=81)	**200 mg twice daily (n=75)**	**Combined 200 mg (N=156)**
Pneumonia	7 (8.6%)	5 (6.7%)	12 (7.7%)
Hypertension	7 (8.6%)	5 (6.7%)	12 (7.7%)
Anaemia	4 (4.9%)	4 (5.3%)	8 (5.1%)
Hyperglycaemia	5 (6.2%)	4 (5.3%)	9 (5.8%)
Gamma-glutamyl transferase increased	5 (6.2%)	2 (2.7%)	7 (4.5%)
Fatigue	2 (2.5%)	4 (5.3%)	6 (3.8%)
Lung infection	1 (1.2%)	4 (5.3%)	5 (3.2%)

AE = adverse event
SOURCE: Kadmon Pharmaceuticals 2022(94)

B.2.11. Ongoing studies

There are two studies including belumosudil in the third-line for chronic GVHD which will provide additional data in the next 4 years:

1) Data collection for the ROCKstar study is still ongoing, with [redacted]. Long-term safety data from the ROCKstar study will be published in due course.

2) We plan to conduct a prospective, observational study to demonstrate the effectiveness and safety of belumosudil compared to BAT in real-world clinical practice globally post-approval. The planned study population will include patients 12 years of age and older with chronic GVHD who received belumosudil or BAT after failure of 2 to 5 prior lines of therapy. This observational study is planned to start in Q4 2023 in countries where the product is launched and would be extended to include other countries after the marketing authorisation application (MAA) is granted. The study is entitled: A Prospective, Observational Global Study of Belumosudil Compared to Best Available Therapy for Chronic GVHD in Patients Who Have Failed at least 2 Lines of Therapy. The estimated study report date is Q1 2027.

B.2.12. Interpretation of clinical effectiveness and safety evidence

B.2.12.1. Summary and interpretation of the evidence

The goals of chronic GVHD treatment are the effective control of symptoms and minimisation of the risk of toxicity and relapse.(47) However, treatment options for patients with chronic GVHD in England who have received at least two lines of systematic therapy are limited, and most are associated with high levels of toxicity.(47)

The most common treatment for chronic GVHD in England after two prior therapies is currently ECP, which requires most patients to travel to one of five main treatment centres in Bristol, Oxford, Manchester, Liverpool and London,(62, 96) thereby placing a substantial burden on patients, caregivers and the NHS. Currently available treatments for chronic GVHD are prescribed according to their organ-specific benefits; however, with chronic GVHD being a complex, heterogenous disease affecting multiple organs, there is a need for treatments which target the underlying causes and are not organ-specific.

Belumosudil is an oral, selective ROCK2 inhibitor that targets both the inflammatory and fibrotic processes of chronic GVHD and is intended as a monotherapy for patients who have received two or more prior therapies.(97) The use of belumosudil for the treatment of patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy is currently supported by two Phase 2 trials.(73, 74)

ROCKstar (KD025-213; NCT03640481): Patients included in the ROCKstar trial had severe chronic GVHD with fibrotic and inflammatory manifestations (e.g., GI and liver) and multi-organ involvement. Overall, patients achieved a best ORR of 76% (95% CI: 68 to 83), demonstrating that belumosudil has a clinically significant impact, including in fibrotic and inflammatory manifestations and difficult-to-treat organs such as the lung and liver.(83) Belumosudil is suitable for the treatment of a wide range of manifestations and led to a best ORR of 72.5% in the joints/fascia, 66.7% in the lower GI tract, 61.5% in the upper GI tract, 53.3% in the mouth, 52.6% in the oesophagus, 39.6% in the eye, 37.5% in the lung, 33.3% in the liver and 30.9% in the skin in patients receiving 200 mg belumosudil once daily.(83) Secondary endpoints in ROCKstar measured longer-term clinical benefits, such as OS and FFS. FFS was defined as the time from the first dose of belumosudil to the time of initiation of new systemic chronic GVHD therapy, non-relapse mortality or recurrent malignancy.(73) FFS is therefore a clinically meaningful and payer-relevant endpoint which captures chronic GVHD control (prevention or delay of the need for chronic GVHD treatment change), underlying disease control and survival information. As such, FFS is the key outcome of the ROCKstar trial used in the presented model (Section B.3.2.2).(82) FFS is recognised as a key endpoint for GVHD for regulatory authorities.(49) It is also a highly relevant endpoint from the patient point of view as it captures the time that a patient spends maintained at a particular line of therapy with an associated QoL, followed by movement to the next stage of their journey however that is characterised (new treatment, recurrent malignancy, or death).

CS remain the mainstay of chronic GVHD treatment, despite the toxicity associated with their longer-term usage, and are often used concomitantly with treatments such as ECP.(51) Reduction in CS use was documented as a secondary endpoint in the ROCKstar trial.(73) As a result of belumosudil treatment, 68% of patients were able to reduce their CS dose.(83) The mean reduction in CS dose from baseline was 50% in the mITT population.(83) ROCKstar also captured the impact of belumosudil on QoL and symptom bother, with LSS summary score assessed as an additional secondary endpoint and PROMIS Global Health 10 (PROMIS-GH) scores included as an exploratory

endpoint. Overall, 62.9% of patients had a clinically meaningful improvement in LSS score, and 47.7% and 50.0% of patients reported improvements in PROMIS raw mental and physical health scores, respectively.(83) PROMIS-GH can be mapped to EQ-5D to derive utility values for health economic modelling using a published algorithm, as described in Section B.3.4.2.(87)

Phase 2a study (KD025-208; NCT02841995): In the Phase 2a, open-label, dose-escalation study of belumosudil, the primary efficacy objective was to evaluate the activity of belumosudil at dose levels of 200 mg QD, 200 mg twice daily (BID), and 400 mg QD.(74) The majority of patients achieved a rapid response, with >75% of the patients who responded to treatment achieving a response by Week 8.(74) Best ORR was ≥50% in all subgroups in the Phase 2a study, including patients with ≥2 lines of prior therapy, patients with ≥4 organs involved at baseline, and patients with severe chronic GVHD.(74) The response was sustained for a median of 35 weeks, which increased to 38 weeks in patients with ≥2 prior systemic therapies.(74)

Across both trials, belumosudil was demonstrated to be well-tolerated and the reported AEs were those that are expected in a population of patients with chronic GVHD receiving CS and other immunosuppressants.(73, 74)

Although the belumosudil 200 mg twice daily dose showed higher responses in certain organs such as the skin (Section B.2.6) and slightly fewer drug-related AEs (Section B.2.10), the difference compared with the 200 mg once daily dose was not deemed significant.(83, 93) The MHRA licence for belumosudil is for 200 mg given orally once daily, except for patients who are receiving PPIs or strong CYP3A inducers where the dose should be increased to 200 mg twice daily.(1) English clinicians consulted at our advisory board estimated that 95% of patients would receive once daily belumosudil (i.e., not be receiving PPIs) due to the steroid sparing effect of belumosudil, the relatively late stage of the disease and cost consciousness.(11) Additionally, replacing PPIs with H2 agonists (e.g., famotidine) was generally seen as a desirable option to avoid PPI-related AEs and a suitable alternative to PPIs for nearly all patients.(11)

Based on the clinical and safety data from the phase II studies, belumosudil is expected to provide durable and sustained multi-organ responses while minimising treatment-limiting toxicities, infections, and other Grade 3-4 AEs that can lead to hospitalisations.(47) These clinical benefits are observed irrespective of organ involvement, prior treatment or duration of chronic GVHD prior to enrolment.(83)

B.2.12.2. Key uncertainties and/or evidence gaps

- **The ROCKstar study is a phase II study with no active control arm**

In the absence of a control arm and published data from which an ITC can be made (Section B.2.9.1), as well as methodological biases that could not be resolved in the ECA data (Section B.2.9.2), we use data from the Phase 3 REACH-3 trial of ruxolitinib vs. investigator's choice after one prior line of therapy to allow comparison to currently available treatments in the economic model through a naïve direct comparison (Section B.3.3).(58)

- **The patient populations included in the ROCKstar and Phase 2a trials are *US-based***

Data from the ROCKstar study and Phase 2a trial are expected to be applicable to patients in routine clinical practice in England, as described in Section B.2.5.2. English clinicians have confirmed the generalisability of the ROCKstar data to the patient population in England.(11)

- **HRQoL in the ROCKstar and Phase 2a study was assessed using the chronic GVHD-specific Lee Symptom Scale and PROMIS-GH. No other QoL measures are available for patients in these studies**

Utility-based HRQoL evidence from the ROCKstar and Phase 2a study are limited. EQ-5D data were not collected in the trials; however, LSS was a secondary endpoint and PROMIS Global Health 10 (PROMIS-GH) was included as an exploratory endpoint in ROCKstar. Due to the uncontrolled nature of the ROCKstar trial and Phase 2a study, these cannot be directly compared with a non-belumosudil treated population. The scarcity of recorded utility data following a failure event in the ROCKstar trial also prevented us from generating meaningful mean PROMIS-GH scores, and therefore utility estimates, in the failure state.

To try and address the lack of utility-based HRQoL evidence, we conducted a utility elicitation exercise within the UK general population (described in Appendix N). While the study highlighted the substantial burden associated with chronic GVHD after two or more prior lines of systemic therapy from the perspective of the general public, these utility values were not used for the reasons described in Section B.1.3.1.4 and Appendix N.

In order to provide utility data for the economic model, we used the mapping algorithm published by Thompson et al. to map PROMIS-GH outcomes from the ROCKstar study to EQ-5D-5L outcomes.(87) This approach provided data for the Failure Free (FF) health state in the model but not for the Failure state. Utility values were sourced from the literature for the failure state (including recurrence of malignancy and move to next treatment; Section B.3.4). In the context of the data from the utility elicitation exercise, these can be considered conservative assumptions.

- **Due to the different organs that may be affected in each patient and the expected heterogeneity in multi-organ manifestations, the sample size per organ and combination of manifestations is low in the Phase 2 trials and it is, therefore, difficult to generalise outcomes for the whole patient population**

The number of patients with involvement of each organ and combination of organ manifestations at baseline varies. Due to the ultra-orphan status of chronic GVHD, sample sizes in trials are limited.

In the ROCKstar study, the involvement at baseline of skin (55/66 patients), eyes (48/66 patients), mouth (30/66 patients), oesophagus (19/66 patients), upper GI tract (13/66 patients), lower GI tract (6/66 patients), liver (9/66 patients), lung (24/66 patients) as well as joint and fascia (51/66 patients) were tracked to evaluate the best organ response.(73) Overall, patients in ROCKstar achieved a best

ORR of 76% (95% CI: 68 to 83), demonstrating that belumosudil has a clinically significant impact, including in fibrotic and inflammatory manifestations and difficult-to-treat organs such as the lung and liver.(83) Belumosudil is suitable for the treatment of a wide range of manifestations and led to a best ORR of 72.5% in the joints/fascia, 66.7% in the lower GI tract, 61.5% in the upper GI tract, 53.3% in the mouth, 52.6% in the oesophagus, 39.6% in the eye, 37.5% in the lung, 33.3% in the liver and 30.9% in the skin in patients receiving 200 mg belumosudil once daily (Section B.2.6.1.1).(83)

B.3. Cost-effectiveness

B.3.1. Published cost-effectiveness studies

An SLR was conducted with a cut-off date of December 2022 to identify economic evaluations and cost-effectiveness studies of therapies for patients with chronic GVHD (details provided in Appendix G). A total of 3,403 abstracts were identified. These included 2,774 records via MEDLINE®, Embase, the NHS Economic Evaluation Database (NHS EED) and EconLit, and 341 additional records through conference proceedings not indexed within Embase: American Society of Transplantation (AST) 2019-2022, American Society of Hematology (ASH) 2021, and International Society for Pharmacoeconomic and Outcomes Research (ISPOR) 2019-2022. Moreover, 288 publications were identified through Health Technology Assessment (HTA) website searching.

Following title/abstract and full-text screening, seven studies that are relevant to this assessment were identified (Table 36): three conducted in the EU (France, Italy and Spain), and one each from Australia, Brazil, Canada and the US. No UK studies were identified. All employed the healthcare payer's perspective. Given the nature of the disease and importance of indirect costs and benefits, a societal perspective that includes, for example, lost productivity or cost to the patient and family would also be valuable. However, no study using this perspective or providing information such as indirect costs was identified in this SLR.

Table 36. Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)			QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
HAS	2022	Markov model comparing ruxolitinib versus BAT (rituximab, ECP, imatinib, methotrexate, MMF, everolimus, sirolimus, ibrutinib, infliximab) over 5 years from French national health system perspective	Chronic GVHD with inadequate response to corticosteroids or other systemic treatments (patients aged ≥12 years; average age 46.5 years)	Ruxolitinib: 2.97 BAT: 2.82	Ruxolitinib: €148,275 BAT: €137,712	Ruxolitinib versus BAT: €66,365/QALY gained		
Crespo et al.	2012	Microsimulation model comparing ECP versus rituximab and versus imatinib at year 5 from Spanish NHS perspective	Chronic GVHD (average age NR)			ECP: 3.335 Rituximab: 3.273 Imatinib: 3.240	ECP: €85,700.66 Rituximab: €85,182.83 Imatinib: €87,438.76	ECP versus rituximab: €8,330.16/QALY gained ECP versus imatinib: dominant
De Waure et al.	2015	Markov model comparing ECP versus pentostatin, versus mycophenolate and versus imatinib over 7 years from Italian NHS perspective	Steroid-refractory/resistant chronic GVHD (average age NR)			ECP: 4.17 Pentostatin: 3.96 Mycophenolate: 4.13 Imatinib: 4.10	ECP: €95,770.36 Pentostatin: €115,673.87 Mycophenolate: €100,284.23 Imatinib: €99,007.45	ECP dominated all comparators
CADTH	2022	Semi-Markov model comparing ruxolitinib versus BAT (rituximab, ECP, imatinib,	Steroid-refractory chronic GVHD (patients aged ≥12 years; mean age [SD] 46.5 years [15.92])			Base case Ruxolitinib: 8.00 BAT: 7.19 CADTH reanalysis	Base case Ruxolitinib: 318,305 BAT: 323,550	Base case Ruxolitinib dominated BAT

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
		methotrexate, MMF, sirolimus, ibrutinib), based on REACH-3, over lifetime horizon (40-year) from Canadian public healthcare payer perspective		Ruxolitinib: 6.59 BAT: 6.49	CADTH reanalysis Ruxolitinib: CAD 304,468 BAT: CAD 198,291	CADTH exploratory reanalysis CAD 1,062,977/QALY
Okumura et al.	2019	Markov model comparing ECP versus MMF, versus sirolimus, versus rituximab and versus imatinib over 1 year from Brazilian perspective	Refractory chronic GVHD (average age NR)	QALYs NR Life-months gained ECP: 11.55 MMF: 11.03 Sirolimus: 10.99 Rituximab: 10.5 Imatinib: 11.2	ECP: USD 85,757 MMF: USD 102,284 Sirolimus: USD 89,138 Rituximab: USD 110,859 Imatinib: 84,689	ECP considered dominant versus comparators
Peacock	2022	Markov model comparing ECP versus SoC (tacrolimus, ciclosporin, MMF) over 10 years from Australian healthcare system perspective	Steroid-refractory chronic GVHD (average age NR)	Incremental gain 1.10 (ECP over SoC)	Decreased AUD 24,006 (ECP over SoC)	ECP dominated SoC
Yalniz et al.	2018	Cost-effectiveness analysis (6-month cost per response) based on meta-analysis of response rates for various	Steroid-refractory chronic GVHD (average age NR)	QALYs NR Response rate Tacrolimus: 30% Sirolimus: 77% Rituximab: 62% Ruxolitinib: 85% HCQ: 32%	Tacrolimus: USD 6,815 Sirolimus: USD 5,731 Rituximab: USD 29,184 Ruxolitinib: USD 83,136	Cost per response Tacrolimus: USD 22,717 Sirolimus: USD 7,443 Rituximab: USD 47,071

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
		treatments in US studies		Imatinib: 46% Bortezomib: 50% Ibrutinib: 33% ECP: 62% Pomalidomide: 78% Methotrexate: 45%	HCQ: USD 2,938 Imatinib: USD 20,224 Bortezomib: USD 46,152 Ibrutinib: USD 79,938 ECP: USD 41,778 Pomalidomide: USD 104,580 Methotrexate: USD 204	Ruxolitinib: USD 97,807 HCQ: USD 9,181 Imatinib: USD 43,965 Bortezomib: USD 92,304 Ibrutinib: USD 242,236 ECP: USD 67,400 Pomalidomide: USD 134,077 Methotrexate: USD 453

AUD = Australian dollars; BAT = best available therapy; CAD = Canadian dollars; CADTH = Canadian Agency for Drugs and Technologies in Health; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; HCQ = hydroxychloroquine; ICER = incremental cost-effectiveness ratio; MMF = mycophenolate mofetil; NHS = national health service; NR = not reported; QALY = quality-adjusted life year; SoC = standard of care

B.3.2. Economic analysis

The SLR reported in Section B.3.1 and Appendix G indicated that there is very limited literature examining the cost-effectiveness of treatments for chronic GVHD on which to base our economic model. It is noteworthy that the source of effectiveness stated in all but two studies (CADTH 2022, HAS 2022) was captured from the literature and not from a single effectiveness study. Only three studies (Crespo 2012, CADTH 2022, HAS 2022) provided justification for the choice of model used and the key parameters on which the model was based.(98-100) Okumura 2019 provided outline details of a Markov based model with four transition states (stable disease, overall response, progression and death) considering ECP in the second-line setting with a Brazilian perspective. However, results were presented in a conference abstract, and no full-text publication is available for this study.(101) Yalniz 2018 is a partial evaluation which provides details of a study conducted in the US to assess the cost per response (ORR, CR, and PR) for frequently used agents in steroid-refractory/chronic GVHD.(102) However, direct non-medical and indirect costs were not included in the analyses and no costs were included related to treatment or medical management of patients other than the comparator treatments.(102) No OS or quality adjustment of response health states were included. The time horizon was 6 months and results were expressed as cost of treatment per response state. Further, considering the similarity in HTA requirements between the UK, EU-4, Canada and Australia, the information most relevant to support this submission comes from three EU based studies (Crespo 2012, De Waure 2015, HAS 2022) and the Australian and Canadian studies (CADTH 2022, Peacock 2022). The Australian study was a conference abstract and did not report several key methodological details, and the source of cost data was not reported.(103) Both the CADTH 2022 and HAS 2022 studies assessed ruxolitinib as the intervention treatment using data from the REACH-3 trial. However, ruxolitinib is not considered a relevant comparator for this assessment (Section B.1.3.2.2), and none of the analyses in these studies were in patients who had received two or more lines of systemic therapy.(99, 100)

Two of the other models taking an EU setting are closely linked. De Waure 2015 is an adaptation of the decision tree model, associated with a Markov model developed by Crespo to the Italian setting and uses many of the inputs from Crespo.(104) The model reported by Crespo was designed to assess the cost-effectiveness of ECP compared with rituximab and with imatinib in Spanish patients with chronic GVHD at 5 years.(98) This study took the perspective of the Spanish National Health System and was a decision tree, associated with a Markov microsimulation over a 5-year time horizon.(98) The published model attempted to reflect disease complexity by considering up to 5 multiple organ systems all of which can respond or progress independently.(98) This complexity along with the paucity of published evidence to populate the health states meant that clinical opinion was used extensively to parameterise this model. The model developed by Crespo was considered to be a potential candidate for development in the English setting. However, the experience of Crespo has shown that an organ-based model requires disaggregated data which are not available in the literature.(98) The belumosudil studies do represent the best available data for the decision problem

in the third-line treatment setting but the patient numbers in the studies and the single-arm design mean that the evidence needed to reflect disaggregated individual and combination of organ responses is very limited. Similarly, in the absence of published data a significant number of assumptions would have to be made in order to develop a matching disaggregated comparator dataset.

We have therefore chosen to take a different, simplified approach and to develop a *de novo* partitioned survival model which is able to make best use of the available data while still reflecting the important elements of the disease and its progression including time to response (TTR) and DOR. The model is described in detail below.

B.3.2.1. Patient population

The population included in our cost-effectiveness model (CEM) is patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy. This reflects the indication for belumosudil as described in the scope and decision problem (Section B.1.1), the Summary of Product Characteristics as well as the patient population in the Phase 2 ROCKstar trial.(1, 73)

B.3.2.2. Model structure

B.3.2.2.1. Introduction

The CEM was developed in Microsoft Excel® using a three-state partitioned survival (or 'area under the curve') structure in both a deterministic and probabilistic framework. A cohort of patients was followed for the remainder of their lifetime through 4-week cycles, corresponding to the treatment cycle length in the ROCKstar study protocol.

To account for the progressive nature of chronic GVHD and potential relapse of malignancies, a partitioned survival structure was considered the most appropriate approach (see Section B.3.2.2.5 for detailed justification), with patients passing through a series of clearly defined and mutually exclusive health states. The model incorporates three health states: Failure-free (FF), Failure and Death, and also considers response outcomes as patients in the FF state are distributed into different response states according to the level of response achieved. These health states are discussed below. Costs and utilities were subsequently assigned to each health state.

B.3.2.2.2. Value of FFS to this assessment

The efficacy of belumosudil was measured in two key ways in our Phase 2 trials:

- ORR, CR and PR at any time
    - ORR was the primary end point in Phase 2 studies, with CR and PR both secondary endpoints.

- FFS
    - This is a composite measure endorsed by the NIH, defined as the interval between the start of treatment and the addition of a new chronic GVHD therapy, relapse, or non-relapse mortality.(49)
    - This was a secondary endpoint in the Phase 2 studies.

FFS was not the primary endpoint in the belumosudil studies but it was selected as the central endpoint to capture disease progression within the model. In line with the regulatory definition of FFS(49) it was defined in the ROCKstar trial as the time between the start of belumosudil and one of the following (whichever occurred first):(73)

- Initiation of a new chronic GVHD systemic treatment,
- Recurrent malignancy, or
- Non-relapse mortality.

We have chosen FFS over the primary endpoint of ORR for several reasons. Previous studies have established that FFS is a simple yet robust and valuable endpoint, that correlates with overall clinical improvement of chronic GVHD.(49) Moreover, FFS incorporates chronic GVHD disease control (i.e., prevention or delay to the need for chronic GVHD treatment change), absence of underlying malignancy and survival information into a single composite endpoint.(82) Capturing the absence of treatment change as well as clinical benefit makes FFS a useful measure which may more accurately reflect real-world practice and is highly relevant from a payer perspective. FFS is recognised as a key endpoint for GVHD by regulatory authorities.(73) It is also a highly relevant endpoint from the patient point of view as it captures the time that a patient spends maintained at a particular line of therapy with an associated QoL, followed by movement to the next stage of their journey, however that is characterised (new treatment, recurrent malignancy, or death).

B.3.2.2.3. Health states and response categories

The three health states included in the model are:

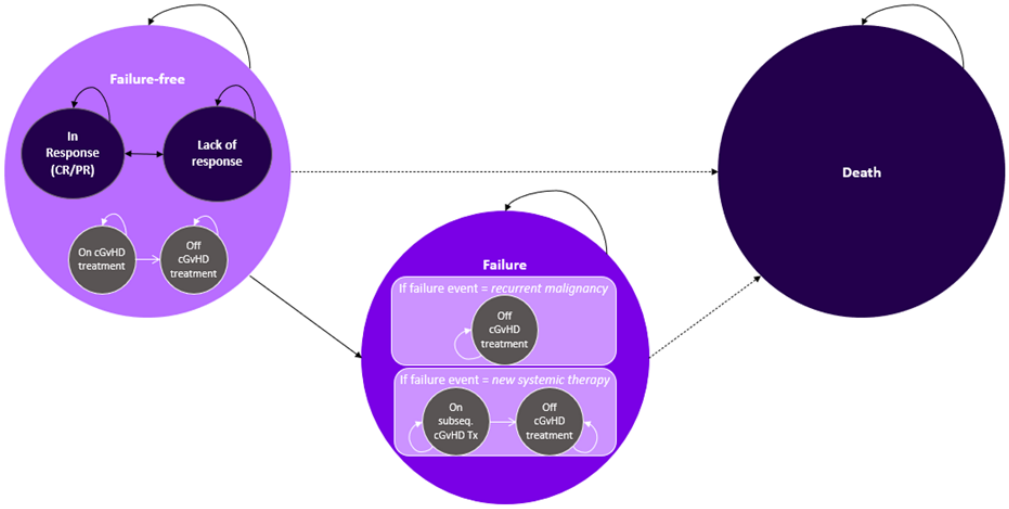
- **Failure-free** (FF): the FF state includes patients who are alive and have not experienced a failure event.
    - Patients in this state are assumed to incur costs associated with treatment (including treatment acquisition costs and costs of drug administration), costs associated with medical management of the condition (i.e., healthcare resource use), and costs associated with management of treatment-related Grade 3 or higher (Grade ≥3) AEs. Patients also experience a higher utility score compared with post-failure disease. Grade ≥3 AEs associated with treatment may impact QoL and result in a reduction in utilities for patients in this health state. Healthcare resource use costs and utilities in the FF state are specified separately for each of the response categories. The impact of response on survival was not explicitly modelled and assumed to be implicitly captured within the survival curves for all patients.
    - Within the FF state, patients can have CR, PR, or lack of response (LR).

- Patients receive their initial treatment while in the FF state and may receive this treatment until failure or until treatment discontinuation due to any reason (modelled using TTD curves) or have reached a maximum treatment duration. When TTD is used to model the time spent on initial treatment, patients who discontinued this treatment for a reason other than failure are also part of the FF health state but stop accruing treatment acquisition and administration costs for the remainder of the time spent in the FF health state.

- **Failure**: this state includes patients who are alive and have experienced a (non-fatal) failure event. Patients in this health state are further separated (in terms of costs and utilities) according to the cause of their failure event: either a recurrence of their malignancy or initiation of a new chronic GVHD systemic therapy.
  - Patients in this state are assumed to incur costs associated with medical management of chronic GVHD (i.e., healthcare resource use). In addition, for patients whose failure event was initiation of a new chronic GVHD systemic therapy, costs of subsequent lines of therapy (acquisition and administration costs) following failure are included in this economic evaluation. Patients whose failure event was recurrence of their malignancy are assumed not to receive further treatment for chronic GVHD. For these patients, the cost of treating the recurrent malignancy is also included in the economic evaluation.
- **Death**: this is an absorbing health state. Transition to the Death state may occur from either the FF or Failure health states.

The model diagram displayed in Figure 14 illustrates the health states, response categories and flow of patients in the model. As our model takes a partitioned survival structure, the elements are not structured as true health states with explicit transition probabilities for each of these elements (as they would be in a Markov model).

Figure 14. Model structure



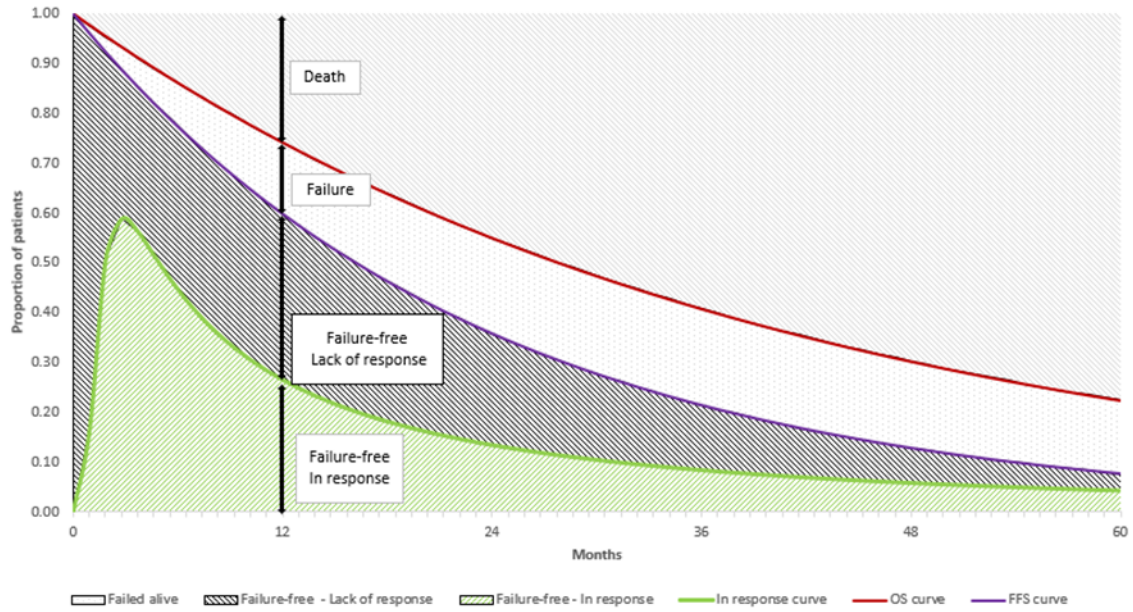
cGVHD = chronic graft-versus-host disease; CR = complete response, PR = partial response; subseq = subsequent; Tx = treatment
Note: Transition probabilities were not estimated explicitly, and an area under the curve approach was used to determine the proportions of patients in the Failure- free – In Response, Failure-free – Lack of response, Failure, and Death health states.

B.3.2.2.4. Implementation of health states over time

Our model was populated by fitting parametric survival curves for FFS, OS, DOR and TTR (e.g., to the belumosudil Phase 2 studies patient-level data for belumosudil and when available to REACH-3 data for the BAT comparator; Section B.3.2.3). The proportions of patients in the FF and Death states over the course of the model time horizon were estimated from these extrapolated survival curves for FFS and OS. By calculating the area under the survival curves at each cycle, the distribution of the cohort of patients between the different health states defined by these curves could be estimated. An "In response" curve was calculated from the TTR and DOR curves, and is used to separate patients who are "In response" (i.e., CR and PR) from those with LR.

The approach to calculating the proportions of patients in the three health states (FF, Failure, Death) and with different response levels at each point in time is illustrated in Figure 15. The area underneath the OS curve determines the proportion of patients who are still alive over time, while the proportion of patients in the FF state was calculated as the area under the FFS curve. This means that the FFS and OS curves were modelled independently of each other. The proportion of patients in the Failure state was calculated as the area between the FFS and OS curves at a given point in time. Due to the overlapping definitions of FFS and OS, non-relapse mortality is not modelled explicitly; instead, these death events are assumed to be captured by the OS curve.

Figure 15. Partitioned survival model approach



FFS = failure-free survival; OS = overall survival

All patients enter the model in the FF health state and initiate the assigned treatment (belumosudil or BAT comparator, see Section B.3.2.3) at model start. As discussed above, treatment-specific FFS and OS curves are used in the model to partition patients in the FF, Failure, and Death health states over time. Based on these curves, at the end of each cycle, patients in the FF state may remain in the FF state or move to the Failure state or the Death state. Patients in the Failure state may either

remain in the Failure state or move to the Death state. Once in the Death state, patients remain in that state for the remainder of the model time horizon.

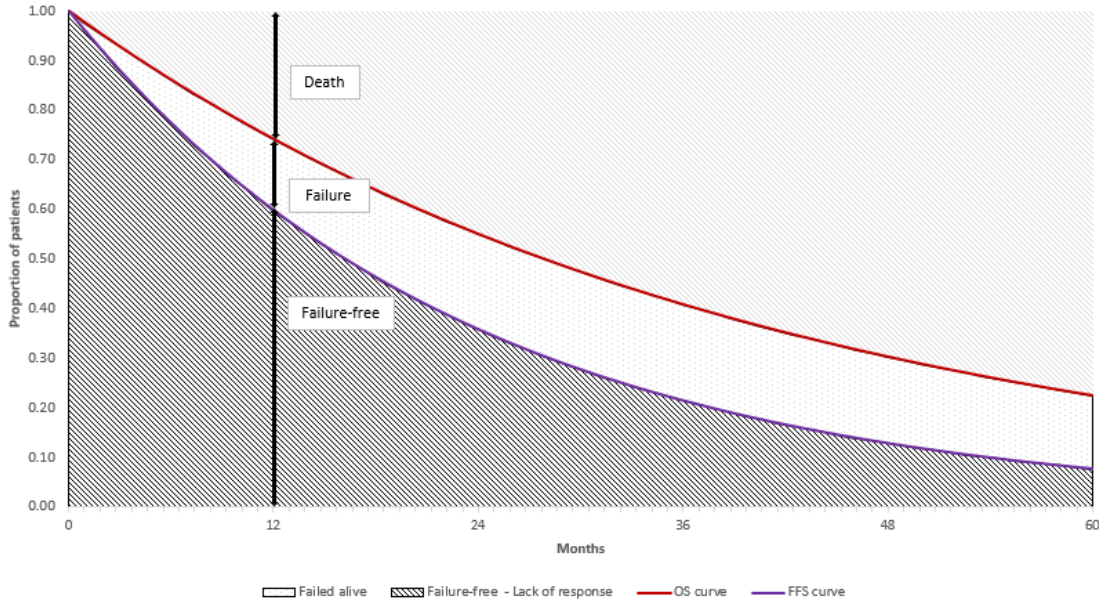
Given the uncertainty regarding the long-term benefit of belumosudil, the model conservatively assumes that after 5 years, patients with belumosudil are exposed to the same mortality risk as those who receive BAT (i.e., same cycle probability of death as the BAT arm post-5 years). This assumption was validated by clinicians in our advisory board (Section B.2.3.3) and is tested in a sensitivity analysis.

Patients entering the Failure state in each model cycle are stratified into one of the two reasons for non-fatal failure: initiation of new systemic chronic GVHD therapy or recurrent malignancy (i.e., relapse of underlying disease). Costs and utilities are assigned to patients accordingly. The distribution of failure events across these two causes of failure varies over time in the model, in line with the trends observed in the clinical trials. The third potential cause of failure, non-relapse mortality, is not modelled explicitly; instead, these death events are assumed to be captured by the OS curve. This is in line with the definition of the Failure state in the model, which is calculated as the difference between the OS and FFS curves and therefore only comprises patients who are under the OS curve, i.e., alive.

Being on or off chronic GVHD treatment is not modelled via explicit health states, i.e., treatment status does not explicitly impact survival or other health outcomes in the model. However, treatment status is tracked for the purpose of treatment-related cost assignment. Patients in the FF health state may discontinue treatment (e.g., due to an AE or for other reasons) at any model cycle. The proportion of patients in the FF health state who are on- versus off-treatment is determined by the time on treatment curve. The latter is capped by the time patients spend in the FF state, which implies that all patients entering the Failure state are off the chronic GVHD treatment they initiated at model start. Patients whose cause of failure is new chronic GVHD systemic therapy receive subsequent chronic GVHD treatment until they move to the Death state or until the end of the time horizon, whichever is earlier. The model design does not explicitly capture efficacy of subsequent treatments after discontinuation from initial therapy. The survival benefit attributable to subsequent treatments is assumed to be implicitly captured by the initial treatment-specific OS estimates (as discussed above). The model applies a maximum treatment duration of 5 years for the majority of treatments, except rituximab with a treatment course duration of 4 weeks. This assumes that all remaining patients stop treatments after 5 years. This is supported by the expert opinions of English clinicians who agreed that no treatment would be continued for more than 3 to 5 years for a patient remaining in the FF state.(11) If patients have persisted in this state for as long as this, the clinicians we consulted felt the remaining patients represent an enriched cohort of responders who would very likely have ceased treatment due to physician advice or patient preference. Indeed, it could be the case that for a small number of patients their chronic GVHD resolves within this time period. This is also reflected in the reduction in resource use associated with this remaining group (see Section B.3.5.3.1).

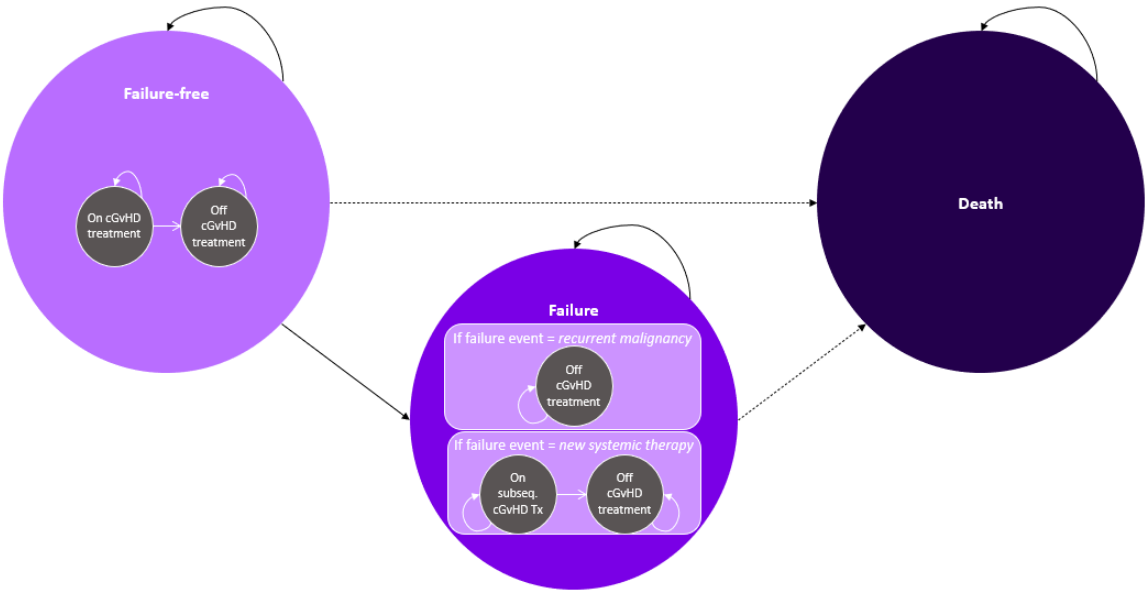
The model also allows for selection of an economic evaluation which does not consider response. This is a simplified option where patients in the FF state are not distinguished by response levels, and incur the same costs and utilities. A schematic representation of the model diagram (Figure 16) and model structure without incorporating response levels (Figure 17) are presented below.

Figure 16. Model diagram of the partitioned survival model approach without response



FFS = failure-free survival; OS = overall survival

Figure 17. Model structure without response



cGVHD = chronic graft-versus-host disease; subseq = subsequent; Tx = treatment
Note: Transition probabilities were not estimated explicitly, and an area under the curve approach was used to determine the proportions of patients in the Failure-free, Failure, and Death health states.

B.3.2.2.5. Justification of the chosen structure

The strengths of the partitioned survival approach are well-documented.(105) The partitioned survival approach was chosen based on the need to capture clinically important outcomes for patients with chronic GVHD (i.e., FFS and OS) and which can be directly obtained from the belumosudil Phase 2 studies and importantly, from the published trial(s) for the comparator. This structure enables the expected clinically important differences in costs and outcomes among patients in failure-free and failure states to be captured. In addition, this approach allows patients in the failure-free state to be segregated by response level, in turn allowing estimations of differences in costs and health outcomes across specific response levels.

An advisory board was held between 20th and 27th January 2023 with 9 English clinicians with expertise in chronic GVHD and one health economist (described in Section B.2.3.3).(11) The experts were consulted on the model structure, underlying assumptions, data sources and inputs and were in agreement with our approach.(11)

B.3.2.2.6. Model features

The model base case employed the perspective of the English payer, i.e., NHS England, which only included direct costs and benefits.

A summary of the features of the analysis is presented in Table 37. Chosen values could not be compared with previous NICE evaluations as none have been published in the same disease area.

Table 37. Features of the economic analysis

Factor	Chosen values	Justification
Model structure	Partitioned survival	Accounts for progressive nature of disease and potential relapse and makes best use of the available data.
Time horizon	40 years	Considered long enough to capture effects/benefits of belumosudil over the life expectancy of patients with chronic GVHD, given the average age of the population in the pooled ROCKstar and Phase 2a trials was approximately 55 years.(83)
Discount rate	3.5%	NICE guidance
Model cycle length	4 weeks	• Short enough to accurately capture differences in cost or health effects between cycles • Aligned with the data collection and reporting in ROCKstar and the pooled ROCKstar and Phase 2a trial • Treatment schedules of comparators can be easily considered • A half-cycle correction was applied to prevent under- or overestimation of costs and QALYs
Central model endpoint	FFS	• Robust and valuable endpoint, that correlates with overall improvement of chronic GVHD • Incorporates disease control (prevent/delay treatment change), absence of underlying malignancy, and survival information • Recognised as a key GVHD endpoint by regulatory authorities(73)
Treatment waning effect	5 years	Conservative assumption. OS treatment benefit associated with belumosudil wanes at five years following treatment initiation and adopts hazard of the BAT arm, to account for uncertainty in long-term treatment benefit of belumosudil

Factor	Chosen values	Justification
Maximum treatment duration (for the modelled initial treatments, except rituximab)	5 years	Maximum treatment duration applies in line with feedback from the advisory board that patients with chronic GVHD who are stable and responding to treatment are unlikely to be on treatment beyond 5 years.(11)

BAT = best available therapy; GVHD = graft-versus-host disease; NICE = National Institute for Health and Care Excellence; OS = overall survival; QALY = quality-adjusted life year

B.3.2.3. Intervention technology and comparators

Belumosudil was administered 200 mg orally once or twice daily in the ROCKstar trial (Section B.2), and therefore the model includes these dosage options. Based on the SmPC, the recommended dose of belumosudil should be increased to 200 mg twice daily when co-administered with strong CYP3A (cytochrome P450, family 3, subfamily A) inducers or PPIs.(1) In an advisory board held between 20th and 27th January 2023 with 9 English clinicians, the consensus was that a low proportion of patients would continue on PPIs due to the steroid sparing effect of belumosudil, the relatively late stage of the disease and cost consciousness.(11) Additionally, replacing PPIs with H2 agonists (e.g., famotidine) was generally seen as a desirable option to avoid PPI-related AEs and a suitable alternative to PPIs for nearly all patients.(11) The advisors estimated that at steady state the proportion of patients on PPIs would be around 5%.(11) The model therefore assumed in the base case that 95% of patients were on the belumosudil 200 mg once daily dose and 5% of patients were on the 200 mg twice daily dose. This assumption is tested in sensitivity analysis. The costs of PPIs (or PPI alternatives) and CYP3A inducers were not modelled as they would be similar with all treatments.

Due to the heterogenous nature of chronic GVHD and the prescription of different medicines according to manifestation and disease stage, it is appropriate to consider and model BAT as a basket of therapies. The list of comparator treatments included within BAT was developed based on a review of the BSH/BSBMTCT 2012 treatment guidelines(47), the 2017 Clinical Commissioning Policy issued by NHS England(48) and consultation with clinical experts in England regarding the treatments currently used at third-line.(11) Further information on current treatments for chronic GVHD and clinical practice in England is provided in Section B.1.3.2. The treatments included as part of the BAT comparator basket are listed below (details of the treatment schedule/dosage used in the model are provided in Appendix N):

- ECP
- MMF (off-label use only)
- Imatinib (off-label use only)
- Sirolimus (off-label use only)
- Rituximab (off-label use only)

In the absence of head-to-head RCTs comparing the efficacy of belumosudil to other treatment options, a feasibility assessment was conducted to explore whether an ITC could be used to estimate the comparative efficacy of belumosudil against BAT (Section B.2.9 and Appendix D). As the network

of evidence identified by the SLR was not connected, an ITC was not feasible. Population-adjusted indirect treatment comparisons, such as matching-adjusted indirect treatment comparisons, were also not feasible due to differences in study designs, in particular the inclusion/exclusion criteria of the populations of patients with chronic GVHD.

As indirect comparisons were not feasible, we constructed an external control arm (ECA) from a study using real-world data from the US Optum Clinformatics® Data Mart database to try and provide a comparison with the results of the Phase 2 belumosudil trials (described in Appendix M).(92) However, due to methodological biases that could not be resolved, the results from the ECA study could not be used to inform the model. The main issues with the collected data were in the coding of relapse and the very extended time to next treatment leading to a lack of face validity in the results. The rationale for this is explored in Appendix M.(92)

The recently published REACH-3 trial, a Phase 3 study for ruxolitinib compared with BAT in second-line only patients, does offer a source of comparator data albeit in an earlier line of treatment (patients who had received two or more systemic treatments were excluded). Therefore, in the absence of a more robust data source, efficacy data for the basket of comparators were sourced from the REACH-3 trial. The economic evaluation performed was based on naïve direct comparison, as an ITC was not feasible.

We are aware of the weaknesses inherent in this approach to the definition and construction of a comparator arm and so we consulted with English clinicians at the advisory board in January 2023 to further validate the choice of the REACH-3 BAT.(11) Two key issues were discussed to address the uncertainty around the naïve nature of the comparison:(11)

1. Patients in REACH-3 were at an earlier line of therapy than those in ROCKstar (second vs. third-line+)

2. The constitution and proportion of treatments across BAT in REACH-3 is not the same as that modelled for UK clinical practice at third-line.

There was general agreement among the clinicians that we engaged that, in the absence of other evidence, REACH-3 could be used as a proxy for BAT and would be appropriate to use in the modelling with some caveats.(11) For example, the expected position for belumosudil in clinical practice is at third-line and beyond and these patients might be expected to have worse prognosis than their REACH-3 counterparts who are at second-line. The advisors agreed this makes the use of the FFS curve from REACH-3 a conservative choice.(11) They told us that their experience of real-world outcomes at this advanced stage (third-line and beyond setting) suggests current treatments perform worse in later lines. FFS may be longer for the second-line patients in REACH-3 and the differential between FFS in ROCKstar and REACH-3 BAT less than the corresponding difference in outcomes that might be expected from a third-line and beyond comparator.(11) The advisors felt that on balance FFS at third-line at 1 year with BAT in the real world would be <20% and at 2 years would

be <10%.(11) This contrasts with the FFS data from REACH-3 where ~30% of patients persist in the failure-free state at one year and ~20% at two years (see Figure 21 in Section B.3.3.1). The advisors also felt that the trajectory of movement from FFS to the failure state would be faster for third-line patients in the first few months after initiating BAT.(11)

It was acknowledged that the difference in the constitution and proportion of treatments across BAT in REACH-3 and clinical practice in England may have an influence.(11) For example, the absence of ibrutinib in England which may have the effect of depressing the initial part of the curve. Other advisors suggested that, at this late stage of the disease, patients move between treatments relatively frequently and because this is a component of FFS, the curve could still be adjusted downwards.(11) They also noted mortality is higher at later lines.(11)

On balance the advisors agreed that an FFS curve based on English BAT in the real world would very likely follow a similar trajectory to the BAT arm of the REACH-3 study, but would be somewhat below it.(11)

These issues were explored in a discussion with the external assessment group (EAG) during the NICE checkpoint meeting on 15th February 2023. We are grateful to the EAG for their advice on this matter. It was recognised that, where there is no direct head-to-head comparison, all efforts to provide a counterfactual arm will result in an unanchored and often unmatched dataset. Data for the belumosudil patients in our modelling are taken from the clinical trials and not from the real world, which is the case for the ECA that we attempted to construct (Section B.2.9.2). The EAG suggested that given the issues with, and origin of the ECA data, the use of the REACH-3 BAT is likely to be the best option for the reasons discussed above and because it is derived from the protocol-driven setting of a clinical, trial which will help to avoid the biases inherent in data collected from the real world.

Therefore due to the absence of a head-to-head study and published data from which an ITC can be made, as well as methodological biases that could not be resolved in the ECA data (Section B.2.9), we have chosen to use data from the REACH-3 BAT arm without adjustment to avoid the introduction of further uncertainty or bias. This may therefore be a conservative approach. The distribution of BAT components (Table 38) was derived based on the distribution of therapies in the BAT arm in the REACH-3 trial adjusted by removing treatments that were not relevant for England according to our advisors.(11) The weights of the removed treatments were then redistributed to ECP, to better reflect the treatment landscape in England. We have tabulated these estimates alongside the REACH-3 study BAT arm below (Table 38).

Table 38. Distribution of BAT components

Treatment	Reported values from the source	Adjusted values for the base case to reflect clinical practice in England*
ECP	34.8%	64.6%
Mycophenolate mofetil	22.2%	22.2%
Imatinib	5.1%	5.1%
Sirolimus	4.4%	4.4%
Rituximab	3.8%	3.8%
Ruxolitinib	0.0%	0.0%
Ibrutinib	17.1%	0.0%
Low-dose methotrexate	6.3%	0.0%
Everolimus	3.2%	0.0%
Infliximab	3.2%	0.0%
Pentostatin	0.0%	0.0%
Others (etanercept, abatacept, hydroxychloroquine, CNIs, IL-2)	0.0%	0.0%

CNI = calcineurin inhibitor; ECP = extracorporeal photopheresis; IL-2 = interleukin-2
*Weighted average of approximately 65% ECP calculated based on feedback from the NICE advisory board, moving the shares of treatments that were not relevant for England to ECP

As described above, the model applies a maximum treatment duration of 5 years for the majority of treatments (except rituximab), based on expert clinician opinion from the advisory board.(11)

B.3.3. Clinical parameters and variables

The key efficacy inputs in the model are FFS, OS, TTR, overall response, DOR and TTD (Table 39). The pooled analysis of the ROCKstar and Phase 2a trials was used as the source of efficacy and safety data for belumosudil (Section B.2.6.3). As described in Section B.3.2.3, our economic analysis had to rely on a naïve direct comparison where the efficacy of belumosudil from the pooled ROCKstar and Phase 2a trials, in patients who received two or more prior lines of therapy (Section B.2.6.3), was directly compared to BAT from the REACH-3 trial.(58)

Table 39. Summary of endpoints across different studies and treatment arms

Outcome	Pooled ROCKstar and Phase 2a – belumosudil(83)	REACH-3 – BAT(58)
OS	✓	✓
FFS	✓	✓
DOR	✓	✓
TTR*	✓	
TTD*	✓	

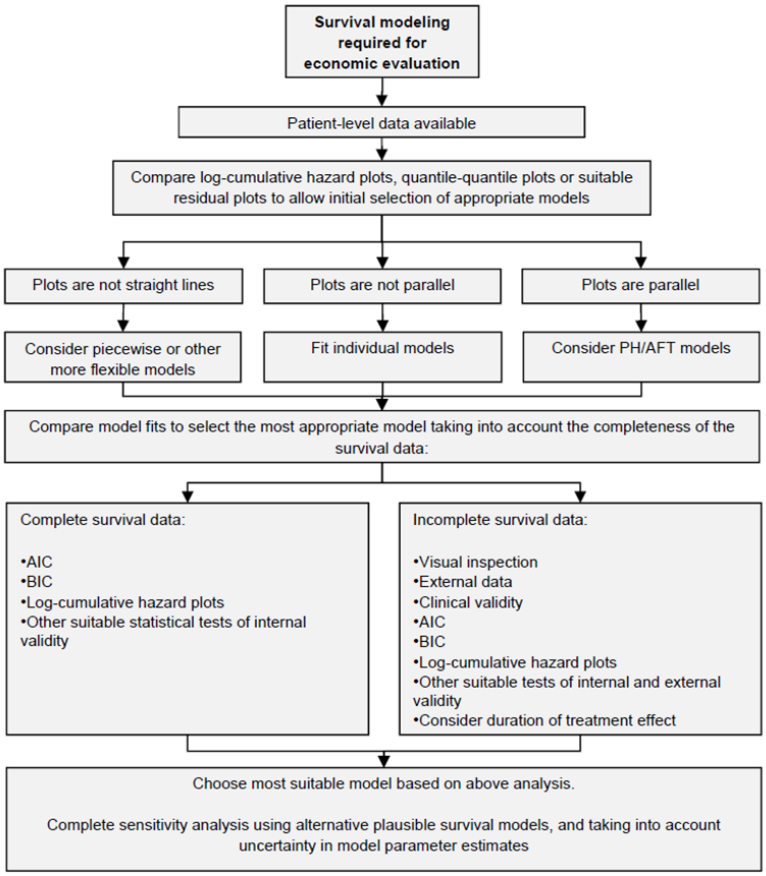
BAT = best available therapy; DOR = duration of response; FFS = failure-free survival; OS = overall survival; TTD = time to treatment discontinuation; TTR = time to response
*While TTR and TTD were not trial endpoints they were derived from the pooled Phase 2 belumosudil studies for the purpose of the economic analysis. For REACH-3, only median statistics for TTR and TTD were available. The methods for deriving TTR and TTD for BAT in the economic analysis are described in Sections B.3.3.4.2 and B.3.3.6.2, respectively.

Parametric survival analyses were conducted by fitting exponential, Weibull, Gompertz, log-logistic, log-normal, generalised gamma and gamma distributions to observed survival data in the pooled analysis of the ROCKstar and Phase 2a trials (for belumosudil) and the REACH-3 trial (for BAT).

While patient-level data were available for the pooled ROCKstar and Phase 2a trials, they were not available for REACH-3.(58) As patient-level data are required in parametric survival analyses, reconstructed individual patient-level data (RIPD) were generated using the Guyot et al.(106) algorithm to obtain data on OS, FFS and DOR for BAT in REACH-3.(58) To create the RIPD, Kaplan-Meier data were digitised to obtain a set of coordinates encoding the observed survival at each time point and supplied to the algorithm together with information (if reported) on the change in the population at risk through time. Note that the Guyot et al.(106) algorithm does not allow to exactly reproduce the patient-level data from the targeted source. Instead, it produces a set of RIPD for a given outcome of interest so that the Kaplan-Meier estimates of the outcome generated using the RIPD set matches the reported Kaplan-Meier estimates for the source.

Properties of these distributions are described in Ishak et al.(107) The parametric survival analysis approach followed the guidelines set out in NICE DSU Technical Support Document 14 (Figure 18).(108)

Figure 18. Selection process algorithm presented by NICE DSU



AFT = accelerated failure time; AIC = Akaike information criteria; BIC = Bayesian information criteria; DSU = Decision Support Unit; NICE = National Institute for Health and Care Excellence; PH = proportional hazard
SOURCE: Latimer et al. 2014(108)

In short:

- A preliminary assessment of fit was made based on diagnostic plots associated with the investigated distributions. A linear pattern observed in these graphs indicates that the distribution may be adequate, and conversely, deviation from linearity indicates poor fit.
- An assessment of the proportional hazard (PH) assumption was done. If the PH assumption holds, joint models with treatment as predictor might be suitable. Otherwise, individual fits should be considered. This step is only used to assess fits from trials that had more than one treatment arms (i.e., belumosudil QD and belumosudil BID arms in the pooled ROCKstar and Phase 2a, and ruxolitinib and BAT arms in REACH-3).
- A plot of ln-negative ln survival vs. ln of time (i.e., cumulative hazard plot) is used to assess if the PH assumption holds. Parallel lines suggest that the PH assumption holds. The PH assumption was also assessed using Schoenfeld residual plots; a straight line (relative to 0) on the plot indicates that the PH assumption holds.
- Joint (i.e., using treatment arm as a predictor) and individual (i.e., each treatment arm is fitted separately) models were fitted to the observed data if appropriate. Depending on the results observed from the PH assumption assessment as described earlier, either joint or stratified models were chosen. Akaike information criteria (AIC) and Bayesian information criteria (BIC) were compared across models, and model(s) achieving the lowest information criteria were chosen (see Appendix N). Interpolation of predicted survival times vs. observed survival times were also produced to assess goodness-of-fit.
- Clinical plausibility of the long-term predictions obtained by the parametric survival models beyond the trial time horizon was assessed by English clinical experts in the advisory board by presenting the extrapolation curves along with predicted mean, median and predicted survival at key landmark points.(11)

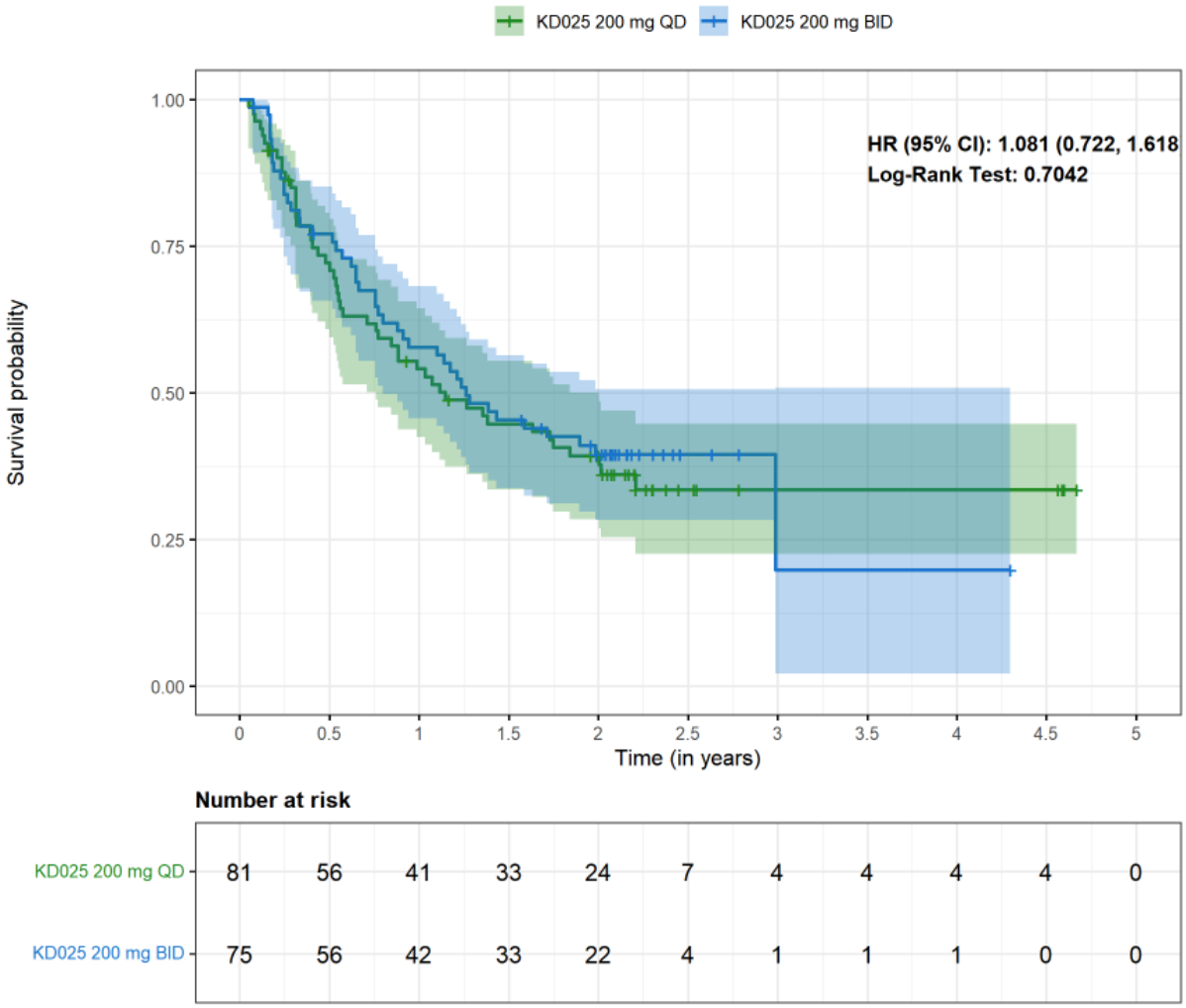
For both the pooled analysis of ROCKstar and the Phase 2a study, and the REACH-3 trial, there were no concerns with respect to PH assumption and thus joint fits of the outcomes were considered. It can be noted that data for FFS and OS were immature for the treatments investigated in the ROCKstar, Phase 2a and REACH-3 trials and thus joint fits provide more reliable estimates of survival.

B.3.3.1. Failure-free survival

B.3.3.1.1. Belumosudil: Pooled ROCKstar and Phase 2a trials

The data included 156 observations from the belumosudil 200 mg once daily and belumosudil 200 mg twice daily arms of the pooled trials and a total of 95 failure events (50 in the once daily arm and 45 in the twice daily arm) over a maximum follow-up duration of 4.7 years.(83, 109) Median FFS was 13.7 months in the once daily arm and 15.1 months in the twice daily arm.(83) The Kaplan-Meier curves for FFS in the individual arms of the pooled trials are displayed in Figure 19. Please refer to Section B.2.6.3 for detailed results of the pooled analysis.

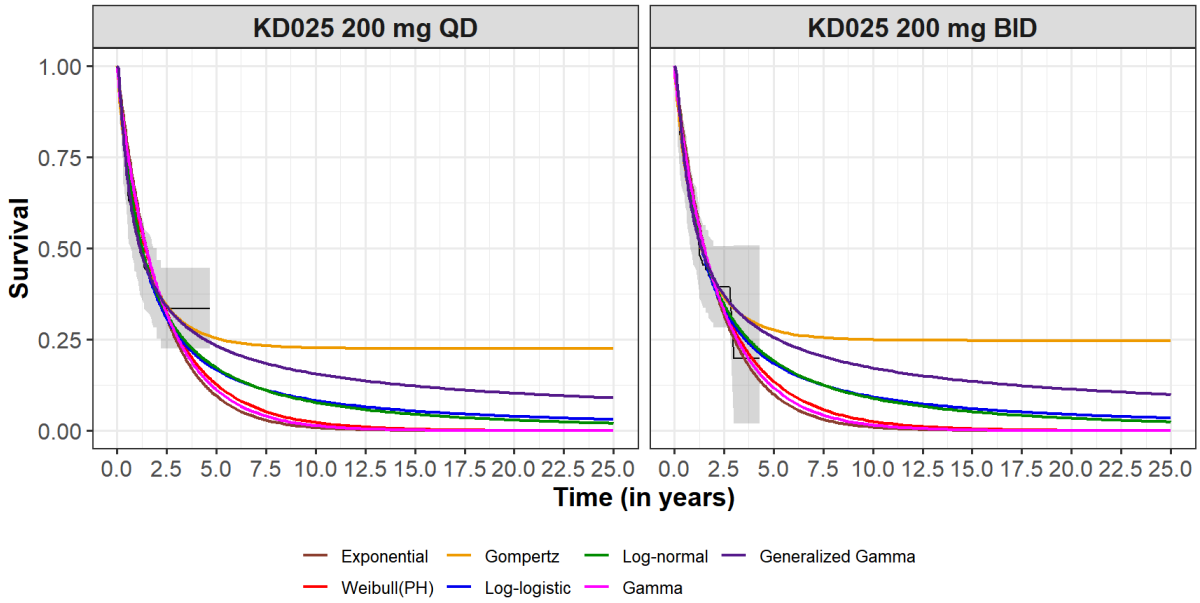
Figure 19. Kaplan-Meier Curves for FFS in belumosudil once daily and twice daily arms of the pooled ROCKstar and Phase 2a trials



CI = confidence interval; HR = hazard ratio; KD025 200 mg BID = belumosudil 200 mg twice daily; KD025 200 mg QD= belumosudil 200 mg once daily

Figure 20 presents the long-term extrapolations of the parametric survival models fitted on FFS for belumosudil once daily and twice daily. The generalised gamma model was selected to estimate FFS for belumosudil based on AIC and BIC fit statistics (Appendix N) and clinical plausibility assessed in the advisory board by English clinical experts.(11) The distribution of failure events by cause for belumosudil QD and BID in the model based on pooled ROCKstar and KD025-208 trials is provided in Appendix N, Table 5, and Figures 9 to 11.

Figure 20. Predicted parametric FFS models for the pooled ROCKstar and Phase 2a trials (belumosudil once daily and twice daily; joint fit)

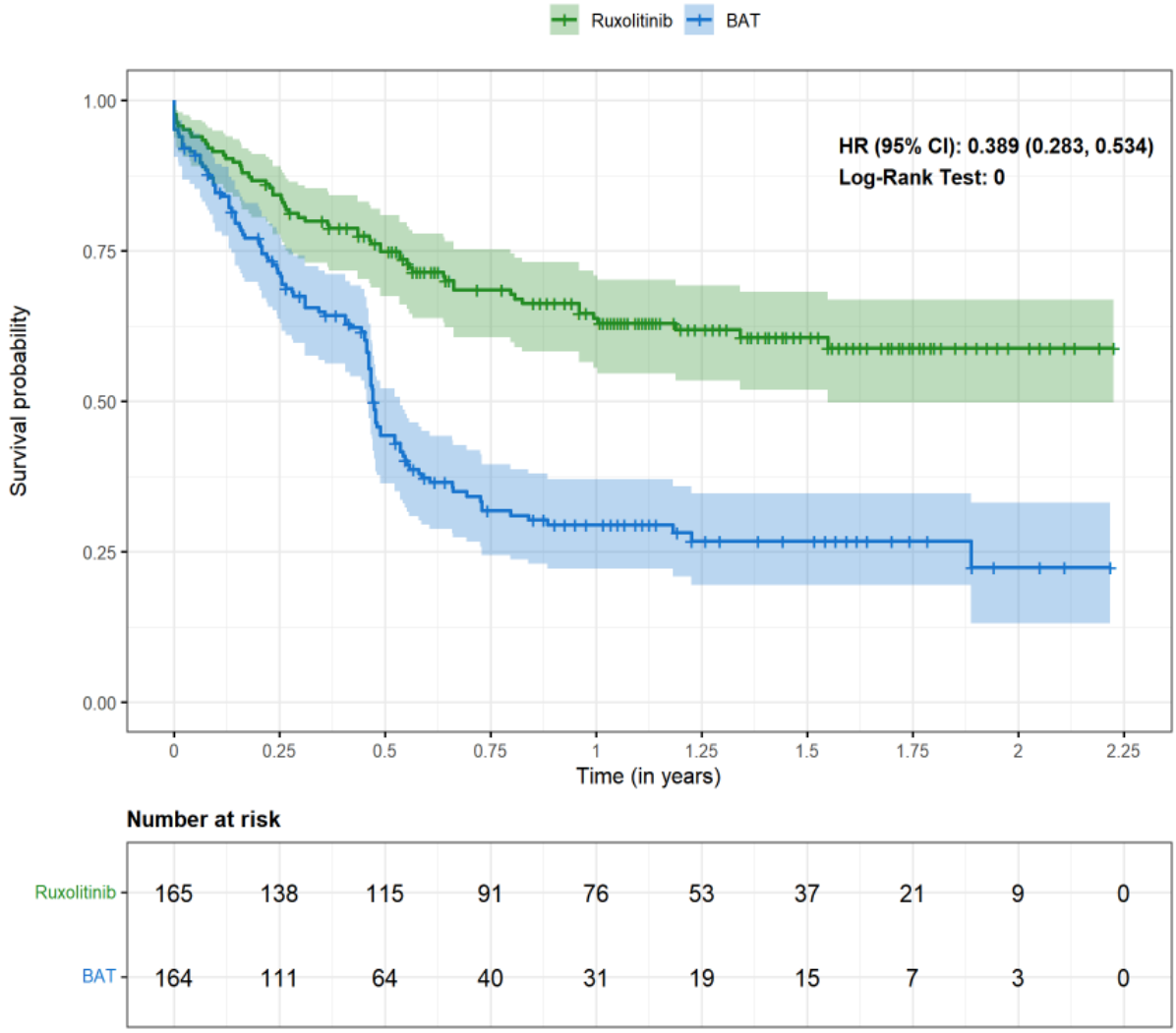


BID = twice daily; FFS = failure-free survival; PH = proportional hazard; QD = once daily

B.3.3.1.2. BAT

The data included 164 observations from the BAT arm of the REACH-3 study, with a total of 109 failure events over a maximum follow-up duration of 2.2 years.(58) Median FFS was 5.7 months in the BAT arm.(58) The Kaplan-Meier curve for FFS in the BAT arm of REACH-3 trial is displayed in Figure 21. Distribution of failure events by cause for BAT used in the model is provided in Appendix N, Table 6.

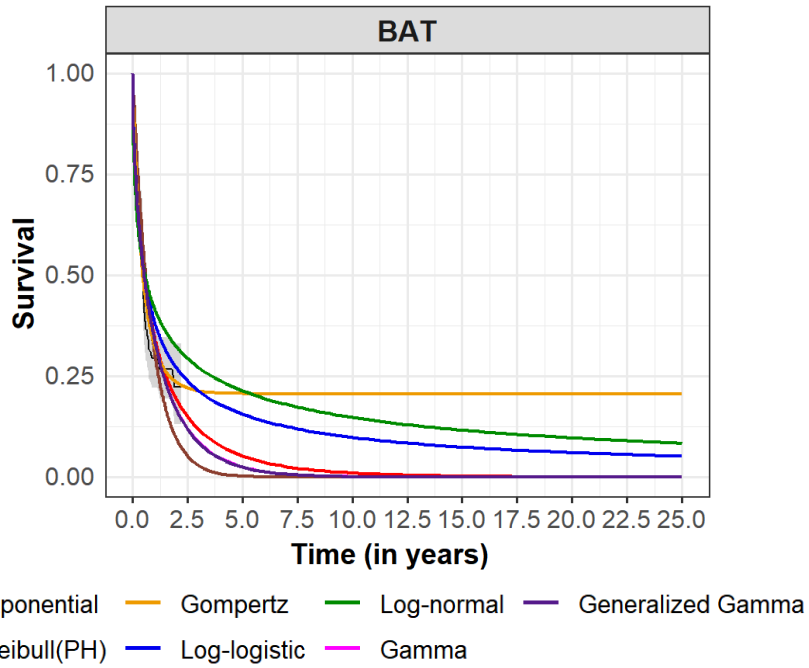
Figure 21. Kaplan-Meier curve for FFS in the BAT arm in REACH-3*



BAT = best available therapy; CI = confidence interval; HR = hazard ratio; KM = Kaplan-Meier
*Kaplan-Meier curve was created based on reconstructed individual patient-level data (RIPD) using the Guyot et al.(106) algorithm. It is important to note that this algorithm does not allow to reproduce exactly the patient-level data from the targeted source, but instead produces a set of RIPD for a given outcome of interest so that the KM estimates of the outcome generated using the RIPD set matches the reported KM estimates for the source. Ruxolitinib data are presented as joint fits were used in order to increase the power with which the ancillary parameters of the distribution are estimated.

Figure 22 presents the long-term extrapolations of the parametric survival models fitted on FFS for BAT. The generalised gamma model was selected to estimate FFS based on AIC and BIC fit statistics (Appendix N) and clinical plausibility assessed by English clinical experts in the advisory board.(11) The same type of parametric model was applied as for belumosudil based on guidance in NICE DSU Technical Support Document 14.(108)

Figure 22. Predicted parametric FFS models for the BAT arm in REACH-3 (joint fit)



BAT = best available therapy; FFS = failure-free survival; PH = proportional hazard

B.3.3.1.3. Distribution of failure events by cause

FFS was defined as the time from the first dose of belumosudil to the time of initiation of new systemic chronic GVHD therapy, non-relapse mortality or recurrent malignancy (whichever occurred first). As described in Section B.3.2.2, patients in the Failure state are distributed into one of the two reasons for non-fatal failure (initiation of new systemic chronic GVHD therapy or recurrent malignancy) and accrue economic consequences (i.e., costs, utilities) according to the reason for failure.

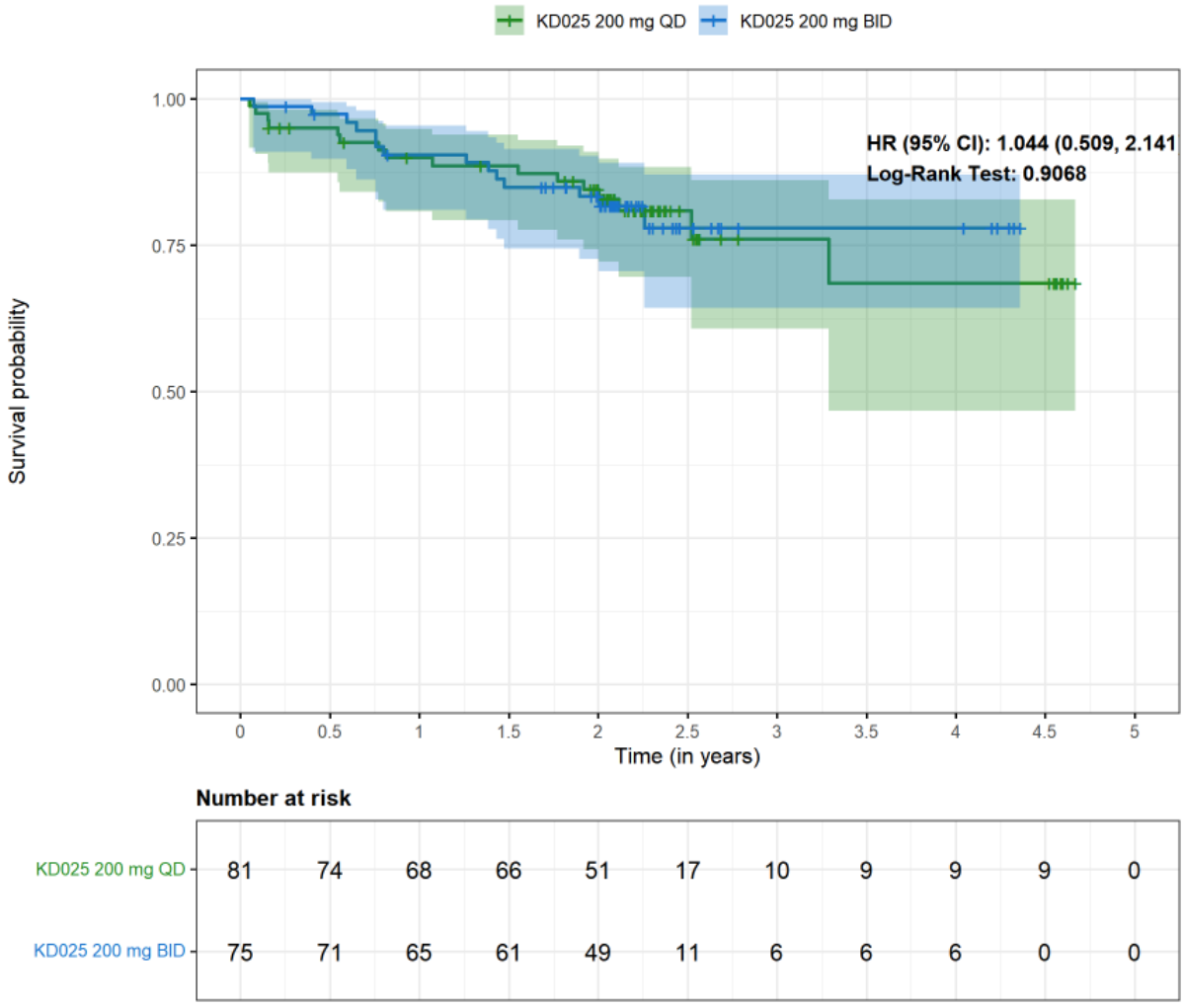
Distributions of failure events by cause were derived from KM plots and data on patients at risk in different time periods. Model inputs for these distributions of failure events and detailed description of the assumptions considered to derive those are provided in Appendix N.

B.3.3.2. Overall survival

B.3.3.2.1. Belumosudil: Pooled ROCKstar and Phase 2a trials

The data included a total of 30 deaths (16 deaths in the belumosudil once daily arm and 14 deaths in the belumosudil twice daily arm) over a maximum follow-up duration of 4.7 years.(83, 109) Median OS was not reached in either arm.(83) The Kaplan-Meier curves for OS in the pooled ROCKstar and Phase 2a trials are displayed in Figure 23. Please refer to Section B.2.6.3 for detailed results of the pooled analysis.

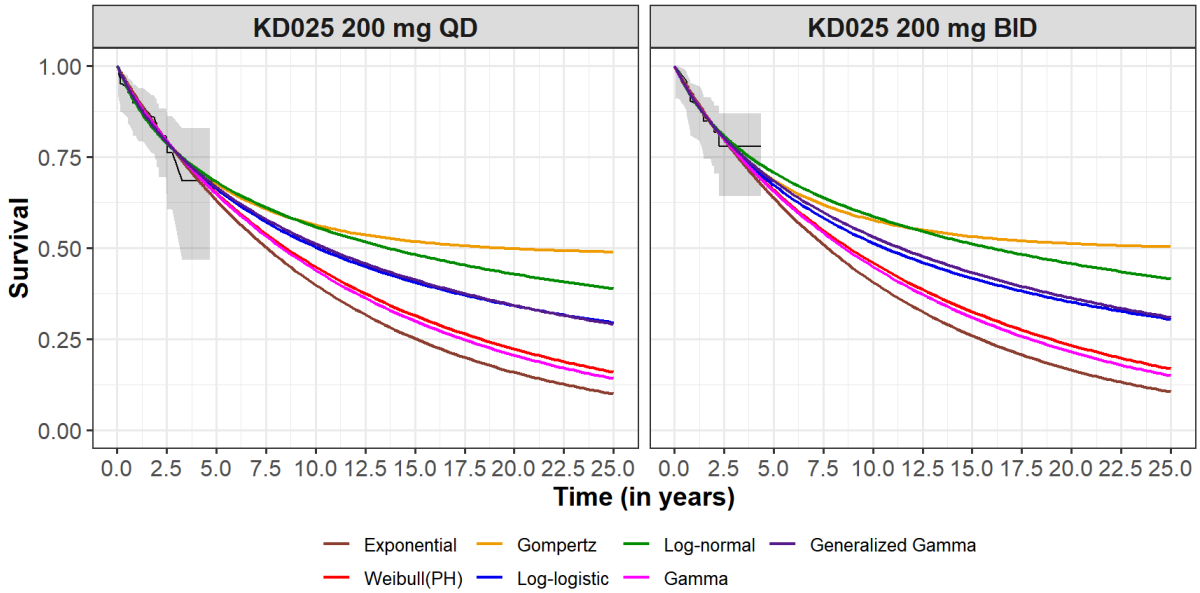
Figure 23. Kaplan-Meier curves for OS in belumosudil once daily and twice daily arms of the pooled ROCKstar and Phase 2a trials



CI = confidence interval; HR = hazard ratio; KD025 200 mg BID = belumosudil 200 mg twice daily; KD025 200 mg QD= belumosudil 200 mg once daily

Figure 24 presents long-term extrapolations of the parametric survival models fitted on OS for belumosudil once daily and twice daily. The exponential model was selected to estimate OS for belumosudil based on AIC and BIC fit statistics (Appendix N) and clinical plausibility assessed by English clinical experts in the advisory board. (11) Please refer to Section B.3.3.2.3 for further details on adjustment of the OS curves.

Figure 24. Predicted parametric OS models for the pooled ROCKstar and Phase 2a trials (belumosudil once daily and twice daily; joint fit)

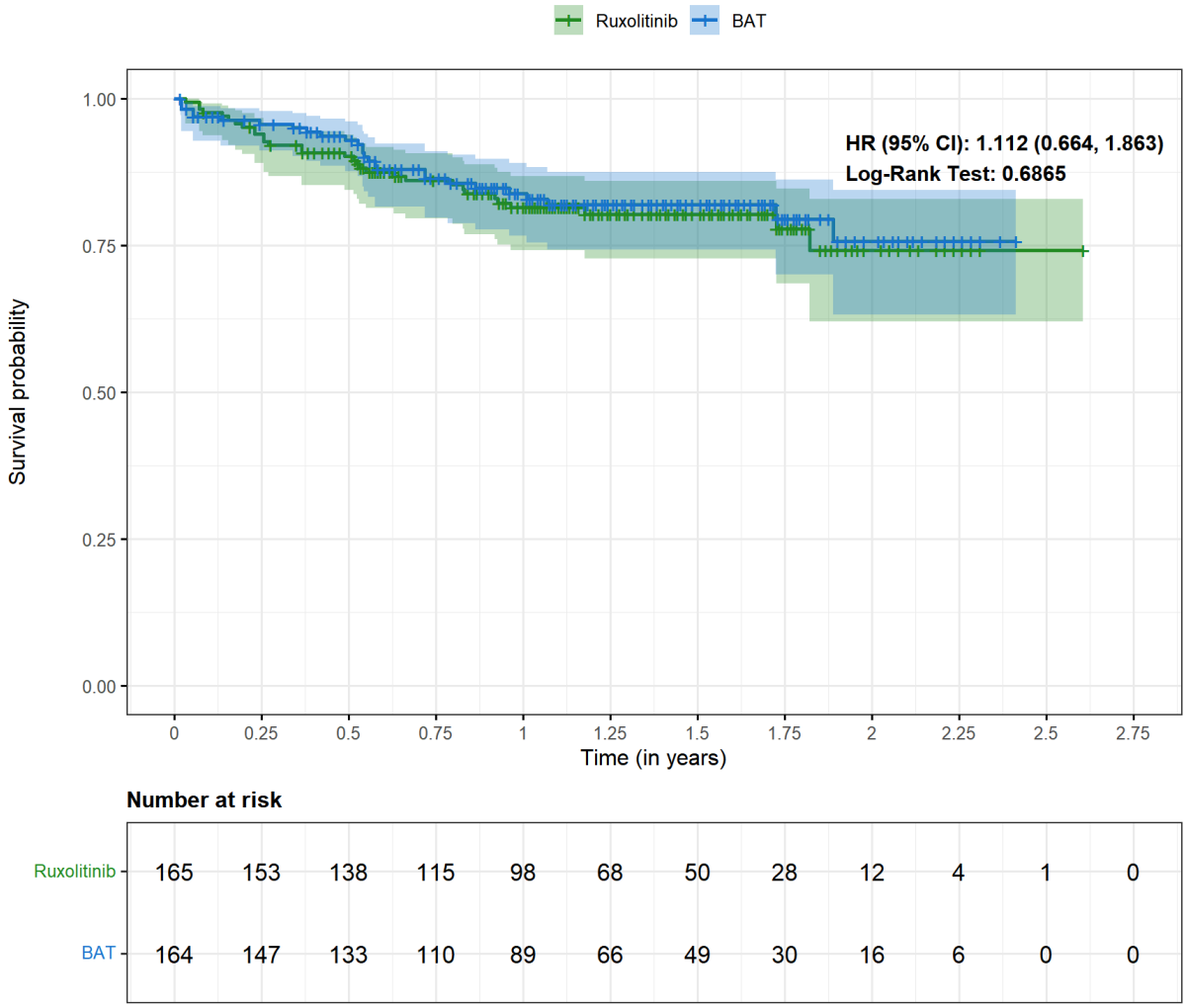


BID = twice daily; KM = Kaplan-Meier; OS = overall survival; PH = proportional hazard; QD = once daily

B.3.3.2.2. BAT

The data used included 164 observations from the BAT arm of the REACH-3 study, with a total of 27 deaths in the BAT arm over a maximum follow-up duration of approximately 2.4 years.(58) Median OS was not reached in the BAT arm.(58) The Kaplan-Meier curve for OS in the BAT arm of REACH-3 trial is displayed in Figure 25.

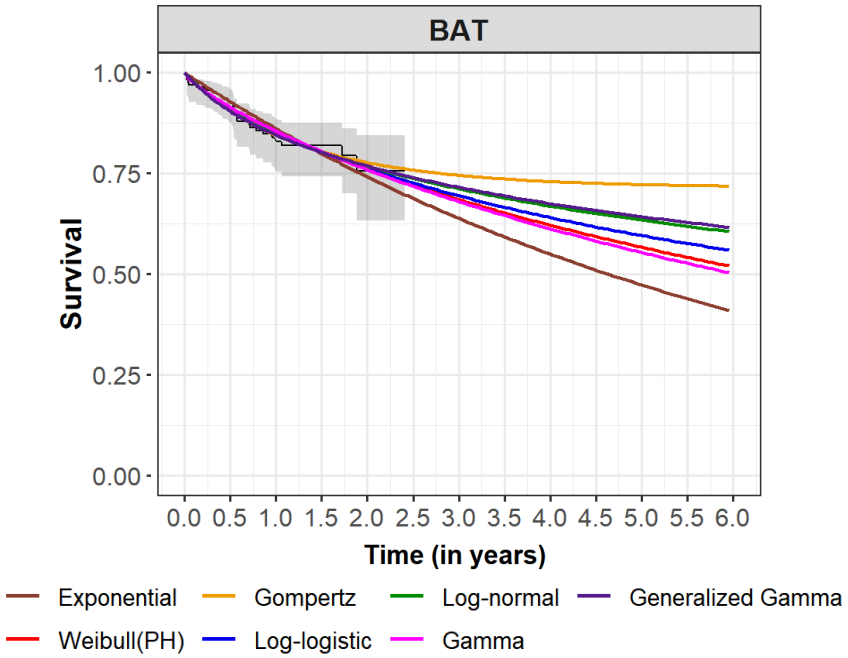
Figure 25. Kaplan-Meier curves for OS in the BAT arm in REACH-3*



BAT = best available therapy; CI = confidence interval; HR = hazard ratio; KM = Kaplan-Meier
*Kaplan-Meier curve was created based on reconstructed individual patient-level data (RIPD) using the Guyot et al.(106) algorithm. It is important to note that this algorithm does not allow to reproduce exactly the patient-level data from the targeted source, but instead produces a set of RIPD for a given outcome of interest so that the KM estimates of the outcome generated using the RIPD set matches the reported KM estimates for the source. Ruxolitinib data are presented as joint fits were used in order to increase the power with which the ancillary parameters of the distribution are estimated.

Figure 26 presents the long-term extrapolations of the parametric survival models fitted on OS for BAT. The exponential model was selected to estimate OS based on AIC and BIC fit statistics (Appendix N) and clinical plausibility assessed by English clinical experts in the advisory board.(11) The same type of parametric model was applied as for belumosudil based on guidance in NICE DSU Technical Support Document 14.(108)

Figure 26. Predicted parametric OS models for the BAT arm in REACH-3 (joint fit)

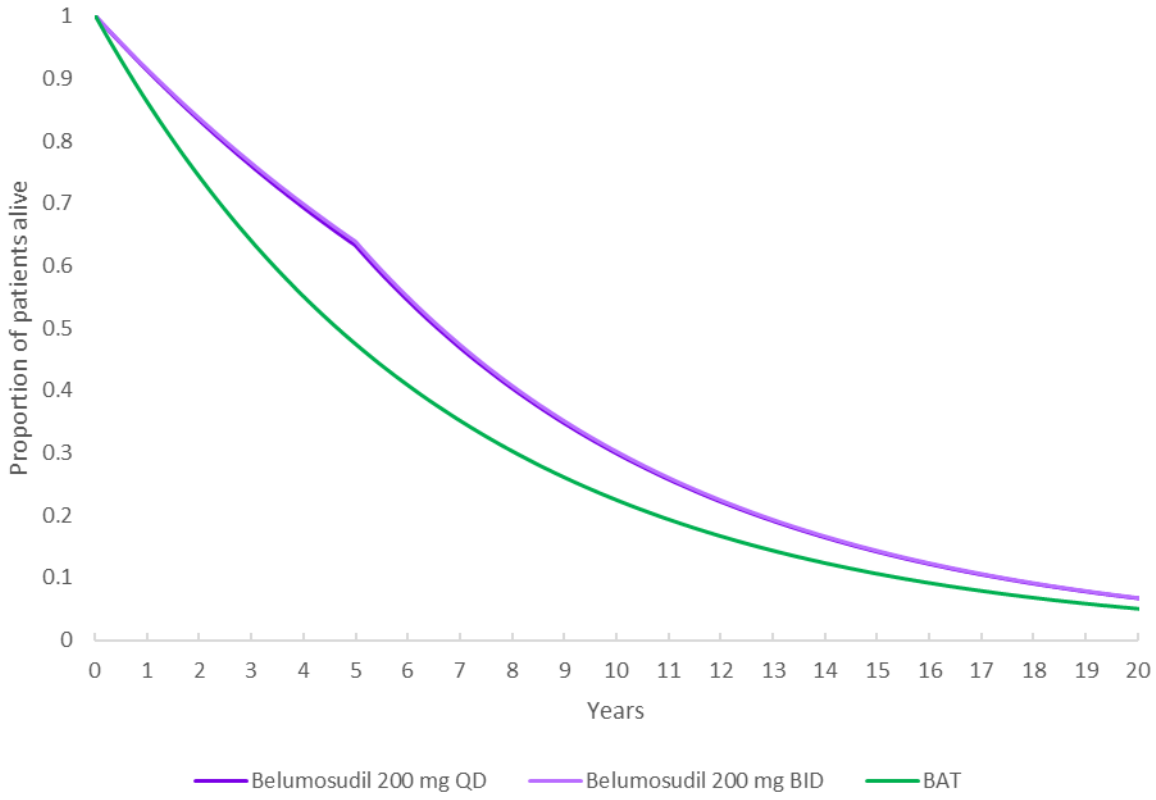


BAT = best available therapy; OS = overall survival; PH = proportional hazard

B.3.3.2.3. Overall survival after 5 years

The model assumes that after 5 years all patients (treated with belumosudil or BAT) have the same probability of death to account for uncertainty in the long-term treatment benefit of belumosudil. This assumption was validated with clinicians at the advisory board.(11) Some experts commented that the exponential fit, which forecasts the shortest OS in the long run, may be optimistic at this stage in the disease pathway. Recognising the uncertainty in this long-term extrapolation based on fit statistics to a relatively small number of observations (only ~25% of patients had died at follow-up), we have taken a parsimonious approach through this adjustment beyond 5 years. This corresponds to the observed data which has a maximum follow-up duration of 4.7 years (Section B.3.3.1.1). Estimated OS curves based on the selected fits (i.e., joint fit with exponential model for both belumosudil and BAT) and assuming the same probability of death as patients initially treated with BAT post-5 years are shown in Figure 27.

Figure 27. Extrapolated OS curves assuming same probability of death after 5 years*



BAT = best available therapy; BID = twice daily; OS = overall survival; QD = once daily
*Based on exponential joint fits for both the belumosudil arms and BAT

B.3.3.3. Response

Overall response, as defined by the 2014 NIH Consensus Criteria(81), was the primary endpoint in both ROCKstar and REACH-3. However, the primary endpoint of ROCKstar was best response at any post-baseline assessment, while response in REACH-3 was assessed at week 24. Thus it is important to note that there is uncertainty regarding the comparability of response outcomes across the trials.

The number and distribution of patients with either CR or PR is included in Table 40 for belumosudil and BAT. The distributions of CR and PR were renormalised for patients with a response (i.e., removing those with no response) and used to estimate the proportion of patients who achieve CR versus PR among those who are in the 'in-response' state in each cycle (Table 41). The model assumes the proportions of CR versus PR among the responders are constant over time.

Table 40. Response data for each treatment comparator

	Pooled ROCKstar and Phase 2a		REACH-3
	Belumosudil 200 mg once daily (n=81)	**Belumosudil 200 mg twice daily (n=75)**	**BAT (n=164)**
Number of patients with response (%)	59 (72.8%)	55 (73.3%)	99 (60.4%)
CR (%)	4 (4.9%)	2 (2.7%)	11 (6.7%)
PR (%)	55 (67.9%)	53 (70.7%)	88 (53.7%)
Number of patients with no response (%)	22 (27.2%)	20 (24.2%)	65 (39.6%)

BAT = best available therapy; CR = complete response; PR = partial response
SOURCE: Kadmon Pharmaceuticals 2022(83); Zeiser et al. 2021(58)

Table 41. Distribution of response level among responders

	Pooled ROCKstar and Phase 2a		REACH-3
	Belumosudil 200 mg once daily	**Belumosudil 200 mg twice daily**	**BAT**
CR (%)	6.78%	3.64%	11.11%
PR (%)	93.22%	96.36%	88.89%

BAT = best available therapy; CR = complete response; PR = partial response

B.3.3.4. Time to response

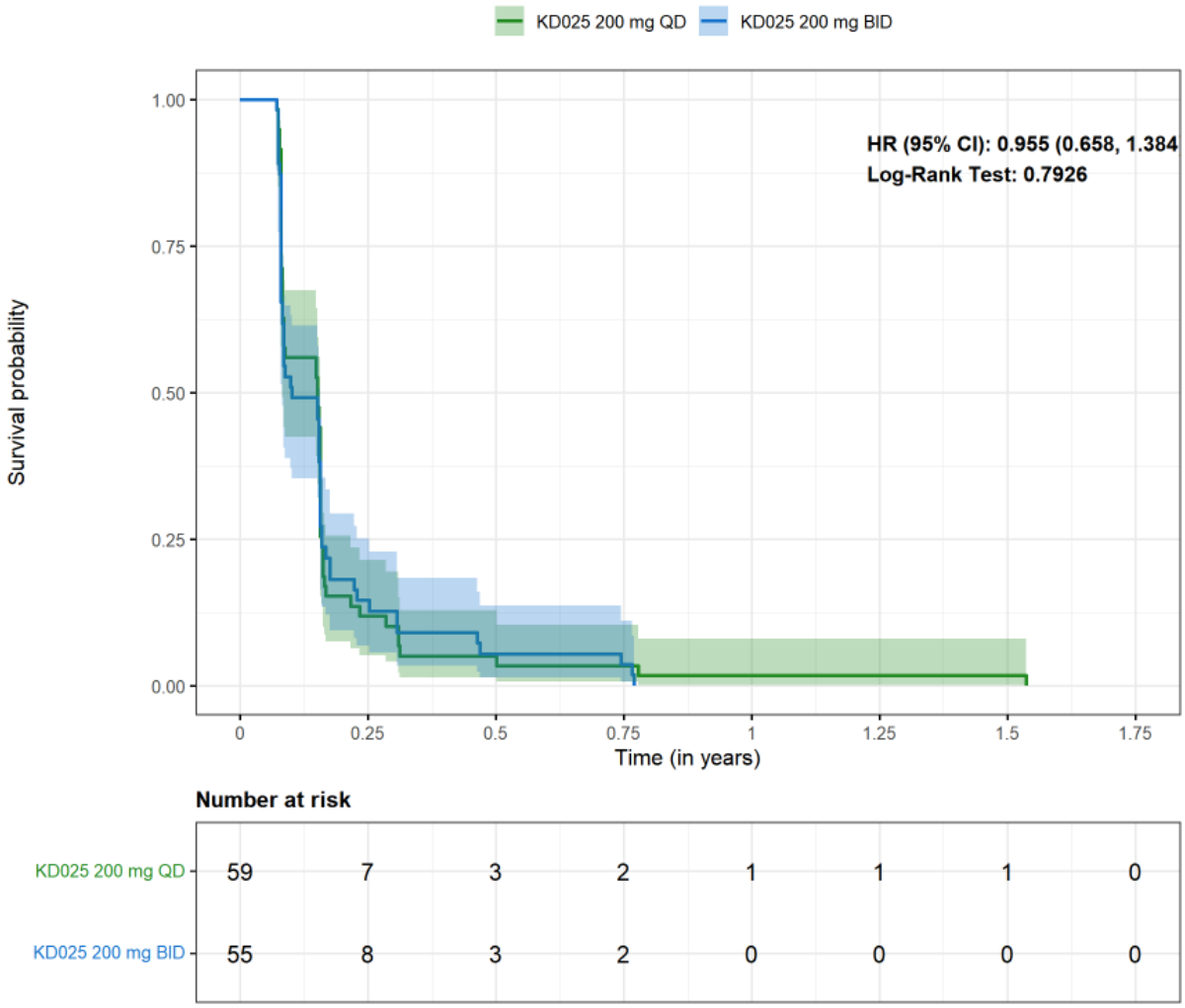
The TTR curve is used in combination with the DOR curve to estimate the 'in-response' curve for each comparator in the model. As shown in Table 39, data to allow estimation of TTR are available for belumosudil. For BAT, only median TTR data are available in published sources, therefore the model uses survival parameter estimates derived from the reported medians assuming an exponential distribution (Section B.3.3.4.2).

B.3.3.4.1. Belumosudil: Pooled ROCKstar and Phase 2a trials

The data included a total of 114 response events (59 from the once daily arm and 55 from the twice daily arm) over a maximum follow-up duration of 1.5 years.(83, 109) Median TTR was 7.86 weeks in the once daily arm and around 5.29 weeks in the twice daily arm.(83) The Kaplan-Meier curves for TTR in the individual arms of the pooled trials are displayed in Figure 28.

By the end of the follow-up period, more than 98% of the patients in both the belumosudil once daily and twice daily arms had an event, thus TTR curves from the pooled ROCKstar and Phase 2a trials data are used directly in the model, without any extrapolation. Please refer to Section B.2.6.3 for detailed results of the pooled analysis.

Figure 28. Kaplan-Meier curves for TTR in belumosudil once daily and twice daily arms of the pooled ROCKstar and Phase 2a trials



CI = confidence interval; HR = hazard ratio; KD025 200 mg BID = belumosudil 200 mg twice daily; KD025 200 mg QD= belumosudil 200 mg once daily; TTR = time to response

B.3.3.4.2. BAT

For BAT, only median TTR statistics were available.(110) Therefore, the model uses survival parameter estimates derived from the reported median TTR assuming an exponential distribution. Median TTR was 4.00 months for BAT.

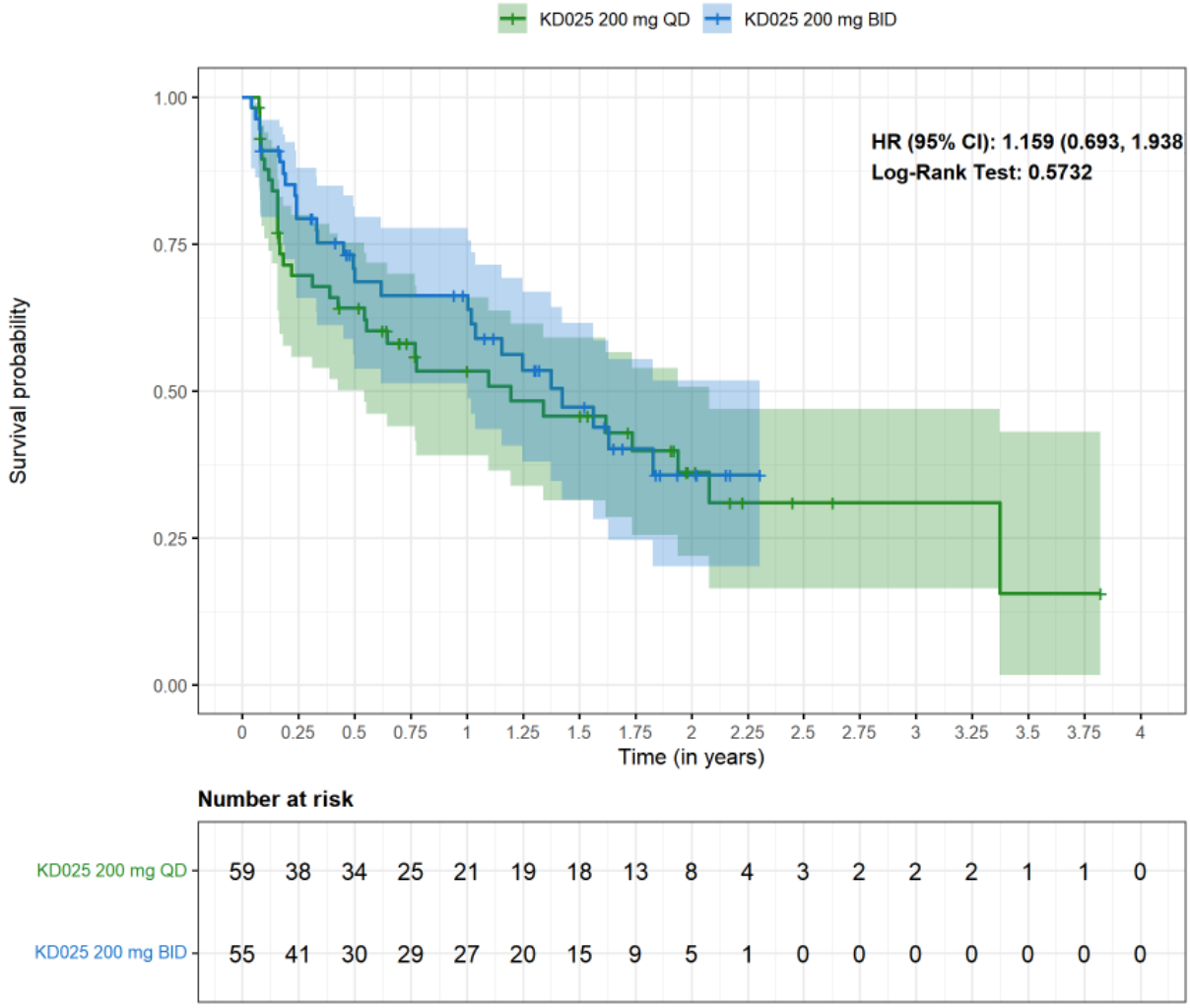
B.3.3.5. Duration of response

DOR for belumosudil was informed through individual fitting of quaternary duration of response curves, defined in ROCKstar as the time from first documentation of response to the time of first documentation of lack of response but with durations summed for multiple responses/LR episodes. Quaternary duration of response was selected for comparability reasons, as it was the closest to the definition of DOR in the REACH-3 trial.

B.3.3.5.1. Belumosudil: Pooled ROCKstar and Phase 2a trials

The data included 114 observations from the belumosudil 200 mg once daily and belumosudil 200 mg twice daily arms of the pooled trials. The data included a total of 60 end of response events (33 from the once daily arm and 27 from the twice daily arm) over a maximum follow-up duration of 3.8 years.(83, 109) Median quaternary DOR was 62.3 weeks (14.3 months) in the once daily arm and 74.3 weeks (17.1 months) in the twice daily arm.(83) The Kaplan-Meier curves for quaternary DOR in the individual arms of the pooled trials are displayed in Figure 29. It is important to note that data for quaternary DOR are among responders and not the overall population in each treatment arm. Please refer to Section B.2.6.3 for detailed results of the pooled analysis.

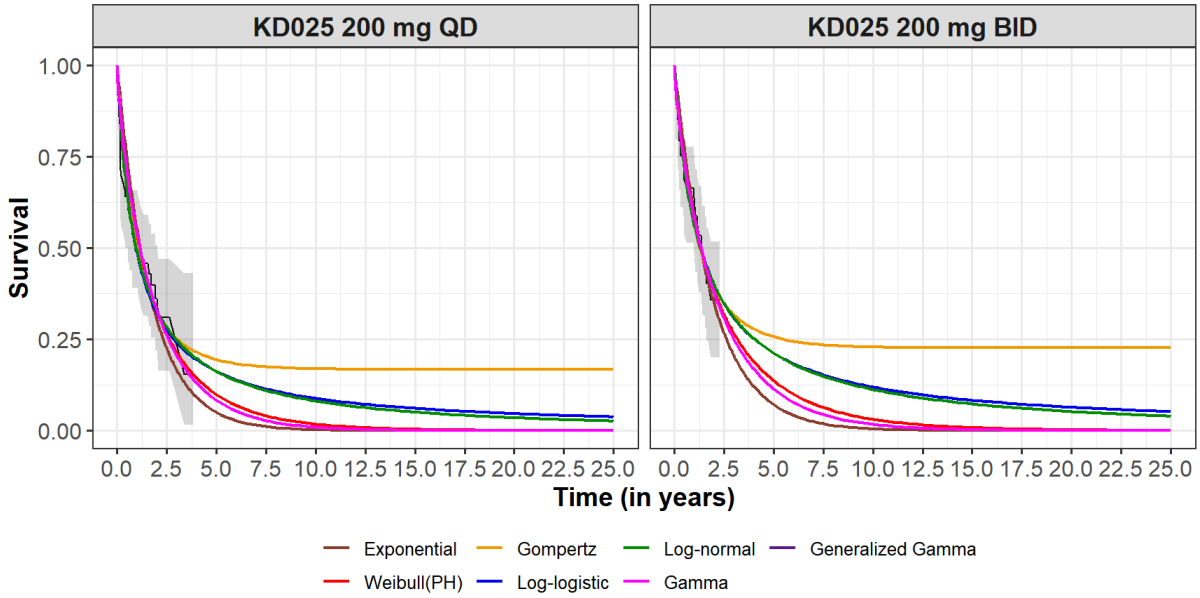
Figure 29. Kaplan-Meier curves for quaternary DOR in belumosudil once daily and twice daily arms of the pooled ROCKstar and Phase 2a trials



CI = confidence interval; HR = hazard ratio; KD025 200 mg BID = belumosudil 200 mg twice daily; KD025 200 mg QD = belumosudil 200 mg once daily

Figure 30 presents the long-term extrapolations of the parametric survival models fitted on DOR for belumosudil once daily and twice daily. The log-normal distribution was selected to estimate DOR based on AIC and BIC fit statistics for belumosudil (Appendix N) and clinical plausibility assessed by English clinical experts in the advisory board.(11)

Figure 30. Predicted parametric DOR models for the pooled ROCKstar and Phase 2a trials (belumosudil once daily and twice daily; joint fit)

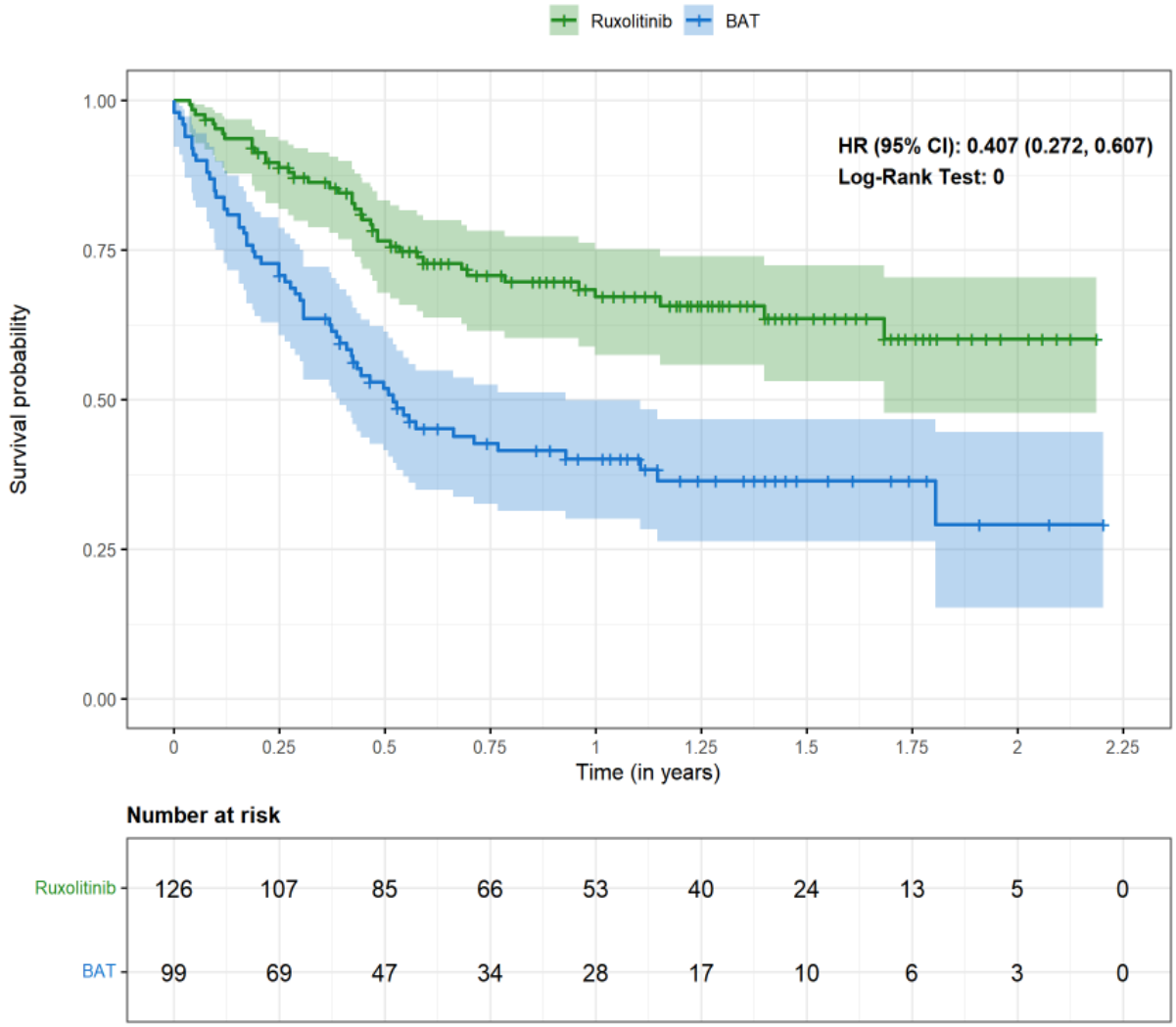


BID = twice daily; DOR = duration of response; PH = proportional hazard; QD = once daily
Note: the generalised gamma model could not be fitted for the 200 mg once daily arm of the trial.

B.3.3.5.2. BAT

The data used included 99 observations from the BAT arm of the REACH-3 study, with a total of 60 response events in the BAT arm over a maximum follow-up duration of approximately 2.2 years.(58) Median DOR was 6.24 months in the BAT arm.(58) The Kaplan-Meier curve for DOR in the BAT arm of the REACH-3 trial is displayed in Figure 31.

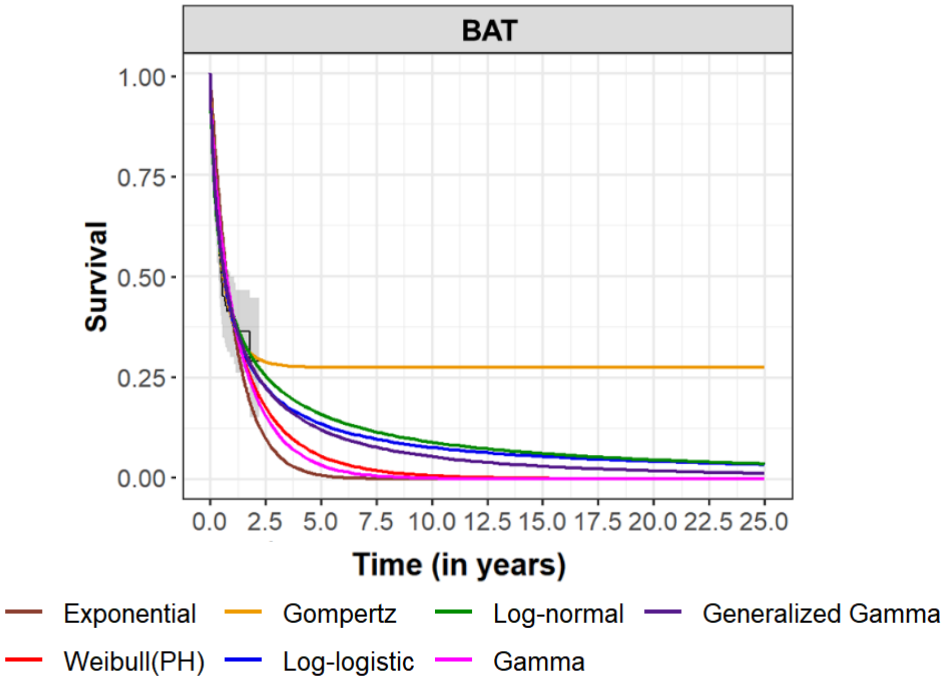
Figure 31. Kaplan-Meier curves for DOR in the BAT arm in REACH-3*



BAT = best available therapy; CI = confidence interval; HR = hazard ratio; KM = Kaplan-Meier
*Kaplan-Meier curve was created based on reconstructed individual patient-level data (RIPD) using the Guyot et al.(106) algorithm. It is important to note that this algorithm does not allow to reproduce exactly the patient-level data from the targeted source, but instead produces a set of RIPD for a given outcome of interest so that the KM estimates of the outcome generated using the RIPD set matches the reported KM estimates for the source. Ruxolitinib data are presented as joint fits were used for parametric fitting.

Figure 32 presents the long-term extrapolations of the parametric survival models fitted on DOR for BAT. The same type of parametric model was applied as for belumosudil based on guidance in NICE DSU Technical Support Document 14.(108) Hence the log-normal distribution was selected to estimate DOR, with clinical plausibility assessed by English clinical experts in the advisory board.(11)

Figure 32. Predicted parametric DOR models for the BAT arm in REACH-3 (joint fit)



BAT = best available therapy; DOR = duration of response; PH = proportional hazard

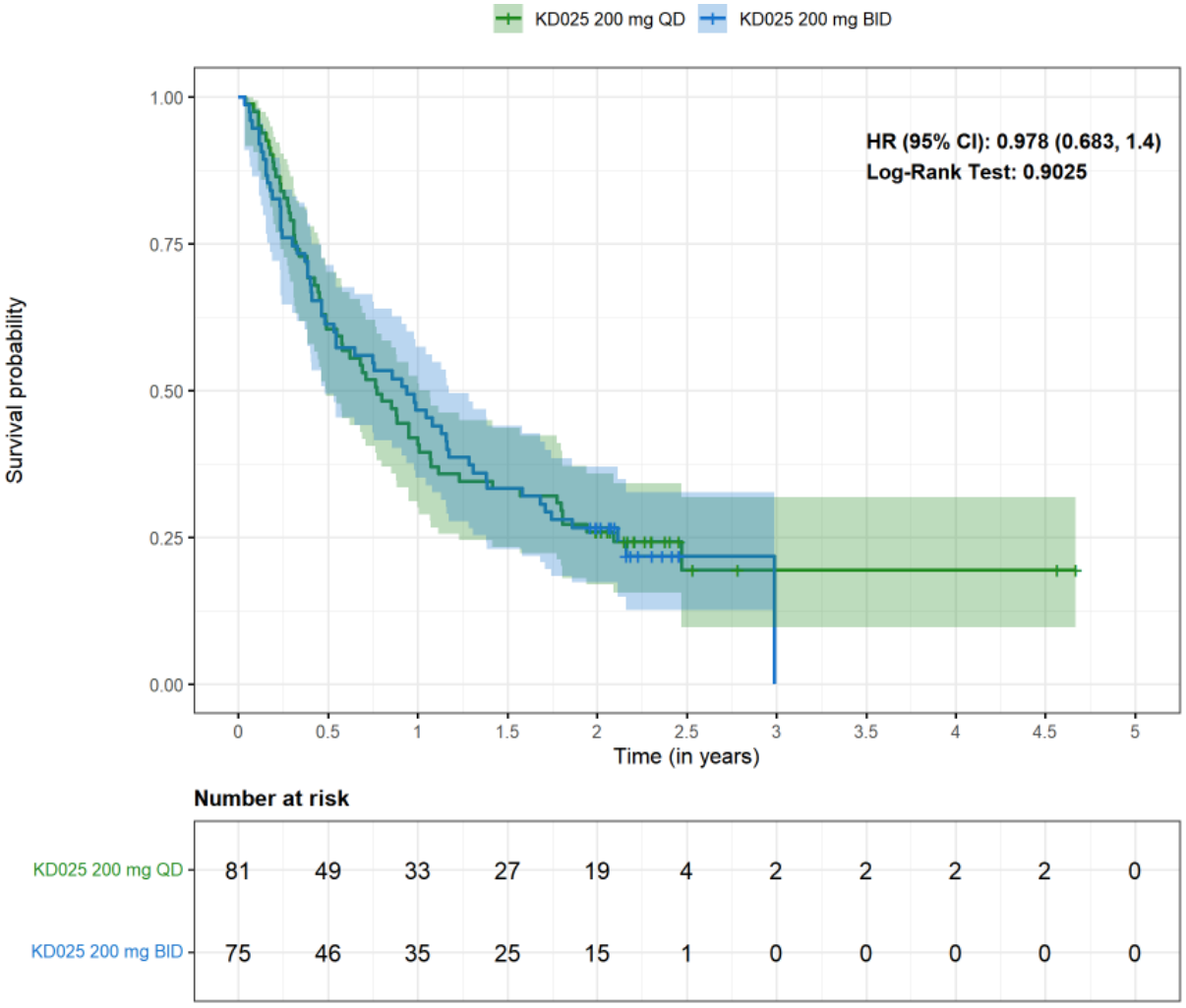
B.3.3.6. Time to treatment discontinuation

Time on treatment in the model was estimated based on TTD. As shown in Table 39, data to allow estimation of the TTD curve were only available for belumosudil. For BAT, only median treatment duration data are available in published sources, thus TTD was estimated by applying a hazard ratio (HR) to the TTD curve of belumosudil QD (details in Section B.3.3.6.2).

B.3.3.6.1. Belumosudil: Pooled ROCKstar and Phase 2a trials

The data included 156 observations from the belumosudil 200 mg once daily and belumosudil 200 mg twice daily arms of the pooled trials.(83) The data included a total of 121 events (63 in the once daily arm and 58 in the twice daily arm) over a maximum follow-up duration of 4.7 years.(83, 109) Median TTD was 9.2 months in the once daily arm and 11.2 months in the twice daily arm.(83) The Kaplan-Meier curves for TTD in the individual arms of the pooled trials are displayed in Figure 33. Please refer to Section B.2.6.3 for detailed results of the pooled analysis.

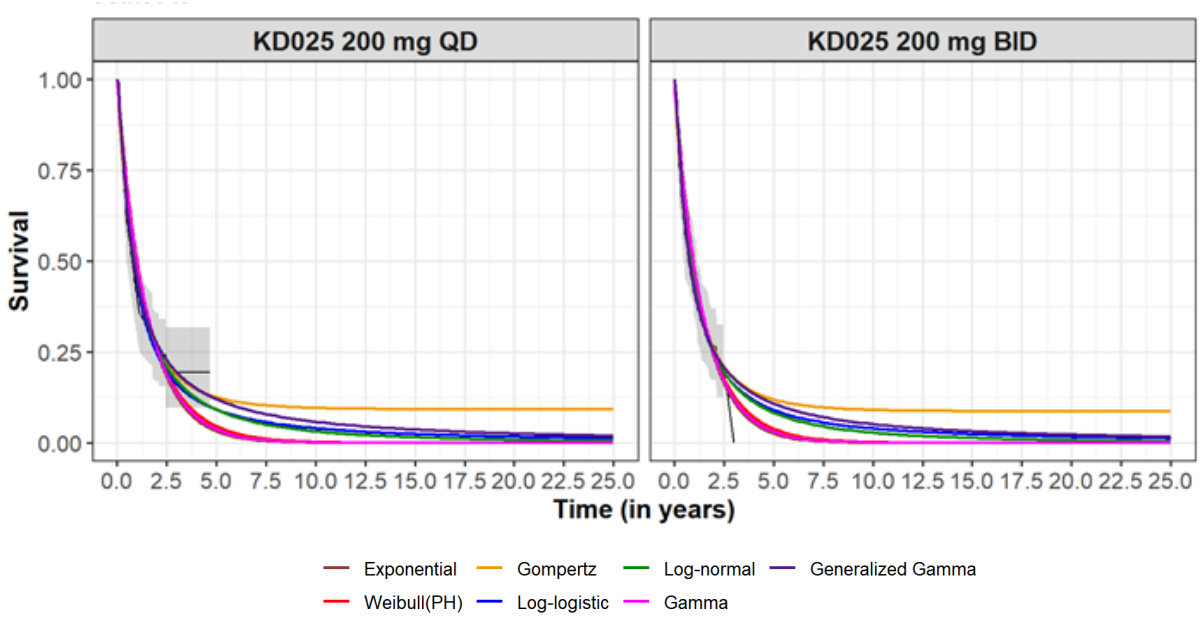
Figure 33. Kaplan-Meier curves for TTD in belumosudil once daily and twice daily arms of the pooled ROCKstar and Phase 2a trials



CI = confidence interval; HR = hazard ratio; KD025 200 mg BID = belumosudil 200 mg twice daily; KD025 200 mg QD= belumosudil 200 mg once daily; TTD = time to treatment discontinuation

Figure 34 presents the long-term extrapolations of the parametric survival models fitted on TTD for belumosudil once daily and twice daily. The log-normal distribution was selected to estimate TTD based on AIC and BIC fit statistics for belumosudil (Appendix N) and clinical plausibility assessed by English clinical experts in the advisory board.(11)

Figure 34. Predicted parametric TTD models for the pooled ROCKstar and Phase 2a trials (belumosudil once daily and twice daily; joint fit)



BID = twice daily; PH = proportional hazard; QD = once daily; TTD = time to treatment discontinuation

B.3.3.6.2. BAT

For BAT, only median treatment duration statistics were available.(58) The TTD curves were therefore estimated by applying an HR to the TTD curve of belumosudil QD from the pooled ROCKstar and Phase 2a trials. The HR was obtained by calibrating the HR versus the Log-normal fit of the belumosudil QD curve, to match the reported median treatment duration of the comparator. The estimated HR for BAT is presented in Table 42. Standard errors are assumed to be 20% of the mean.

Table 42. HR of TTD for BAT versus the Log-normal fit of the belumosudil QD curve

Treatment	Reported median treatment duration (months)	Hazard ratio compared to belumosudil 200 mg QD from the pooled ROCKstar and Phase 2a trials*		
		Mean	Lower bound	Upper bound
BAT	5.54	1.80	1.44	2.16

BAT = best available therapy; HR = hazard ratio; QD = once daily; TTD = time to treatment discontinuation
*HRs were generated based on Log-normal model of TTD for belumosudil QD

B.3.3.6.3. Maximum treatment duration

Based on feedback from the advisory board that patients with chronic GVHD who are stable and responding to treatment are unlikely to be on treatment beyond 5 years(11), the base case assumes duration of treatment is capped at 5 years. This is applicable for all treatments except rituximab, where patients are treated for up to 4 weeks.(111) A maximum treatment duration of 3 years was explored in scenario analysis. We believe this is an important scenario to consider. Experts consulted at the English advisory board in January 2023 suggested that patients at this stage of their journey often do not stay on treatments for very long periods of time.(11) They felt that 3 to 5 years was the maximum duration that any therapy would reasonably be sustained for.(11) Where failure events

occur patients will come off treatment anyway and for those patients doing well on therapy, it is likely they would discontinue at some point either through personal choice or clinician advice.(11)

B.3.3.7. Safety inputs

The model includes Grade ≥3 AEs occurring in at least 5% of patients in either of the treatment arms of the pooled ROCKstar and Phase 2a trials or REACH-3 trial. This is a commonly accepted approach as Grade ≥3 AEs reflect events that are likely to require substantial healthcare resource use and have a significant impact on QoL.

In general, for each treatment, it is preferable for safety data to be derived from the same study that was used to determine efficacy. This ensures that the AEs accurately reflect those that are relevant to the treatment, as observed in the safety and efficacy assessment in clinical trials. Furthermore, using the same data source for safety and efficacy inputs avoids introducing uncertainty related to cross-study differences (e.g., differences in trial populations or drug administration).

The AE incidence data used in our presented model are informed by Grade ≥3 treatment-emergent AEs reported in the pooled analysis of ROCKstar and Phase 2a trials (for belumosudil) and Grade ≥3 AEs up to 24 weeks in the REACH-3 trial (for BAT).

Table 43 presents the frequencies of Grade ≥3 AEs as per the trials in terms of the percentage of patients that experienced each event. Please refer to Section B.2.10.3 for detailed results of the pooled analysis.

There was limited information available in the literature on AEs associated with ECP. Based on expert clinical opinion at the advisory board, an additional AE was included in the model to represent the central line infections associated with ECP treatment. The proportion of patients affected by this AE in the BAT arm was calculated based on the assumption that 64.6% of patients in the BAT arm are treated with ECP (Section B.3.2.3) and based on feedback from the advisory board that approximately 20% of patients who undergo ECP have a central line infection.(11)

Table 43. List of Grade ≥3 AEs occurring in >5%* patients in any treatment arm**

	Pooled ROCKstar and Phase 2a		REACH-3
	Belumosudil 200 mg once daily (n=81)	**Belumosudil 200 mg twice daily (n=75)**	**BAT (n=158)**
Pneumonia	8.6%	6.7%	9.5%
Hypertension	8.6%	6.7%	7.0%
Anaemia	4.9%	5.3%	7.6%
Thrombocytopenia and decreased platelet counts†	3.7%	0.0%	10.1%
Neutropenia	1.2%	1.3%	3.8%
Hyperglycaemia	6.2%	5.3%	1.9%
Gamma-glutamyl transferase increased	6.2%	2.7%	1.9%
Fatigue	2.5%	5.3%	1.9%
Central line-related infections	N/A	N/A	12.9%‡

BAT = best available therapy; ECP = extracorporeal photopheresis; N/A = not applicable TEAE = treatment-emergent adverse event
*Other than central line-related infections
**Lung function was not included as the definition and distinction from pneumonia is unclear, and it is not reported in a similar way across trials.
†For BAT, thrombocytopenia and decreased platelet count events were reported aggregated. For belumosudil, the category only includes decreased platelet count events as there were no Grade ≥3 TEAEs of thrombocytopenia.
‡Calculated value based on the assumptions that 64.6% of patients in the BAT arm are treated with ECP (Section B.3.2.3) and approximately 20% of patients have a central line-related infection based on feedback from the NICE advisory board.(11)

B.3.4. Measurement and valuation of health effects

In the model, utility values are applied to the failure-free and failure health states to capture patients' QoL associated with treatment and disease outcomes. Utility values are also applied to the different response categories (CR, PR, LR) within the failure-free health states.

B.3.4.1. Health-related quality of life data from clinical trials

The impact of belumosudil on QoL and symptom burden/bother was captured in ROCKstar by the LSS summary score assessed as an additional secondary endpoint and PROMIS-GH included as an exploratory endpoint (PRO instruments are described in detail in Appendix M).(73) At the August 2021 data cut, 62.9% of patients had a clinically meaningful improvement in LSS score, and 47.7% and 50.0% of patients reported improvements in PROMIS raw mental and physical health scores, respectively (Section B.2.6.1.2).(83)

To derive utility values for health economic modelling, PROMIS-GH results available for the FF state are mapped to EQ-5D using a published algorithm.(87)

B.3.4.2. Mapping

Utility values for the failure-free health states are based on an analysis of EQ-5D-3L utility scores that were mapped from the PROMIS-GH results of the ROCKstar trial (Appendix N). The model

differentiates between response and LR health states within the FF health state. Given the small number of observations of utility values specific to CR, utility values for CR and PR are derived based on pooled CR and PR observations and are assumed to be equal in the model. This introduces an element of uncertainty in the assumption that the utility of response is the same whether it is a CR or PR. However, the CR portion of the FF health state is small relative to PR and so this assumption is not expected to materially impact the results. Utility values for the FF health state are detailed in Table 44.

Table 44. Utility values for failure-free health states mapped from PROMIS results (ROCKstar)

Health state	Mean	SE	Source
Failure-free – In response – CR	[redacted]	0.007	ROCKstar trial post-hoc analysis, mapped from PROMIS to EQ-5D-3L
Failure-free – In response – PR	[redacted]	0.007	
Failure-free – LR	[redacted]	0.008	

CR = complete response; LR = lack of response; PR = partial response; SE = standard error

The scarcity of recorded QoL data following a failure event in the ROCKstar trial prevented us from generating meaningful mean utility estimates in the failure state. Indeed, only 23 patients had utility measurements recorded in 52 visits following a failure event. These data only capture the utility at the point of or soon after failure, rather than the full time spent in this health state (i.e., until death) as represented in the model. Section B.3.4.3 provides details of the efforts made to identify other sources for utility estimates for the failure state.

B.3.4.3. Health-related quality of life studies

An SLR was conducted with a cut-off date of December 2022 to identify studies reporting health utility values and HRQoL measures in patients with chronic GVHD (Appendix H). In total, 6,106 abstracts were identified from the SLR, including 2,690 records via Embase, 1,245 records via MEDLINE, and 2,171 records via CENTRAL. Six relevant publications were identified through hand-searching conference proceedings and HTA websites. Following title/abstract and full-text screening, a total of 94 publications, pertaining to 81 studies, were included in the SLR. Only three of the identified publications reported utility values in patients with chronic GVHD. An abstract by Lachance et al. 2021(44) reported the results of an international cross-sectional survey investigating QoL in patients with chronic GVHD. This study was based on DSP methodology(44) and used primary data from the UK, Australia, France, Canada and Switzerland, which we were able to access.(28, 31) A poster by Lee et al. 2021 reporting PRO results from the REACH-3 trial for ruxolitinib and a paper by Matza et al. 2020 reporting utilities associated with treatment approaches for transfusion-dependent β-thalassemia (TDT) including alloHSCT and the impact of chronic GVHD were also screened (Table 45). None of the studies reported utilities by response or FFS status for patients who had received at least two prior lines of systemic therapy. All three studies were therefore excluded from further consideration for the base case.

Table 45. Studies reporting utilities in patients with chronic GVHD identified by the SLR

Author	Year	Design	Study countries	Title
Lachance	2021	Cross-sectional	UK, Australia, France, Canada, Switzerland	Impact of chronic GVHD severity and steroid response on the quality of life in patients following allogeneic stem cell transplantation: Findings from a real-world study
Lee	2021	RCT	US, Austria, Australia, Belgium, Bulgaria, Canada, Czechia, Denmark, France, Germany, Greece, Hungary, India, Israel, Italy, Japan, Jordan, South Korea, Netherlands, Norway, Poland, Portugal, Puerto Rico, Romania, Russian Federation, Saudi Arabia, Spain, Sweden, Switzerland, Turkey, UK	Patient-reported outcomes (PROs) among patients with steroid-refractory or -dependent chronic graft-vs-host disease (cGVHD) randomized to ruxolitinib (RUX) vs. BAT
Matza	2020	Cross-sectional	England	Health state utilities associated with treatment for transfusion-dependent β thalassemia

BAT = best available therapy; GVHD = graft-versus-host disease; RCT = randomised controlled trial; SLR = systematic literature review; UK = United Kingdom; US= United States

None of the publications identified through hand-searching conference proceedings and HTA websites for chronic GVHD as part of the SLR were relevant for the base case. We therefore conducted additional searches of HTA websites in related disease areas (the indications for the most recent transplants [i.e., the underlying disease] of patients in ROCKstar) to identify relevant technology appraisals reporting utility values. We also conducted a utility elicitation exercise within the UK general population; however, these utility values were not used for the reasons described in Section B.1.3.1.4 and Appendix N. The utility estimate for the recurrent malignancy health state in our model was calculated as a weighted average based on utility values for progression/relapse health states in the relevant technology appraisals (see Appendix N for detailed information). Utility estimates from previous technology appraisals for recurrent malignancy are likely to be higher than may be expected for the population included in our model, since they do not account for the presence of advanced chronic GVHD as a comorbidity of the malignancy itself. Neither do they account for the severe mental and emotional burden of disease relapse following remission of the malignancy which the alloHSCT was intended to treat.

Unlike the published values for recurrent malignancy, there are no good estimates for the utility associated with the move from third to fourth-line treatment available in the literature. Therefore, we turned to clinical opinion. The clinicians we spoke to noted that, considering the time spent in the FFS health state and the likelihood of disease progression with recurrent malignancy in the absence of effective later line therapies, the utility of patients needing to start fourth-line treatment would be just as low as that of patients with recurrent malignancy. One clinician commented that the move to next therapy is likely prompted by a high number of infections and hospital inpatient attendances, a significant reduction in performance status and real concerns about the risk of death; all of which contribute to reduced QoL.(11) Therefore we have chosen to implement a uniform utility value for both a new systemic therapy after failure and recurrent malignancy. This assumption is tested in sensitivity analysis.

Mean utility values used in the base case and the corresponding standard errors for failure health states are presented in Table 46.

Table 46. Model utility values for failure health states based on literature

Health state	Mean	SE	Source
Failure - New chronic GVHD systemic therapy	0.479	0.036	Assumption
Failure - Recurrent malignancy	0.479	0.036	Calculated as a weighted average based on utility values of progression/relapse health state of indications for the most recent transplant (i.e., AML(112, 113), ALL(114), CML(115, 116), CLL(117))

ALL = acute lymphoblastic leukaemia; AML = acute myelogenous leukaemia; CLL = chronic lymphocytic leukaemia, CML = chronic myelogenous leukaemia; GVHD = graft-versus-host disease; SE = standard error

B.3.4.4. Adverse reactions

Given the lack of published disutility values identified by the SLR, AE disutility estimates were taken from technology appraisals in indications related to the underlying disease of patients in ROCKstar and are assumed to be the same for all treatments in the model. QALY loss for each AE was calculated by multiplying the associated disutility with the duration of the AE. Disutilities, duration per AE, and QALY loss associated with each AE are presented in Table 47. A disutility value for central line infections was not included in the base case as no estimates were identified in the literature. This results in a conservate overall estimate for disutility since we expect that 12.9% of patients on ECP will experience this event. The total one-off AE-related QALY loss associated with each treatment was calculated as the sum product of the disutility associated with each AE, the duration of experiencing the disutility and the rate of experiencing an AE with a given treatment. In a simplifying assumption, disutilities due to AEs are not considered for subsequent treatments.

Table 47. AE disutilities

AE	Mean disutility	SE*	Duration (days)	QALY loss	Source
Pneumonia	-0.195	0.039	18.2	-0.010	TA359(118)
Hypertension	-0.020	0.004	21.0	-0.001	TA689(119)
Anaemia	-0.090	0.018	23.2	-0.006	TA689(119)
Thrombocytopenia	-0.110	0.022	23.2	-0.007	TA689(119)
Neutropenia	-0.160	0.032	15.1	-0.007	TA689(119)
Hyperglycaemia	0.000	0.000	0.0	0.000	Assumption. Assume no disutility for abnormal lab test, consistent with an assumption used in TA642(112)
Gamma-glutamyl transferase increased	0.000	0.000	0.0	0.000	
Fatigue	-0.115	-0.023	30.4	-0.010	TA642(112)
Sepsis	-0.195	-0.039	23.2	-0.012	TA359(118)
Leukopenia	-0.090	-0.018	30.4	-0.007	TA642(112)
Dyspnoea	-0.050	-0.010	30.4	-0.004	TA642(112)
Central line-related infections	0.000	0.010	28.0	0.000	Assumption

AE = adverse event; ECP = extracorporeal photopheresis; QALY = quality-adjusted life year; SE = standard error; TA = technology appraisal
* Standard errors were not reported, are assumed to be 20% of the mean in the model.

B.3.4.5. Disutility associated with IV infusion

Based on clinical opinion received in the advisory board, a disutility was assumed for patients receiving IV treatments (ECP and rituximab).(11) This disutility was applied to all patients receiving IV treatments for as long as they are on treatment in the model, defined by the overall time on treatment curve for BAT, taking into account any maximum duration of treatment applied to the specific BAT components (i.e., capped by 4 weeks for rituximab; Table 38). The disutility associated with IV administration is based on Matza et al. 2013.(63) Disutility values are displayed in Table 48.

Table 48. Disutility associated with IV infusion

Parameter	Mean disutility	SE	Source
Disutility associated with IV infusion	-0.037	0.010	Matza et al. 2013(63)

IV = intravenous; QALY = quality-adjusted life year; SE = standard error

B.3.4.6. Caregiver disutility

Whilst not normally included in the reference case, the NICE manual (Section 4.3.17) states that all health effects for carers can be considered when the condition is associated with substantial effects on the carers' HRQoL.(120) Carers of patients with chronic GVHD who have progressed after two prior lines of therapy report that a substantial amount of time is spent looking after their loved one (28) and that, as the disease progresses, anxiety caused by the serious and potentially life-threatening nature of the condition is significant.(5, 6) Hence it is anticipated that caregivers will suffer significant disutility, but this has not been quantified in the literature. Nevertheless, this assumption is supported by feedback from English clinical experts consulted at the advisory board, who agreed that chronic GVHD is likely to have substantial emotional, financial and social impacts on the QoL of caregivers and that the effect on carers reported in the literature for MS is a good proxy for the impact of chronic GVHD.(11) Therefore, in the base case, the model assumes a caregiver burden for caregivers of patients in the FF health state with partial or lack of response, and for caregivers of patients in the failure state based on reported MS values. In the model calculations, the utility decrements of caregivers are added to the patients' utility score at each model cycle in the different health states. Utility decrements for caregiver burden were sourced from Acaster et al. 2013, a study on caregivers of patients with MS.(41) It was assumed that caregivers of patients in the failure-free PR and LR health states have similar disutilities as caregivers of MS patients in patient determined disease steps (PPDS) levels 2-3, and caregivers of patients in the failure states have similar disutilities as caregivers of MS patients with PPDS level 4, both of which are significant. These assumptions and the level of caregiver-related utility decrements were confirmed to be appropriate or even conservative by clinical experts at the advisory board. The respective caregiver-related utility decrements are shown in Table 49.

Table 49. Caregiver-related utility decrements

Health state	Mean	SE	Source
Failure-free			
Failure-free – In response – CR	0.000	0.000	Assumption
Failure-free – In response – PR	-0.045	0.057	Acaster et al. 2013(41)
Failure-free – LR	-0.045	0.057	
Failure			
Failure - New chronic GVHD systemic therapy	-0.142	0.062	Acaster et al. 2013(41)
Failure - Recurrent malignancy	-0.142	0.062	

CR = complete response; GVHD = graft-versus-host disease; LR = lack of response; PR = partial response; SE = standard error

B.3.4.7. Age- and gender-related utility adjustment

Utility values of the model health states were adjusted to account for the natural decrease in QoL associated with age. Adjusting utilities for age prevents the overestimation of benefits associated with treatment that can occur if otherwise perfect health is assumed at baseline.

Utilities were adjusted in the model using a multiplicative approach. General population utility estimates were obtained from the expected EQ-5D-3L values by age and sex published on the NICE DSU website.(121) These utility estimates are based on Adjusted Limited Dependent Variable Mixture Models (ALDVMM) of the Hernández Alava 2022(122) study, that used various regression models to estimate the change in utility values according to age and sex for the UK population. Hernández Alava 2022(122) uses the Health Survey for England (HSE) database from 2014 for calculation of the utility values, due to the fact that the 2014 version was the latest available HSE dataset before the COVID-19 pandemic that included EQ-5D-3L valuations.

For each health state, a utility decrement between the health state utility (Table 44 and Table 46) and the utility in the general population at the age at which the health state utility was measured was first calculated. For these calculations, the general population utility was estimated assuming a gender distribution equivalent to that of the modelled population at baseline (i.e., as a weighted average of gender-specific general population utility estimates, using 58.3% males and 41.7% females as weights based on baseline characteristics of the patients from the pooled ROCKstar and Phase 2a trials). This decrement was then applied consistently to the general population utility (which evolved as the population aged through the model time horizon) in order to obtain the age- and gender-adjusted health state utility at each model cycle.

B.3.4.8. Health-related quality of life data used in the cost-effectiveness analysis

A summary of all utility values used in the cost-effectiveness analysis is presented in Table 50.

Table 50. Summary of utility values for cost-effectiveness analysis

State	Utility value: mean	Standard error	Reference in submission (section and page number)	Justification
Failure-free – In response – CR	[redacted]	0.007	Section B.3.4.2 (page 109)	Based on analysis of EQ-5D-3L utility scores mapped from the PROMIS-GH results of the ROCKstar trial
Failure-free – In response – PR	[redacted]	0.007		
Failure-free – LR	[redacted]	0.008		
Failure - New chronic GVHD systemic therapy	0.479	0.036	Section B.3.4.3 (page 111)	Scarcity of utility data following a failure event in ROCKstar precluded generating meaningful utility estimates. Instead, estimates were derived from technology appraisals in other disease areas (indications for the most recent transplants of patients in ROCKstar) for the recurrent malignancy health state. New systemic therapy after failure state assumed to have the same utility values as recurrent malignancy based on clinician opinion received in advisory board.(11)
Failure - Recurrent malignancy	0.479	0.036		
AE disutilities				
Pneumonia	-0.195	0.039	Section B.3.4.4 (page 111)	To account for potential loss of QALYs due to adverse events. Included grade ≥3 AEs occurring in ≥5% of patients in any of the treatments included in the analysis (Phase 2a, ROCKstar or REACH-3). Disutility estimates taken from technology appraisals in indications related to the underlying disease of patients in ROCKstar, due to lack of published disutility values for AEs in chronic GVHD.
Hypertension	-0.020	0.004		
Anaemia	-0.090	0.018		
Thrombocytopenia	-0.110	0.022		
Neutropenia	-0.160	0.032		
Hyperglycaemia	0.000	0.000		
Gamma-glutamyl transferase increased	0.000	0.000		
Fatigue	-0.115	-0.023		
Sepsis	-0.195	-0.039		
Leukopenia	-0.090	-0.018		
Dyspnoea	-0.050	-0.010		
Central line-related infections	0.000	0.010		Assumption
Disutility associated with IV infusion	-0.037	0.010	Section B.3.4.5 (page 112)	Disutility for IV treatments included based on clinical opinion received in advisory board.(11) Disutility value based on Matza et al. 2013.(63)
Caregiver disutility				
Failure-free – In response – CR	0.000	0.000	Section B.3.4.6 (page 113)	Based on clinical opinion, a disutility for caregivers of patients on treatment was applied (assumed to exclude caregivers of patients with CR).(11) Disutility values sourced from Acaster et al. 2013.(41)
Failure-free – In response – PR	-0.045	0.057		
Failure-free – LR	-0.045	0.057		
Failure - New chronic GVHD systemic therapy	-0.142	0.062		
Failure - Recurrent malignancy	-0.142	0.062		

AE = adverse event; CEM = cost-effectiveness model; CR = complete response; ECP = extracorporeal photopheresis; EQ-5D-3L = European Quality of Life 5-Dimensions 3-Level version; GVHD = graft versus host disease; IV = intravenous; LR = lack of response; PR = partial response; PROMIS-GH = Patient-Reported Outcomes Measurement Information System Global Health scale

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

B.3.5.1. Resource identification, measurement and valuation studies

An SLR was conducted with a cut-off date of December 2022 to identify studies presenting the economic burden and healthcare resource utilisation (HCRU, e.g., hospital length of stay, intensive care length of stay, emergency department visits, prescriptions, etc.) in patients with chronic GVHD (Appendix I). A total of 4,419 abstracts were identified from the SLR, including 2,173 records via CENTRAL, 1,220 via Embase, 441 via MEDLINE®, 57 via NHS EED and 6 via EconLit. An additional 224 records were added through conference proceedings not indexed within Embase, 288 publications were identified through HTA websites, nine publications were added through hand-searching and one clinical trial was added from https://clinicaltrials.gov. After screening these records, a total of 23 primary studies, across 31 publications, were included for qualitative synthesis. Most studies were based in the US (n=11), two were multi-national, three were based in Sweden, two were based in France and one study was based in each of the following countries: Canada, Germany, Spain, Thailand and Tunisia. No studies were identified for England or other parts of the UK that could be used to inform our economic model.

Due to the unavailability of data on long-term disease management costs from the belumosudil clinical trials or SLR, disease management costs in the model were primarily estimated based on the results of our HES study (described in Section B.1.3.1.5).(45) Further details are provided in Section B.3.5.3.

B.3.5.2. Intervention and comparator costs and resource use

A summary of intervention and comparator's costs included in the cost-effectiveness analysis is presented in Table 51.

Table 51. Unit costs associated with the technology in the economic model

Items	Belumosudil 200 mg once daily	Belumosudil 200 mg twice daily	BAT	Reference in submission
Treatment acquisition cost (per cycle)	[redacted]	[redacted]	1st cycle: £4,285.56 2nd cycle: £4,265.76 3rd cycle: £4,265.76 4th to 6th cycle: £3,415.81 7th cycles onwards: £3,203.08	Section B.3.5.2.1 (page 116)
Administration cost (per cycle), including accommodation costs for ECP	0 (tablet taken at home)	0 (tablet taken at home)	1st cycle: £445.72 2nd cycle: £395.92 3rd cycle: £395.92 4th to 6th cycle: £316.74 7th cycles onwards: £296.94	Section B.3.5.2.2 (page 117)

BAT = best available therapy; ECP = extracorporeal photopheresis; NA = not applicable
Note, the model applies a maximum treatment duration of 5 years for the majority of treatments (except rituximab which has a treatment course duration of 4 weeks) in the base case, based on expert clinician opinion from the advisory board.(11)

B.3.5.2.1. Treatment acquisition

Treatment acquisition costs were calculated based on the package price costs and the drug quantity required based on the treatment schedules (detailed in Appendix K).

The package price cost was extracted from the electronic market information tool (eMIT) of the Department of Health and Social Care, from the NHS Electronic Drug Tariff Database or from the British National Formulary (BNF) database of NICE.(123-125) If multiple package sizes were available, then the package with the minimum price per mg was selected. The price per session for ECP was sourced from Button et al. 2021.(52) Package price for belumosudil was provided by Sanofi.

The treatment schedule for belumosudil was taken from ROCKstar. For the treatments considered in BAT, the treatment schedules were extracted from published literature and clinical expert guidance (as the treatment schedules are not available in the REACH-3 trial).

Treatment acquisition costs were calculated for every 4-week cycle. The costs were calculated using the drug package costs, dosing amount and the treatment administration schedule (Appendix K). The base case analysis considered no vial sharing (with wastage) for treatments where dosage was weight or BSA dependent (i.e., rituximab). Table 52 summarises estimates of treatment acquisition cost per 4-week cycle. The cost of BAT was estimated as a weighted average cost based on a distribution of BAT treatments and treatment acquisition cost of each treatment within the basket. The distribution of BAT components was derived based on the distribution of therapies in the BAT arm of the REACH-3 trial adjusted by removing treatments that were not relevant for England and re-allocating the proportions of the removed treatments to ECP (Section B.3.2.3). The distribution of BAT components based on adjusted values of the REACH-3 trial was aligned with feedback from the English clinical experts.(11)

Table 52. Treatment acquisition costs per 4-week cycle

Treatment	1st cycle (Weeks 1-4)	2nd cycle (Weeks 5-8)	3rd cycle (Weeks 9-12)	4th to 6th cycle (Weeks 13-24)	7th cycle onwards (Weeks 25+)
Belumosudil 200 mg once daily*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Belumosudil 200 mg twice daily*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAT	£4,285.56	£4,265.76	£4,265.76	£3,415.81	£3,203.08

BAT = best available therapy; ECP = extracorporeal photopheresis
*The presented costs for belumosudil are inclusive of the PAS discount of [redacted] to a pack price of £6,708.

B.3.5.2.2. Drug administration

Drug administration costs per cycle were calculated based on their respective route of administration and dosing schedule. Administration costs for oral drugs were assumed to be zero. ECP requires at least 2 hours of specialist nurse time per session.(11, 52) Administration costs for ECP were therefore assumed to be similar to the cost of 2 working hours for a specialist nurse, with the respective costs extracted from PSSRU 2021.(65) The cost of IV infusion for rituximab was based on a previous NICE technology appraisal (TA627)(126) which specifies two reference codes (SB13Z and SB15Z) from the list of NHS reference costs to correspond to the administration cost for rituximab (SB13Z: Deliver more Complex Parenteral Chemotherapy at First Attendance; SB15Z: Deliver Subsequent Elements of a Chemotherapy Cycle). The costs for these two reference codes were obtained from the latest list of NHS reference costs (2020-21)(127), and the cheapest was considered as the administration cost for rituximab in the model (£426.80 for each IV administration). The total administration costs per cycle for each comparator are displayed in Table 53.

Table 53. Administration costs per 4-week cycle

Treatment	1st cycle (Weeks 1-4)	2nd cycle (Weeks 5-8)	3rd cycle (Weeks 9-12)	4th to 6th cycle (Weeks 13-24)	7th cycle onwards (Weeks 25+)
Belumosudil 200 mg once daily	Tablet taken at home: No administration cost				
Belumosudil 200 mg twice daily	Tablet taken at home: No administration cost				
BAT*	£445.72	£395.92	£395.92	£316.74	£296.94

BAT = best available therapy; ECP = extracorporeal photopheresis
*Administration costs for BAT include the costs associated with overnight accommodation for ECP

The CEM does not consider monitoring costs associated with treatments.

B.3.5.2.3. Travel and accommodation for ECP

As ECP can only be administered at specialist centres, patients who are not local residents are expected to incur significant travel and accommodation costs.(11) It is our understanding that accommodation costs can be reimbursed by NHS England. Expert opinion was obtained to understand what proportion of patients on ECP would incur these costs. Clinical experts consulted at our advisory board estimated that 50% of patients undergoing ECP require overnight stays (often with

a family member).(11) As each ECP cycle consists of two ECP sessions administered on consecutive days, a one-night stay is expected. An accommodation cost of £150 per night was assumed, based on a previous CAR-T submission.(66)

Patients with chronic GVHD are usually eligible for NHS-funded hospital transport, and it is our understanding that some Trusts will reimburse public transport costs. Both of these forms of transport can be associated with significant distress for patients, many of whom limit their social contacts due to heightened risk of infection transmission, so many patients will travel to the clinic by car in order to avoid the distress and inconvenience.(11) Therefore, we did not include travel costs within the model.

B.3.5.3. Health state unit costs and resource use

B.3.5.3.1. Disease management costs

Disease management costs refer to the medical resource use associated with management of the disease, excluding any treatment costs or costs of treating treatment-related AEs. Due to the unavailability of data on long-term disease management costs from the clinical trials, disease management costs in the model were primarily estimated based on the results of the HES study (described in Section B.1.3.1.5), with an assumption that disease management costs for patients in the FF health state with PR and LR would decrease over time based on clinical opinion and a published real world study.(11, 45, 128)

Disease management costs were differentiated by health state in the model:

- **Patients in the failure-free health state with CR:** assumed to be the mean cost incurred by HSCT patients without GVHD in the HES study(45) throughout the time horizon of the model
- **Patients in the failure-free health state with PR and LR:** assumed to be the mean cost incurred by all HSCT patients with chronic GVHD in the HES study(45) in the first year, with a linear decrease in each year to reach the disease management cost of patients with CR in the fifth year. The model assumes that patients remaining failure-free incur the same costs regardless of response status after the fifth year
- **Patients in the failure state with a new systemic therapy**: assumed to incur the mean cost of HSCT patients with two or more records of high-cost therapy in the HES study. Treatments considered as high-cost therapy in the analysis included ECP, rituximab and protein tyrosine kinase inhibitors (i.e., ruxolitinib and imatinib)(45)
- **Patients in the failure state with recurrent malignancy.** These were not available from the HES study and so were sourced from TA642 that included the total costs incurred by patients with acute myelogenous leukaemia (AML)-related inpatient admissions, ICU, emergency department, outpatient visits, diagnostic procedures, lab tests, and blood transfusions(112). AML was the most common underlying malignancy in ROCKstar.

[redacted] Therefore, it was necessary to introduce some assumptions on the longer-term disease management costs. In particular, if patients have persisted in the FF health state for five years or more, the clinicians we consulted felt the remaining patients represent an enriched cohort who would very likely have ceased treatment due to physician advice or patient preference. Indeed, it could be the case that for a small number of patients, their chronic GVHD resolves within this time period. Clinicians told us that it is reasonable to assume that these patients would consume less and less healthcare resource over time.(11) This assumption is supported by the study from Schain et al in which costs for chronic GVHD patients were tracked over time in the Swedish healthcare setting and observed to decrease significantly.(128)

The yearly disease management costs by health state are presented in Table 54. Yearly disease management costs were converted to 4-weekly costs of disease management assuming that a year includes 52.14 weeks (Table 55).

Table 54. Summary of yearly disease management costs by health states

Health states	Mean cost per year					Source
	1st year	2nd year	3rd year	4th year	≥5th year	
Failure-free						
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Partial response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Failure						
New chronic GVHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Recurrent malignancy	£35,474.42	£35,474.42	£35,474.42	£35,474.42	£35,474.42	Calculated from TA642(112)

GVHD = chronic graft-versus-host disease; HES = Hospital Episode Statistics

Table 55. Disease management costs per 4-week cycles by health states

Health states	Mean cost per year					Source
	1st year	2nd year	3rd year	4th year	≥5th year	
Failure-free						
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Partial response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Failure						
New chronic GVHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Calculated from HES database (Section B.1.3.1.5)(45, 46)
Recurrent malignancy	£2,719.46	£2,719.46	£2,719.46	£2,719.46	£2,719.46	Calculated from TA642(112)

GVHD = chronic graft-versus-host disease; HES = Hospital Episode Statistics

B.3.5.3.2. Cost of subsequent treatments

For patients whose failure event was initiation of a new chronic GVHD systemic therapy, costs of subsequent lines of therapy following failure were included in the model. Evidence and clinical guidelines on treatments for chronic GVHD in fourth-line therapy and beyond were inconclusive (Section B.1.3.2). The treatment pathway is highly patient/case-dependent, and patients cycle through multiple therapies. Therefore, using simplifying assumptions, the base case implements a uniform distribution of subsequent treatments (based on the treatments in BAT) regardless of the initial treatment received by a patient, and patients spend an equivalent time on various treatment options, except for rituximab that has a treatment duration of 4 weeks. Clinicians in the advisory board explained that patients could stop receiving treatment altogether for reasons including patient choice and treatment burden outweighing benefits at later lines.(11) Therefore, it was assumed that patients spend 60% of their remaining lifetime on treatment. The approximate proportion of time spent on rituximab as a subsequent treatment was obtained by dividing the duration of a rituximab course by the mean time spent in the new chronic GVHD systemic therapy health state. The proportions of other treatment options were calculated by summing the proportions to 60%. The distribution of subsequent treatments in the model base case is shown in Table 56. Drug acquisition costs and drug administration costs for each subsequent treatment are multiplied by the proportion of patients receiving that subsequent treatment and are applied at each model cycle from the time of the Failure - new chronic GVHD systemic therapy event until death or the end of the time horizon, whichever is earlier.

Table 56. Distribution of subsequent treatments

	Initial treatment		
Subsequent treatment	**Belumosudil 200 mg QD**	**Belumosudil 200 mg BID**	**BAT**
ECP	14.5%	14.5%	14.5%
Mycophenolate mofetil	14.5%	14.5%	14.5%
Sirolimus	14.5%	14.5%	14.5%
Rituximab	2.0%	2.0%	2.0%
Imatinib	14.5%	14.5%	14.5%

BAT = best available therapy; BID = twice daily; ECP = extracorporeal photopheresis; QD = once daily

B.3.5.3.3. Cost of recurrent malignancy

The cost of treatment and management of a recurrent malignancy among patients with chronic GVHD who are currently on treatment is considered in the model. As evidence for the long-term cost of recurrence was beyond the scope of the clinical trials, published literature was used to inform this input. Given that a substantial proportion of patients in the ROCKstar trial received their initial transplant for AML (43.9%), a simplifying assumption was made that the cost of treating recurrence to AML will represent the cost of all recurrent malignancy in the model. A literature review was performed to identify the cost of treatment among AML patients with recurrence. Estimates of a one-time cost for post-progression treatment for AML patients were identified from TA642, which includes the cost of drug, test, event and the health state cost.(112) The costs were converted to 2021 GBP from the reported 2018 prices using the inflation factor of 1.08 and computed to be a one-time cost of £8,908.(65) This is likely to be conservative, since it does not include the costs of disease management for concomitant, late stage chronic GVHD on top of the recurrent malignancy.

B.3.5.3.4. Adverse reaction unit costs and resource use

In the base case, the cost of each AE was estimated using a micro-costing approach (by specifying resource use requirements and unit cost per resource use by adverse event). The AE cost for each treatment was calculated based on per event unit costs and the probability of experiencing AEs (Section B.3.3.7).The costs associated with the management of AEs were primarily derived from the National Schedule of NHS Costs (Year 2020-21) database.(127) Based on expert opinion from the advisory board with English clinical experts, it was assumed that all AEs can be managed in the outpatient setting, except central line-related infections which are treated in an inpatient setting.(11) Costs for central line-related infections experienced by a proportion of patients receiving ECP (12.9%; Section B.3.3.7) were sourced from Manoukian et al. 2021(129), based on the assumption that the cost of treating bloodstream infections (£5,917 in 2019) is an appropriate proxy. The costs were inflated to reflect 2021 price levels with an inflation index of 1.054(65), calculated to be a per event cost of £6,234. Per event unit costs are presented in Table 57. AE costs are applied as a one-off cost to the proportion of patients on treatment at the beginning of the model.

Table 57. Cost of adverse event management (2021 prices)

AE	Outpatient management unit cost per event	Inpatient management unit cost per event	Source
Pneumonia	£559.05	£2,644.23	National Schedule of NHS Costs (year 2020-21)
Hypertension	£518.93	£924.08	
Anaemia	£398.57	£1,667.92	
Thrombocytopenia	£414.46	£2,534.21	
Neutropenia	£366.66	£2,719.97	
Hyperglycaemia	£415.40	£1,103.36	
Gamma-glutamyl transferase increased	£415.40	£1,103.36	
Fatigue	£780.55	£1,579.39	
Sepsis	£302.39	£3,239.79	
Leukopenia	£557.42	£2,039.05	
Dyspnoea	£442.66	£971.68	
Central line-related infections	-	£6,234.04	Calculated based on Manoukian et al. 2021(129)

AE = adverse event; NHS = National Health Service

B.3.6. Severity

In order to investigate whether chronic GVHD in the third-line and beyond setting meets the criteria to be classified as a severe disease under the severity modifier framework, the lifetime QALY gain of patients receiving standard of care (assumed to be BAT, as estimated by the CEM) is expressed as a proportional (i.e., proportional QALY shortfall) and an absolute decrement (i.e., absolute QALY shortfall) of the estimated lifetime QALY gain of healthy patients of the same age and gender distribution. Results from the economic model base case and the subsequent calculation are presented in Table 58.

Table 58. Summary features of QALY shortfall analysis

Factor	Value	Reference
Sex distribution	58.3% males	Pooled ROCKstar and Phase 2a
Starting age	54.5 years	Pooled ROCKstar and Phase 2a
QALYs of population without the disease	14.432	Calculated by summing the product of the probability of being alive by age in the general population at each cycle of the model using the UK life tables(130) with the half-cycle corrected general population utilities over the model time horizon, adjusted for the model's four-week cycle length
QALYs with BAT	[redacted]	Estimated from the model (Section B.3.10)
Absolute QALY shortfall	[redacted]	Calculated
Proportional QALY shortfall	[redacted]	Calculated
QALY weight based on absolute QALY shortfall	1.2	NICE health technology evaluations: the manual (PMG36)(120)
QALY weight based on proportional QALY shortfall	1.2	NICE health technology evaluations: the manual (PMG36)(120)

BAT = best available therapy; NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year

This shows that chronic GVHD does meet the severity modifier criteria with a QALY weight of 1.2, demonstrating the significant impact on patients QoL, when using the utility values in the economic model base case. When using the public valuation from the utility elicitation exercise (described in Appendix N), the QALY shortfall analysis generates the highest 1.7 modifier, bringing the effective WTP to £51,000/QALY (see Section B.1.3.1.4 and Appendix N).

B.3.7. Uncertainty

Due to the ultra-orphan nature of chronic GVHD, there are several challenges with data availability and generation for this condition. Whilst the belumosudil trials are high quality in nature, the sample sizes are limited (N=132 in the ROCKstar study and N=54 in the Phase 2a study; Section B.2.2) and the studies are single armed.(73, 74) There is also a paucity of robust RCT evidence for existing treatments used and reimbursed in England for the treatment of chronic GVHD (Section B.1.3.2) with which to compare our uncontrolled Phase 2 data. An ITC to estimate the comparative efficacy of belumosudil against current treatments was not feasible due to lack of a common comparator arm between the trials, plus key differences between the study populations (Section B.2.9). In light of this, we attempted to construct an ECA based on real-world data to provide a comparison with the results of the Phase 2 belumosudil trials (Section B.2.9.2). However, methodological biases including the accurate identification of relapse due to coding issues along with the apparently very long time observed between treatment discontinuation and next treatment, were identified which significantly impact the interpretation of the findings and lead to limited face validity. This is explored further in Appendix M. The EAG suggested that, given the issues with and origin of the ECA data, the use of the REACH-3 BAT arm is likely to be the best option for a source of comparator data and will help to avoid the biases inherent in data collected from the real world (described further in Section B.3.2.3).

Furthermore, chronic GVHD is a complex and heterogeneous disease that affects multiple organs (Section B.1.3.1).(47) Due to the different organs that may be affected in each patient and the expected heterogeneity in multi-organ manifestations, the sample size per organ and combination of manifestations in our Phase 2 trials make it difficult to provide an analysis which examines all of the different combinations possible and so we have chosen to simplify the approach using all the available evidence to the best effect in an overarching analysis.

The prescription of medicines for chronic GVHD differs according to manifestation as well as access to treatment (e.g., proximity to ECP treatment centres; Section B.1.3.2). It is therefore necessary to consider BAT in the form of a basket of therapies as the standard of care. The REACH-3 BAT outcome data used in our model has not been adjusted to align with the UK-based BAT we derived from clinical opinion for costing purposes. However, clinicians told us that at this point in the pathway, outcomes are generally poor regardless of treatment. Nonetheless, we recognise that the BAT approach taken may lead to some uncertainty under the assumption that treatment outcomes are generalisable across BAT.

Limited utility-based HRQoL data were collected in the ROCKstar and Phase 2a studies. While PROMIS-GH was an exploratory endpoint in ROCKstar, the trial was not controlled and thus these data cannot be directly compared with a non-belumosudil-treated population. Furthermore, a lack of recorded utility data following a failure event precluded the generation of meaningful utility estimates in the failure state. In an effort to address this lack of utility-based HRQoL evidence, we conducted a utility elicitation exercise within the UK general population (described in Appendix N). While the study highlighted the substantial burden associated with chronic GVHD after two or more prior lines of systemic therapy from the perspective of the general public, these utility values were not appropriate for use in the model for the reasons described in Section B.1.3.1.4 and Appendix N. We therefore mapped PROMIS-GH outcomes from the ROCKstar study to EQ-5D-5L outcomes for patients in the FF state, with simplifying assumptions applied to the failure state (Section B.3.4), which can be considered conservative.

Finally, due to the ultra-orphan nature of the condition, there is a lack of published cost and healthcare resource use data for chronic GVHD in England (Section B.1.3.1.5). To address this, we conducted a large and robust study using the most up-to-date secondary care data from the HES database to estimate disease management costs for the submission model and validated our findings with expert clinicians at the advisory board.(11) Whilst this study provides the best available evidence from an English cohort of patients, the duration of follow-up was less than two years and so we have made assumptions about the decline in healthcare resource costs over time based on clinical opinion and literature precedent.

B.3.8. Summary of base case analysis inputs and assumptions

B.3.8.1. Summary of base case analysis inputs

A summary of base case analysis inputs used in the model is presented in Table 59. A summary of parameters varied in the PSA can be found in Appendix N.

Table 59: Summary of base case inputs

Setting	Base case	Rationale/Comments
Time horizon (Section B.3.2.2.6)	40 years	Long enough to capture treatment effect and benefits of belumosudil over the life expectancy of patients with chronic GVHD, given the average age of the population in the pooled ROCKstar and Phase 2a trials(83) was 55 years.
Discount rate for health outcomes and cost outcomes (Section B.3.2.2.6)	3.50%	In line with NICE guidance.(120)
Model population (Section B.3.2.1)	Patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy	In line with the indicated population.(1)
Model structure (Section B.3.2.2)	Partitioned survival model with response	Best use of the available data. Allows consideration of response outcomes, the primary endpoints in the ROCKstar(77) and REACH-3(58) trials, in addition to FFS. Accounts for progressive nature of disease and potential relapse
Proportion of belumosudil once daily vs. twice daily (Section B.3.2.3)	95% once daily and 5% twice daily	Based on feedback from an advisory board with expert clinicians in England, a low proportion of patients would continue on PPIs because of the steroid sparing effect of belumosudil, the relatively late stage of the disease and cost consciousness (famotidine is a suitable alternative for nearly all patients).(11) The advisors estimated that at steady state this would be around 5%.(11)
Comparator (Section B.3.2.3)	BAT	Due to the heterogenous nature of chronic GVHD and the prescription of different medicines according to manifestation and disease stage, it is appropriate to consider and model BAT as a basket of therapies.
Treatments within the BAT basket (Section B.3.2.3)	ECP (64.6%) Mycophenolate mofetil (22.2%) Imatinib (5.1%) Sirolimus (4.4%) Rituximab (3.8%)	Conservative assumption. Distribution of BAT in the REACH-3 study, adjusted to only retain treatments that are used in England (according to advisors) by redistributing the weights of the removed treatments to ECP. Weighted average of approximately 65% ECP calculated based on feedback from the NICE advisory board.
Source of efficacy and safety data: belumosudil (Section B.2.6.3)	Pooled analysis of ROCKstar and Phase 2a trials, in patients with chronic GVHD with 2 or more prior lines of therapy	Best available data. To maximise the sample size to estimate efficacy and safety inputs for belumosudil.
Source of efficacy and safety data: BAT	REACH-3	Best available data. Phase 3 clinical trial for BAT.

Setting	Base case	Rationale/Comments
(Section B.3.3)		
OS approach – belumosudil once daily (Section B.3.3.2.1)	Parametric: Joint Fit - Exponential distribution	Same type of parametric model applied across all comparators, based on guidance in NICE DSU Technical Support Document 14.(108) Exponential model selected to estimate OS, based on AIC and BIC fit statistics for belumosudil.
OS approach – belumosudil twice daily (Section B.3.3.2.1)	Parametric: Joint Fit - Exponential distribution	
OS approach – BAT (Section B.3.3.2.2)	Parametric: Joint Fit – Exponential distribution	
FFS approach – belumosudil once daily (Section B.3.3.1.1)	Parametric: Joint Fit – Generalised Gamma distribution	Same type of parametric model applied across all comparators, based on guidance in NICE DSU Technical Support Document 14.(108) Generalised gamma model selected to estimate FFS, based on AIC and BIC fit statistics for belumosudil.
FFS approach – belumosudil twice daily (Section B.3.3.1.1)	Parametric: Joint Fit – Generalised Gamma	
FFS approach – BAT (Section B.3.3.1.2)	Parametric: Joint Fit – Generalised Gamma	
Time on treatment approach – belumosudil once daily (Section B.3.3.6.1)	Parametric: Joint Fit – Log-Normal distribution	Log-normal model selected to estimate TTD for belumosudil, based on AIC and BIC fit statistics.
Time on treatment approach – belumosudil twice daily (Section B.3.3.6.1)	Parametric: Joint Fit – Log-Normal	
Time on treatment approach – BAT (Section B.3.3.6.2)	TTD curve was estimated by applying a HR to the TTD curve of belumosudil once daily. HR was derived based on the reported median TTD from REACH-3	
Time to response approach – belumosudil once daily (Section B.3.3.4.1)	Using Kaplan-Meier curve, no extrapolation	For BAT, only median TTR statistics were available from Le et al. 2022(110)
Time to response approach – belumosudil twice daily (Section B.3.3.4.1)	Using Kaplan-Meier curve, no extrapolation	
Time to response approach – BAT (Section B.3.3.4.2)	Derived from the reported median TTR (4 weeks) in Le et al.(110), assuming an exponential distribution	

Setting	Base case	Rationale/Comments
Duration of response approach – belumosudil once daily (Section B.3.3.5.1)	Parametric: Joint Fit – Log-Normal distribution	Log-normal distribution (joint fit) selected to estimate DOR, based on AIC and BIC fit statistics for belumosudil.
Duration of response approach – belumosudil twice daily (Section B.3.3.5.1)	Parametric: Joint Fit – Log-Normal distribution	
Duration of response approach – BAT (Section B.3.3.5.2)	Parametric: Joint Fit – Log-Normal distribution	
Utilities (Section B.3.4.8)	Failure-free, complete response: [redacted] Failure-free, partial response: [redacted] Failure-free, lack of response: [redacted] Failure, new systemic therapy: 0.479 Failure, recurrent malignancy: 0.479	Utility values for the failure-free health states are based on an analysis of EQ-5D-3L utility scores that were mapped from the PROMIS-GH scores collected in the ROCKstar trial. Due to limited number of observations, utility values for the failure health states were sourced from published sources.
Utility age and gender adjustment (Section B.3.4.7)	Multiplicative approach	A preferred adjustment method by NICE.(120)
AE disutilities (Section B.3.4.4)	Included	To account for loss of QALYs that could be due to AEs.
QALY loss due to AEs by treatment (one-off) (Section B.3.4.4)	Belumosudil once daily: 0.002 Belumosudil twice daily: 0.002 BAT: 0.002	Calculated as the sum product of the disutility associated with each AE considered in the model, the duration of experiencing the disutility, and the rate of experiencing an AE with a given treatment.
AEs included in the analysis (Section B.3.4.4)	Pneumonia, hypertension, anaemia, thrombocytopenia, neutropenia, hyperglycaemia, gamma-glutamyl transferase increase, fatigue, sepsis, leukopenia, dyspnoea, central line-related infections	Included Grade ≥3 AEs occurring in at least 5% of patients receiving any of the treatments included in the analysis (belumosudil, BAT) based on the incidence reported in the pooled analysis of the ROCKstar and Phase 2a trials(83) and REACH-3.(58)
Treatment acquisition costs (Section B.3.5.2.1)	Belumosudil [redacted] per cycle (once daily) Belumosudil [redacted] per cycle (twice daily) BAT £3,203.08 to £4,285.56 (depending on cycle)	Assumed [redacted] discount of the list price of £6,708 per pack for belumosudil.
Administration cost of oral drugs (Section B.3.5.2.2)	£0	Assumed that oral drugs do not require administration costs.

Setting	Base case	Rationale/Comments
Administration cost of IV drug and ECP (Section B.3.5.2.3)	ECP: £110 for nurse administration, £37.50 for accommodation costs, per ECP session Rituximab: £426.80 for each IV administration	ECP administration assumed to cost 2 hours specialist nurse time per session.(52) ECP accommodation cost based on a previous CAR-T submission.(66) Cost of rituximab IV infusion based on previous NICE technology appraisal (TA627).(126)
Disease management costs (per 4-week cycle) (Section B.3.5.3.1)	[redacted] Failure-free, partial response and lack of response: 1st year: [redacted] 2nd year: [redacted] 3rd year: [redacted] 4th year: [redacted] ≥5th year: [redacted] [redacted] Failure, recurrent malignancy: £2,719.46 in all cycles	Best available data. Costs for failure-free health states and failure, new systemic therapy were mapped from HES database.(45) Cost for recurrent malignancy was calculated from TA642.(112)
AE management costs by treatment (one-off cost) (Section B.3.5.3.4)	Belumosudil once daily: £186.94 Belumosudil twice daily: £142.48 BAT: £987.86 Central line-related infections: £6,234	Derived as a sum product of Grade ≥3 AE incidence and unit AE management costs(127) assuming all AEs (except central line-related infections) were treated in an outpatient setting. Based on expert opinion from the advisory board with English clinical experts, it was assumed that all AEs can be managed in the outpatient setting, except central line-related infections which are treated in an inpatient setting.(11) The model calculates a weighted average one-off cost that includes both inpatient and outpatient events.
Drug wastage (Section B.3.5.2.1)	Included	Relevant for rituximab where dosing was BSA-based.
Subsequent treatments included in the analysis (Section B.3.5.3.2)	ECP Mycophenolate mofetil Sirolimus Rituximab Imatinib	Standard of care in England for patients with chronic GVHD with 3 or more prior lines of treatments. Frequency inputs were based on the assumption that over their remaining lifetime patients spend 60% of their time on treatment.
Subsequent treatment frequency and duration (Section B.3.5.3.2)	ECP (14.5%, lifetime) Mycophenolate mofetil (14.5%, lifetime) Sirolimus (14.5%, lifetime) Rituximab (2.0%, lifetime) Imatinib (14.5%, lifetime)	No clear published recommendations on treatments for fourth-line and beyond. Thus, the assumption was that patients spend an equivalent time on different treatment options, except for rituximab, that has a one course treatment duration of 4 weeks. Therefore, it was assumed that patients spend around 2% of their time on rituximab, and the rest of the treatments were equally distributed to add up to 60%.

AE = adverse event; BAT = best available therapy; BSA = body surface area; CAR-T = Chimeric antigen receptor T-cell therapy; ECP = extracorporeal photopheresis; GVHD = graft versus host disease; IV = intravenous; NICE = National Institute for Health and Care Excellence; QALY= quality-adjusted life year

B.3.8.2. Assumptions

A summary of assumptions made in the model, alongside their justifications, is provided in Table 60.

Table 60. Summary of model assumptions

Assumptions	Justifications	Addressed in scenario analyses
95% of patients receiving once daily belumosudil, 5% twice daily (concomitant PPIs and/or strong CYP3A inducers)	Based on feedback from an advisory board with expert clinicians in England, a low proportion of patients would continue on PPIs because of the steroid sparing effect of belumosudil, the relatively late stage of the disease and cost consciousness (famotidine is a suitable alternative for nearly all patients).(11) The advisors estimated that at steady state this would be around 5%.(11)	A scenario with 10% of patients receiving PPIs and/or strong CYP3A inducers was explored (Table 65: Scenario #6)
Assume same cycle probability of death after 5 years as for BAT	To account for the uncertainty in the long-term treatment benefit of belumosudil. As discussed in Section B.3.3.2.3, this assumption was validated with clinicians at the advisory board.(11)	Use of the parametric survival curve fitted to belumosudil data for the entire time horizon was explored in a scenario analysis (Table 65: Scenario #13)
Utility of failure state among patients receiving new systemic therapy was assumed to be the same as the value for recurrent malignancy	This assumption was made given the scarcity of utility data following a failure event in ROCKstar, and was supported by clinician opinion received in advisory board.(11)	Alternate sources/assumptions on utility value for new systemic therapy state were explored in scenario analysis (Table 65: Scenario #48, 49, 50)
Five-year maximum duration of treatment for all treatments except rituximab (i.e., belumosudil, ECP, mycophenolate mofetil, imatinib, and sirolimus)	Based on feedback from the advisory board that patients with chronic GVHD who are stable and responding to treatment are unlikely to be on treatment beyond 5 years.(11)	Two different scenarios were explored: 1) Maximum treatment duration of three years; 2) No limit in treatment duration (i.e., purely driven by time to treatment discontinuation curve) (Table 65: Scenario #30, 31)
Disease management costs based on HES data, with a linear decline over 5-years to reach costs of CR	If patients have persisted in the FF state for five years or more the clinicians, we consulted felt the remaining patients represent an enriched cohort who would very likely have ceased treatment due to physician advice or patient preference. Indeed, it could be the case that for a small number of patients their chronic GVHD resolves within this time period. Clinicians told us that it is reasonable to assume that these patients would consume less and less healthcare resource over time. This assumption is supported by the study from Schain et al in which costs for chronic GVHD patients were tracked over time in the Swedish healthcare setting and observed to decrease significantly.(128) The model assumes that patients remaining failure-free incur the same costs regardless of response status after the fifth year	Different assumptions on disease management cost (i.e., cost values and the pattern of cost change over time to reflect change in health care resource use) were explored in several scenarios (Table 65: Scenario# 58-60)
AEs treated in outpatient settings (except central line-related infections)	All AEs are conservatively assumed to be managed in the outpatient setting.	A scenario where AEs are treated 50% in outpatient and 50% in inpatient settings (except central

Assumptions	Justifications	Addressed in scenario analyses
		line-related infections) was explored. (Table 65: Scenario #43)
Among patients entering failure state with new systemic therapy, it was assumed that patients spend 60% of their remaining lifetime on chronic GVHD treatments	Evidence and clinical guidelines on treatments for chronic GVHD in fourth-line therapy and beyond were inconclusive. Our assumption was supported by feedback from the clinicians in the advisory board which indicated patients may not continue to be on treatment over lifetime and could stop receiving treatment altogether for reasons including patient choice and treatment burden outweighing benefits at later lines.(11)	A scenario assuming patients spending 100% of remaining life years on subsequent treatments was explored (Table 65: Scenario #27)
The model accounted for disutility associated with caregiver burden for caregivers of patients in the FF health state with partial or lack of response, and in the failure state, using values based on reported values based on multiple sclerosis.	This assumption is supported by feedback from English clinical experts consulted at the advisory board, who agreed that chronic GVHD is likely to have substantial emotional, financial and social impacts on the QoL of caregivers and that the effect on carers reported in the literature for multiple sclerosis is a good proxy for the impact of chronic GVHD.(11)	Scenario analyses using different values for caregiver disutility and excluding caregiver disutility were performed (Table 65: Scenario #58-60)
Disutility was applied over treatment duration for patients receiving IV treatments (ECP and rituximab)	This assumption was supported by the clinical opinion received in the advisory board (11)	Scenario analyses using alternate value for disutility and excluding IV infusion disutility were performed (Table 65: Scenario #56-57)

AE = adverse event; BAT = best available therapy; CR = complete response; CYP3A = Cytochrome P450, family 3, subfamily A; ECP = extracorporeal photopheresis; FF = failure free; GVHD = graft versus host disease; HES = Hospital Episode Statistics; IV = intravenous; PPI = proton pump inhibitor; QoL = quality of life

B.3.9. Base case results

Results from the probabilistic analysis for the base case without applying a severity modifier are presented in Table 61. Table 62 presents the base case results with a severity modifier, in which the QALY weight of 1.2 (see Section B.3.6) was applied to the average QALYs obtained from the PSA. The PSA results for each treatment were generated over 5,000 iterations. Parameters varied in the analysis are detailed in Appendix N.

Compared to BAT, average incremental costs per patient with belumosudil were estimated to be [redacted], with average incremental LYs of [redacted]. Incremental QALYs were estimated to be [redacted] without the severity modifier and [redacted] when the severity modifier was applied. This leads to mean probabilistic ICERs of £15,032 (without severity modifier) and £12,526 (with severity modifier) per QALY. The base-case ICERs generated from probabilistic analysis are very close to the deterministic base-case ICERs (Section B.3.10). Based on WTP thresholds of £20,000 or £30,000 per QALY, the incremental net health benefit was estimated at [redacted] and [redacted], respectively, when the severity modifier was not included, and [redacted] and [redacted], respectively, with the severity modifier. Based on WTP thresholds of £20,000 or £30,000 per QALY, the incremental net monetary benefit was estimated at [redacted] and [redacted], respectively, when the severity modifier was not included, and [redacted], and [redacted], respectively, with the severity modifier.

Table 61. Base-case probabilistic results (PAS, without severity modifier)

Outcome	Belumosudil	BAT
Total costs	[redacted]	£251,222
Total LYs	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
Incremental costs	[redacted]	
Incremental LYs	[redacted]	
Incremental QALYs	[redacted]	
ICER (£/QALY)	£15,032	
INHB (£20,000/QALY)	[redacted]	
INHB (£30,000/QALY)	[redacted]	
INMB (£20,000/QALY)	[redacted]	
INMB (£30,000/QALY)	[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

Table 62. Base-case Probabilistic Results (PAS, with severity modifier [1.2 QALY weight])

Outcome	Belumosudil	BAT
Total costs	[redacted]	£251,222
Total LYs	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
Incremental costs	[redacted]	
Incremental LYs	[redacted]	
Incremental QALYs	[redacted]	
ICER (£/QALY)	£ 12,526	
INHB (£20,000/QALY)	[redacted]	
INHB (£30,000/QALY)	[redacted]	
INMB (£20,000/QALY)	[redacted]	
INMB (£30,000/QALY)	[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

B.3.10. Base case deterministic cost-effectiveness analysis results

Results from the deterministic analysis for the base case are presented in Table 63 (without severity modifier) and Table 64 (with severity modifier [1.2 QALY weight]). Compared to BAT, patients treated with belumosudil were estimated to have [redacted] LYs gained along with increased costs of [redacted] per patient. Incremental QALYs were estimated at [redacted] and [redacted], without and with severity modifier, respectively. These led to ICERs per QALY gained of £15,086 (without severity modifier) and £12,572 (with severity modifier). Based on WTP thresholds of £20,000 and £30,000 per QALY, the incremental net health benefit was estimated at [redacted] and [redacted], respectively, when the severity modifier was not included, and [redacted] and [redacted], respectively, with the severity modifier. Based on WTP thresholds of

£20,000 and £30,000 per QALY, the incremental net monetary benefit was estimated at [redacted] and [redacted], respectively, when the severity modifier was not included, and [redacted], and [redacted], respectively, with the severity modifier.

Table 63. Base-case deterministic results (PAS, without severity modifier)

Outcome	Belumosudil	BAT
Total costs	[redacted]	£251,396
Total LYs	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
Incremental costs	[redacted]	
Incremental LYs	[redacted]	
Incremental QALYs	[redacted]	
ICER (£/QALY)	£15,086	
INHB (£20,000/QALY)	[redacted]	
INHB (£30,000/QALY)	[redacted]	
INMB (£20,000/QALY)	[redacted]	
INMB (£30,000/QALY)	[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; INHB = incremental net health benefit; INMB = incremental net monetary benefit; LY = life year; QALY = quality-adjusted life year

Table 64. Base case deterministic results (PAS, with severity modifier [1.2 QALY weight])

Outcome	Belumosudil	BAT
Total costs	[redacted]	£251,396
Total LYs	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]
Incremental costs	[redacted]	
Incremental LYs	[redacted]	
Incremental QALYs	[redacted]	
ICER (£/QALY)	£12,572	
INHB (£20,000/QALY)	[redacted]	
INHB (£30,000/QALY)	[redacted]	
INMB (£20,000/QALY)	[redacted]	
INMB (£30,000/QALY)	[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; INHB = incremental net health benefit; INMB = incremental net monetary benefit; LY = life year; QALY = quality-adjusted life year

Disaggregated deterministic health and cost outcomes are available in Appendix J. Base-case deterministic results at list price are available in Appendix N.

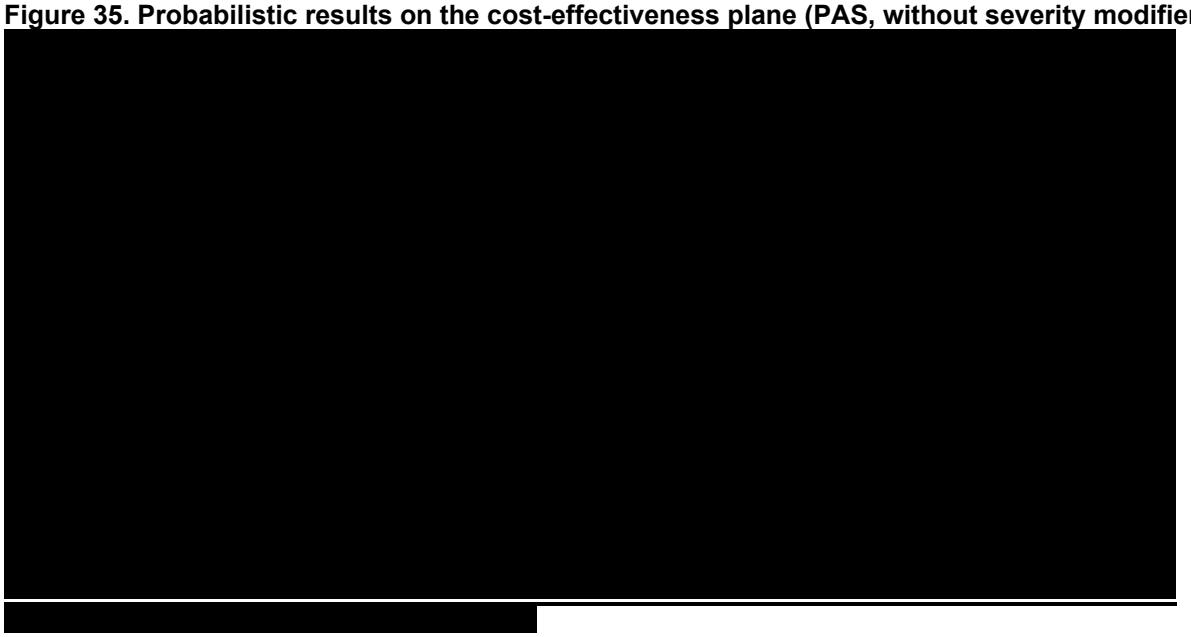
B.3.11. Exploring uncertainty

B.3.11.1. Probabilistic sensitivity analysis

In addition to the results from the probabilistic analysis presented in Table 61, the results for the probabilistic analysis (without severity modifier) are plotted on the cost-effectiveness plane for

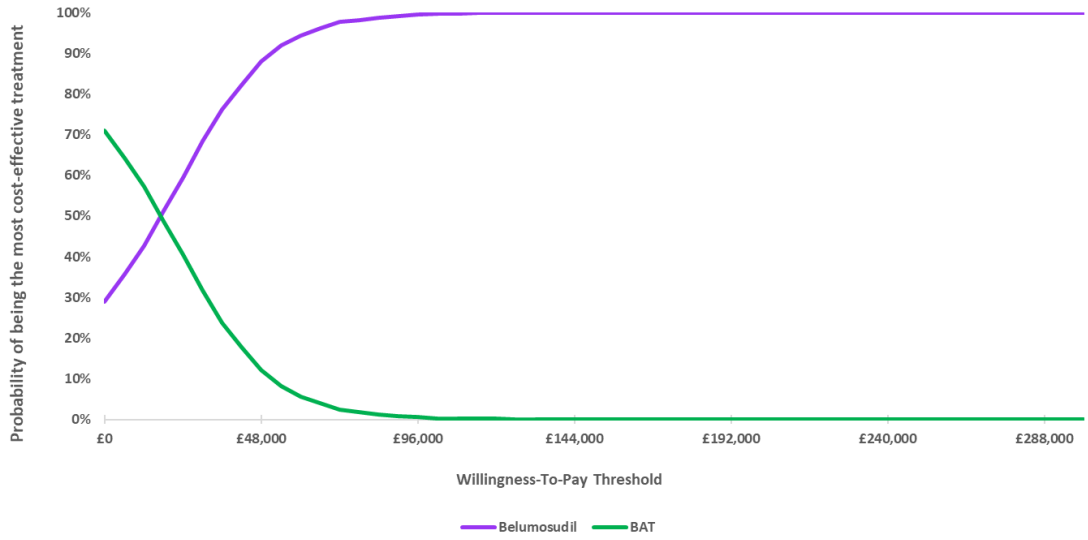
belumosudil compared to BAT over 5,000 iterations in Figure 35. The majority of iterations (71%) fell in the north-east quadrant, indicating that treatment with belumosudil was consistently more effective and more costly than BAT. However, 29% of the iterations fell in the south-east quadrant which indicates belumosudil being a dominant treatment option, providing better health outcomes at lower costs compared to BAT.

Figure 35. Probabilistic results on the cost-effectiveness plane (PAS, without severity modifier)



The cost-effectiveness acceptability curves in Figure 36 show the probability of each treatment being an optimal treatment choice across a range of WTP thresholds. It suggests that belumosudil is more likely to be the optimal treatment choice vs. BAT at WTP thresholds above £18,000/QALY. The probability of belumosudil being the optimal treatment choice vs. BAT at a £30,000/QALY WTP threshold is 68.3%.

Figure 36. Cost-effectiveness acceptability curves for belumosudil vs. BAT (PAS, without severity modifier)



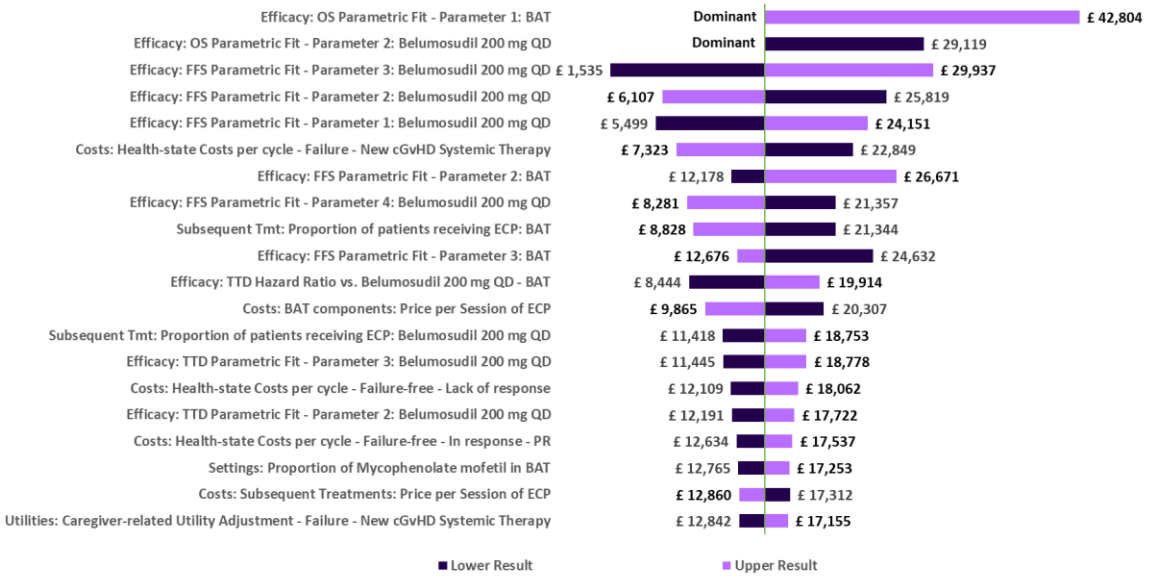
BAT = best available therapy; PAS = patient access scheme

B.3.11.2. One-way deterministic sensitivity analysis

A one-way deterministic sensitivity analysis (DSA) was performed to investigate key drivers of the base-case results. In this analysis, input parameters were individually increased and decreased with deterministic results generated for the higher and lower values. The higher and lower values were based on 95% CIs or published ranges. In the absence of such data, the higher and lower values were calculated as ± 20% of the mean base-case value. Parameters varied in the analysis are detailed in Appendix N.

The 20 most influential parameters on the ICER of belumosudil vs. BAT are shown in Figure 37. The most influential parameters included those related to OS and FFS estimates for belumosudil 200 mg once daily and for BAT, disease management costs in the failure health state for patients receiving new chronic GVHD systemic therapy, proportion of ECP among subsequent treatments, and cost of ECP. Parameters related to TTD also appeared among the top 20 most influential factors, though with a smaller impact on the ICER compared to those related to OS and FFS.

Figure 37. Tornado diagram of ICER (incremental cost per QALY gained) for belumosudil vs. BAT (PAS, without severity modifier)



BAT = best available therapy; BID = twice daily; cGvHD = chronic graft-versus-host disease; ECP = extracorporeal photopheresis; FFS = failure-free survival; ICER = incremental cost-effectiveness ratio; OS = overall survival; PAS = patient access scheme; PR = partial response; QALY = quality-adjusted life year; QD = once daily; Tmt = treatment; TTD = time to treatment discontinuation
Notes: FFS Parametric Fit – Parameter 1 = mu (Generalised Gamma distribution); FFS Parametric Fit – Parameter 2 = sigma (Generalised Gamma distribution); FFS Parametric Fit – Parameter 3 = Q (Generalised Gamma distribution); FFS Parametric Fit – Parameter 4 = treatment coefficient; OS Parametric Fit – Parameter 1 = rate (exponential distribution); OS Parametric Fit – Parameter 2 = treatment coefficient; TTD Parametric Fit – Parameter 1 = mean (log-normal distribution); TTD Parametric Fit – Parameter 2 = standard deviation (log-normal distribution); TTD Parametric Fit – Parameter 3 = treatment coefficient

DSA results for incremental costs and incremental QALYs are available in Appendix N.

B.3.11.3. Sensitivity and scenario analysis

The ICERs for each sensitivity analysis, with and without the application of the severity modifier (i.e., QALY weight of 1.2), alongside the percentage change from the base-case ICERs, are presented in Table 65. Detailed description of the scenarios is included in Appendix N.

Deterministic ICERs from the sensitivity and scenario analyses listed ranged from -71% to 489% variance from the base-case ICER without applying a severity modifier and ranged from -80% to 522% of the base-case ICER when applying a 1.2 QALY weight. In some scenarios around OS estimates, belumosudil was dominant vs. BAT. The majority of sensitivity analyses fell below the threshold of £30,000 per QALY gained (even more so if the 1.2 QALY weight was applied).

The most sensitive scenarios are those where different fitted distributions were used for OS and FFS extrapolation. Belumosudil became dominant, providing higher number of QALYs at lower cost compared to BAT in all scenarios using alternate distributions for OS extrapolation. On the other hand, the ICERs increased up to £88,794 per QALY (without a severity modifier) when different

extrapolations were used for FFS. However, we believe the choice of the extrapolations used in the base case is robust because they were based on the lowest AIC/BIC combinations. Unlike many oncology economic models where fits are based on very limited data, these data sets are more than 50% complete providing confidence that the fit statistics can be used reliably. The fits were also chosen by the clinicians we consulted.

Scenarios that resulted in ICERs above £30,000 per QALY without a severity modifier included those related to treatment duration (scenarios #21 and 31), scenario without OS waning at 5 years (scenario #13), scenario with 5-year time horizon (scenario #2), and scenarios with different distributions for FFS extrapolation (scenarios #8-10). Only four of these (scenarios #2, 8-10) had ICERs above £30,000 per QALY when applying the severity modifier.

Table 65. Sensitivity analyses with percentage change from base-case ICER (PAS)

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
	Base case	£15,086	£12,572	
1	Time horizon: 20 years	£19,046	£15,872	+26.25%
2	Time horizon: 5 years	£78,249	£65,207	+418.69%
3	Discount rates: 0%	£4,398	£3,665	-70.85%
4	Discount rates: 1.5%	£9,122	£7,602	-39.53%
5	No response	£14,894	£12,411	-1.27%
6	Proportion of patients receiving PPIs and/or strong CYP3A inducers: 10%	£17,725	£14,771	+17.50%
7	Alternative distribution of BAT components	£21,485	£17,904	+42.42%
8	FFS for all treatments: Joint Fit - Gamma	£88,794	£73,995	+488.59%
9	FFS for belumosudil QD and BID: Joint Fit - Log-normal; FFS for BAT: Joint Fit - Weibull	£55,417	£46,181	+267.35%
10	FFS for belumosudil QD and BID: Independent Fit - Log-normal; FFS for BAT: Independent Fit - Gamma	£39,632	£33,027	+162.71%
11	FFS - Distribution of failure events for BAT assumed to be the same as for Belumosudil QD arm, in all periods	£17,155	£14,296	+13.71%
12	FFS distribution - 100% new systemic therapy after 12 months	£15,240	£12,700	+1.02%
13	OS long-term assumption for belumosudil: Do not assume same probability of death as for BAT after 5 years.	£35,025	£29,187	+132.17%

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
14	OS for all treatments: Joint Fit - Log-normal	Belumosudil is Dominant, Incremental QALYs: 1.325, Incremental Costs: -£14,034	Belumosudil is Dominant, Incremental QALYs: 1.59, Incremental Costs: -£14,034	
15	OS for all treatments: Joint Fit - Log-logistic	Belumosudil is Dominant, Incremental QALYs: 1.378, Incremental Costs: -£6,337	Belumosudil is Dominant, Incremental QALYs: 1.653, Incremental Costs: -£6,337	
16	OS for all treatments: Joint Fit - Weibull	Belumosudil is Dominant, Incremental QALYs: 1.368, Incremental Costs: -£5,039	Belumosudil is Dominant, Incremental QALYs: 1.642, Incremental Costs: -£5,039	
17	OS for belumosudil QD and BID: Independent Fit - Log-normal; OS for BAT: Independent Fit - Log-normal	Belumosudil is Dominant, Incremental QALYs: 1.323, Incremental Costs: -£14,263	Belumosudil is Dominant, Incremental QALYs: 1.588, Incremental Costs: -£14,263	
18	TTD for belumosudil QD and BID: Joint Fit - Generalised Gamma	£15,370	£12,808	+1.88%
19	TTD for belumosudil QD and BID: Joint Fit - Exponential	£15,830	£13,192	+4.93%
20	TTD for BAT: exponential curve fitted to median	£18,900	£15,750	+25.28%
21	Treat until failure (all treatments)	£30,985	£25,820	+105.39%
22	Treat until failure (BAT only)	£7,015	£5,846	-53.50%
23	DOR for all treatments: Joint Fit - Log-logistic	£15,060	£12,550	-0.17%
24	DOR for all treatments: Joint Fit - Weibull	£15,268	£12,723	+1.21%
25	DOR for belumosudil QD and BID: Independent Fit - Log-normal; DOR for BAT: Independent Fit - Log-normal	£15,075	£12,563	-0.07%
26	No AE for central line-related infections	£15,701	£13,084	+4.08%
27	Alternate distribution of subsequent treatments (applied for all initial treatments)	£5,745	£4,787	-61.92%

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
28	Alternate approach to costing of subsequent treatments	£28,481	£23,734	+88.79%
29	Full vial sharing	£15,094	£12,578	+0.05%
30	Maximum duration of treatment for all treatments (except rituximab): 3 years	£6,536	£5,446	-56.68%
31	No maximum duration of treatment for all treatments (except rituximab)	£30,430	£25,358	+101.71%
32	Alternate proportions of responders to ECP assumed for drug cost calculations	£20,452	£17,043	+35.57%
33	No accommodation cost is reimbursed	£15,703	£13,086	+4.09%
34	Disease management costs for all Failure-free health states follow the decrease observed in Schain et al. 2021(128)	£5,157	£4,297	-65.82%
35	Disease management costs for all Failure-free health states reduced in Years 5+	£5,748	£4,790	-61.90%
36	Disease management costs for Failure-free - CR assumed to be higher than base case in first year with a linear decline to base case values in Year 5+	£15,013	£12,511	-0.48%
37	Disease management costs for Failure-free - PR assumed to be equal to costs of Failure-free - CR	£12,726	£10,605	-15.64%
38	Disease Management Costs for Failure-free - PR and for Failure-free - Lack of response assumed to be higher than base case in first year with a linear decline to base case values in Year 5+	£16,972	£14,143	+12.50%
39	Disease Management Costs for Failure-free - PR and for Failure-free - Lack of response assumed to be equal to costs of Failure-free - CR from Year 3 onwards	£13,219	£11,016	-12.37%
40	Disease Management Costs for Failure-free - PR and for Failure-free - Lack of response assumed to be equal to costs of Failure-free - CR from Year 5 onwards with no reduction from Year 1 to Year 4	£17,858	£14,882	+18.38%
41	Disease Management Costs for Failure-free - Lack of response assumed to be higher than	£15,994	£13,329	+6.02%

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
	base case in first year with a linear decline to base case values in Year 5+			
42	Disease management costs for Failure - New chronic GVHD Systemic Therapy: based on chronic GVHD patients with 1 high-cost therapy	£18,658	£15,548	+23.68%
43	Alternate AE management setting	£15,040	£12,533	-0.31%
44	Health state utilities: Adelphi DSP data(28, 31)	£17,262	£14,385	+14.42%
45	Health state utilities: Lachance et al. 2021(44)	£18,663	£15,553	+23.71%
46	Health state utilities: Crespo et al. 2012(98)	£18,781	£15,650	+24.49%
47	Different health state utilities for CR and PR	£15,065	£12,554	-0.14%
48	Health state utility for Failure - New chronic GVHD Systemic Therapy: Adelphi DSP data (using absolute value)(28, 31)	£15,846	£13,205	+5.04%
49	Health state utility for Failure - New chronic GVHD Systemic Therapy: Adelphi DSP data (using decrement)(28, 31)	£16,764	£13,970	+11.12%
50	Health state utility for Failure - New chronic GVHD Systemic Therapy: Crespo et al. 2012(98)	£19,618	£16,348	+30.04%
51	Value of health state utility for Failure - Recurrent Malignancy and for Failure - New chronic GVHD Systemic Therapy: apply utility decrement	£12,923	£10,770	-14.33%
52	Value of health state utility for Failure - Recurrent Malignancy and for Failure - New chronic GVHD Systemic Therapy: lower bound of range	£12,756	£10,630	-15.45%
53	Value of health state utility for Failure - Recurrent Malignancy and for Failure - New chronic GVHD Systemic Therapy: upper bound of range	£19,000	£15,833	+25.95%
54	No age- and gender-related utility adjustment	£14,840	£12,367	-1.63%
55	No AE disutilities	£15,312	£12,760	+1.50%
56	Alternate disutility associated with IV infusion treatment	£14,999	£12,499	-0.58%
57	No disutility associated with IV infusion treatment	£15,305	£12,754	+1.45%

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
58	Same caregiver-related disutilities (-0.045) are applied in PR, Lack of response, and Failure health states	£16,880	£14,066	+11.89%
59	Same caregiver-related disutilities (-0.142) are applied in PR, Lack of response, and Failure health states	£18,639	£15,533	+23.56%
60	No caregiver-related disutilities	£16,171	£13,476	+7.20%

AE = adverse event; BAT = best available therapy; BID = twice daily; CR = complete response; CYP3A = cytochrome P450, family 3, subfamily A; DOR = duration of response; DSP = Disease Specific Programmes; ECP = extracorporeal photopheresis; FFS = failure-free survival; HES = Hospital Episode Statistics; ICER = incremental cost-effectiveness ratio; OS = overall survival; PPI = proton pump inhibitor; PR = partial response; QD = once daily; TTD = time to treatment discontinuation

B.3.11.3.1. Comparison of costs vs. ECP

ECP is an expensive procedure that consumes considerable time and resource because patients are required to attend for infusion on consecutive days. This is unpleasant and inconvenient for patients and both costly and burdensome to the NHS, which is facing unprecedented resourcing and capacity challenges. Due to the cost of the procedure, it is worth considering a side-by-side naïve comparison of the costs for ECP and belumosudil treatment over 1 year. (Note, in the economic model, the median time on treatment for belumosudil was 10 months).

According to clinicians we have spoken to, ECP continues to form the mainstay of treatment for chronic GVHD patients at 3rd line.(11) The number of ECP sessions per cycle used in the economic model are shown in Table 66. This is based on a consensus statement update from the UK Photopheresis Society and feedback from clinicians in England.(11, 51)

Table 66. Number of ECP sessions per cycle

Model cycle	Timeframe	Number of ECP sessions
1st cycle	Weeks 1-4	4
2nd cycle	Weeks 5-8	4
3rd cycle	Weeks 9-12	4
4th to 6th cycle	Weeks 13-24	3.2
7th cycle onwards	Weeks 25+	3

ECP = extracorporeal photopheresis

One cycle corresponds to two treatments on consecutive days, typically performed via peripheral venous access catheter; however, when this is not possible, a central-venous catheter must be used. The costs associated with acquisition, administration and accommodation for ECP treatment are reproduced in Table 67.

Table 67. Acquisition and associated costs for ECP

	Cost	Source
Cost of an ECP session (Appendix K)	£1,585	Cost taken from Button et al.2021.(52)
Administration cost (Section B.3.5.2.3)	£110	ECP administration assumed to cost 2 hours specialist nurse time per session.(52)
Accommodation costs (Section B.3.5.2.3)	£150	ECP accommodation cost based on a previous CAR-T submission.(66) Note that two treatments on consecutive days are given and accommodation costs are pro rata per session in the model (£75 per session)
Proportion of patients with accommodation costs	50%	Based on feedback from clinicians.(11)

ECP = extracorporeal photopheresis

Assuming 1 year is equivalent to 13 x four weekly cycles the total annual ECP costs for acquisition, administration and accommodation are presented in Table 68.

Table 68. Total costs for 1 year of ECP treatment

Cycle	No. of ECP sessions per cycle	Acquisition cost per cycle	Administration Cost per cycle	Accommodation costs per cycle	Total cost
1	4	£6,340	£440	£150	£6,930
2	4	£6,340	£440	£150	£6,930
3	4	£6,340	£440	£150	£6,930
4	4	£6,340	£440	£150	£6,930
5	3.2	£5,072	£352	£120	£5,544
6	3.2	£5,072	£352	£120	£5,544
7	3	£4,755	£330	£113	£5,198
8	3	£4,755	£330	£113	£5,198
9	3	£4,755	£330	£113	£5,198
10	3	£4,755	£330	£113	£5,198
11	3	£4,755	£330	£113	£5,198
12	3	£4,755	£330	£113	£5,198
13	3	£4,755	£330	£113	£5,198
Total	**43.4**	**£68,789**	**£4,774**	**£1,628**	**£75,191**

The acquisition cost for belumosudil per cycle used in the model is [redacted] for once daily dosing and [redacted] for twice daily dosing (Table 52). We have assumed 5% of patients would receive belumosudil twice daily based on feedback from clinicians(11) and so the weighted average cost of belumosudil per cycle can be assumed to be [redacted]. No administration or accommodation costs are associated with belumosudil as it is an oral preparation. On this basis, the total annual cost for belumosudil (acquisition, administration, and accommodation) is 13 x [redacted] = [redacted]. This compares favourably with the equivalent annual cost of ECP of £75,191 (Table 68). This simple side-by-side cost comparison does not account for the costs of inserting central lines, managing adverse events (including line infections) and patient transport costs. This cost comparison should also be

interpreted in the context of extreme capacity pressures for the NHS, and the opportunity costs associated with regular in hospital infusions.(11)

B.3.12. Subgroup analysis

Two subgroups were suggested for consideration in the scope but only if the evidence allows:

- Different organs or tissues affected by chronic GVHD
- Number and type of previous treatments

However, no subgroups have been considered in this submission for the following reasons:

- Section 5.2 of the SmPC states that: 'No clinically relevant differences in belumosudil pharmacokinetics were observed with regard to age, race, sex, weight or renal impairment (mild or moderate; severe renal impairment has not been studied)'.(1)
- The forest plots of pre-specified analyses for ORR are provided in Appendix E for the ROCKstar, Phase 2a and pooled analyses. These indicate that high ORRs were observed in all subgroups analysed in the mITT population comprising severity, duration of disease, number of organs at baseline and prior systemic treatments. Albeit in relatively small cohorts, these analyses suggest that efficacy was maintained irrespective of which subgroup was considered.
- Based on the above pharmacokinetics and clinical trial outcomes evidence, we do not expect belumosudil to be more clinically or cost effective than BAT in the subgroups suggested in the scope and therefore no subgroups have been considered in this submission.
- In addition, whilst the pooled evidence does provide overall a suitable number of patients for the cost-effectiveness analysis presented in this submission, constructing FFS and OS was considered not to be feasible for subgroups due to sample size from the belumosudil studies and availability of granular data for BAT making subgroup analyses difficult to perform robustly.

B.3.13. Benefits not captured in the QALY calculation

We consider belumosudil to be an innovative treatment, which was licensed under the Project Orbis programme(1) and granted an innovation passport by the MHRA in April 2021 (ILAP reference number ILAP/IP/21/53904/01). We believe there are benefits associated with belumosudil which are not captured by the QALY calculation. ECP, which constitutes a large proportion of the BAT comparator arm, is a high cost and resource-intensive therapy which can cause significant disruption to patients (Section B.1.3.2.2). Some important aspects of ECP administration were not included in the QALY calculation. These included the disruption and anxiety associated with public or hospital transport for patients and their caregivers attending regular outpatient appointments, lost workdays for caregivers due to taking time off to help the patient attend ECP appointments, the disutility associated

with insertion and removal of central lines where peripheral venous access was not possible, and the need for blood transfusions and anticoagulation therapy.(11) All of the above would be avoided through the use of an oral treatment such as belumosudil. Additionally, at a time of intense staffing and resourcing pressure for the NHS, the capacity benefit of an oral treatment versus regular, lengthy appointments for ECP, should not be underestimated.

Similarly, it is also worthwhile noting that a rapid COVID policy was implemented by NHSE during the pandemic for another treatment, ruxolitinib. This policy for an oral treatment that could be taken at home was intended to protect patients from COVID infection when presenting at hospital for ECP treatment. The policy has since been withdrawn and the ruxolitinib topic for both acute and chronic GVHD has been terminated in the NICE programme leaving patients with no licensed NICE-recommended oral treatment in this position.

B.3.14. Validation

The model validation process followed the current guidelines from the International Society for Pharmacoeconomics and Outcomes Research-Society of Medical Decision Making (ISPOR-SMDM).

Conceptual validity assesses if each model component (i.e., model concept/structure, data sources, problem formulation, and outcomes) reflects the underlying disease course, available evidence, and the clinical or administrative question at issue. The model concept was explained and discussed with a team of senior modellers from an external vendor. In addition, we discussed details on patient characteristics, clinical management, clinical outcomes, and resource use for the target population with English clinical and economic experts at an advisory board.(11) An additional clinical validation of the model inputs (i.e., extrapolation of survival outcomes) was also performed by our external vendor with a clinical KOL.

A technical review of the model was performed by an internal peer reviewer not involved in the original programming of the model. The review included:

- Detailed review of the formulas and sequence of calculations
- Checking the functionality of any built-in VBA macros and subroutines
- Extreme-value testing to identify and correct potential inconsistencies in model behaviour which could have been the result of programming or typing errors
- Checking the intermediate calculations for references (e.g., whether they are linked to correct cells) and implementation (e.g., whether correct signs for the parameters were used)
- Checking data inputs against references and sources
- Evaluation of the face validity of predicted results
- Following the validation, any identified errors were corrected before the model was finalised.

B.3.15. Interpretation and conclusions of economic evidence

Our cost-effectiveness analysis demonstrates the economic value of belumosudil as a treatment for patients with chronic GVHD who have received at least two prior lines of systemic therapy compared to current treatment options in England (i.e., BAT). In the base case, the probabilistic results indicated that belumosudil was associated with an incremental cost of [redacted] per patient, with incremental QALYs of [redacted] (without severity modifier) and [redacted] (with severity modifier [1.2 QALY weight]), resulting in probabilistic ICERs of £15,032 (without severity modifier) and £12,526 (with severity modifier) per QALY compared to BAT (Section B.3.10).

The benefit of belumosudil on health outcomes was driven by longer OS and FFS. Higher LYs and QALYs were estimated with belumosudil in the FF health state while in the failure state it was estimated to have lower LYs and QALYs, compared to BAT. Incremental costs with belumosudil were also mostly accumulated in the FF health state and were mainly attributed to increased treatment acquisition costs and disease management costs due to prolonged failure-free life years. Cost offsets were mainly driven by reduction in disease management costs in failure state. Results from one-way sensitivity analyses illustrated that the ICER was most sensitive to OS and FFS estimates, and disease management costs during failure state for patients receiving new systemic therapy (Section B.3.11.2). Results from the deterministic analyses (Section B.3.10) were similar to the probabilistic results, indicating that the results were robust to any uncertainty associated with the input parameters.

Results from scenario analyses indicated that changing assumptions around FFS and OS extrapolation, and time on treatment had a significant impact on the ICER (Section B.3.11.3):

- Sensitivity analyses showed that the ICER increased significantly (up to almost five times of the base case results), when different extrapolations for FFS were used (gamma and log-normal joint fits; and log-normal independent fits for belumosudil with gamma independent fits for BAT). However, the choice of the extrapolations used in the base case were based on the lowest AIC/BIC combinations, and the selected distribution was also validated by clinicians.
- Assuming that all treatments were given until failure (as opposed to modelling time on treatment based on a parametric TTD curve for belumosudil and median treatment durations reported in Zeiser et al. for BAT) increased the ICER by 105.4% (£30,985 per QALY gained with severity modifier [1.2 QALY weight]). However, this assumption is unlikely to be realistic for belumosudil given that observed data from the clinical trials indicated significantly shorter times on treatment compared to FFS (i.e., median treatment duration vs. median FFS: 9.2 vs. 13.7 months for belumosudil once daily; 11.2 vs. 15.1 months for belumosudil twice daily; 5.5 vs. 5.7 months for BAT).(58, 83)
- In sensitivity analyses that explored alternate distributions (log-normal, log-logistic, Weibull; and the use of independently fitted curves as opposed to jointly fitted curves) for OS extrapolation, belumosudil dominates BAT with incremental QALY gains (without severity

modifier) of [redacted] to [redacted] and incremental cost of [redacted] to [redacted]. However, we believe the choice of the extrapolations used in the base case is robust because they were based on the lowest AIC/BIC combinations. Unlike many oncology submissions where fits are based on very limited data, in most cases the data sets here are more than 50% complete providing confidence that the fit statistics can be used reliably. The fits were also chosen by the clinicians we consulted.

We have set the discount price of belumosudil at a responsible level to mitigate against the uncertainty inherent in the analysis of a rare disease with an evidence base derived from single-arm studies. It is critical to note that, of all the 60 sensitivity analyses carried out, only seven fall above £30,000/QALY without application of the severity modifier and four fall above this threshold with application of the severity modifier (Section B.3.11.3). All but one of these analyses are associated with choice of fit for the extrapolation of data and we have provided strong rationale above for the choices we have made, indicating that other fits are less plausible (even if some do take the ICER into the southeast quadrant where belumosudil dominates). The remaining high ICER is due to a very short time horizon which is less plausible given the observed OS data from the studies.

B.3.15.1. Strengths of the cost-effectiveness analysis

Our CEM was designed based on careful consideration of the clinical characteristics and treatment pathway of patients with chronic GVHD to ensure that key aspects of the disease and English treatment practices were captured in our analysis. The partitioned survival structure is well-suited to represent the progressive nature of chronic GVHD, with patients passing through a series of health states as their condition evolves, and with the impact of this clinical progression on patients' well-being and disease management being reflected via health state-specific utilities and costs. In addition to capturing differences in health and cost outcomes in terms of FFS – a clinically meaningful endpoint to clinicians, payers and patients (Section B.3.2.2.2) – the model structure allows further differences among failure-free patients to be captured depending on their level of response to treatment.

Additional strengths of our CEM include:

- The model was developed based on a thorough review of published economic modelling approaches, available data and consultation with clinical experts.(11) The model was designed to provide extensive flexibility to estimate clinical benefits of belumosudil. Clinical validation was carried out to substantiate the long-term survival predictions of the OS, FFS, and time in response projections.(11)
- The modelling approach, logical structure, expressions and sequences of calculations, and model inputs were validated by the team who conceptualised and implemented the model, and by a peer reviewer not initially involved with the model concept and programming.

B.3.15.2. Limitations of the cost-effectiveness analysis

Given the ultra-orphan nature of chronic GVHD and the inherent challenges with data availability for such a rare condition, there are limitations of our CEM that should be highlighted and discussed to properly interpret the results of the analysis:

- **Clinical inputs were based on naïve direct comparisons**

Comparison of belumosudil to BAT was based on a naïve direct comparison. This was necessary as an ITC to estimate the comparative efficacy of belumosudil against BAT was not feasible due to lack of a common comparator arm between the trials and key differences in the study populations, and due to methodological biases that could not be resolved in the ECA data (Section B.2.9.1). The pooled analysis of ROCKstar and the Phase 2a trial included patients who had 2 or more prior lines of systemic therapy, aligned with the MHRA indication for belumosudil, while the REACH-3 trial specifically excluded patients with 2 or more prior lines of therapy. Therefore, patients in REACH-3 likely had earlier stage, less severe chronic GVHD than patients in the belumosudil trials. As a result, use of naïve direct comparison data in our CEM likely favours the BAT comparator and, because of this, the results can be considered conservative.

- **Treatment tapering or discontinuation after a sustained treatment response was not explicitly modelled**

As per the ROCKstar study protocol, belumosudil can be tapered after a sustained response for 6 months and cessation of all other immunosuppressants for at least 3 months. Similarly, in the REACH-3 study, glucocorticoids and CNIs could be tapered after patients had a CR or PR. In addition, feedback from English clinical and economic experts consulted at an advisory board in January 2023 suggested that patients with sustained treatment response would eventually come off treatment and that 3 to 5 years should be considered as the maximum treatment duration for these patients. In our base case analysis, time on treatment was derived from the TTD curve and median treatment duration reported in the trials with an extrapolation of the curve over the long-term, with maximum treatment duration capped at 5 years for belumosudil, ECP, mycophenolate mofetil, imatinib and sirolimus. Thus, our base case may still slightly overestimate treatment costs among responders who were still on treatment in the longer term. A sensitivity analysis was performed to better reflect treatment tapering among responders by capping treatment duration at 3 years, and the ICER decreased to £5,446 with a 1.2 QALY weight (and to £6,536 without applying a severity modifier). Removing the treatment cap and assuming lifetime duration for treatments increased the ICER to £25,358 and to £30,430 with and without applying a severity modifier, respectively (Section B.3.11.3).

- **OS extrapolation was based on immature OS data**

OS data from the pooled ROCKstar and Phase 2a trials for belumosudil and from the REACH-3 study for BAT were immature, which created challenges in extrapolating long-term life expectancy (Section B.3.3.2). Clinical plausibility of the long-term predictions obtained from the parametric survival models

beyond the trial time horizon was assessed at the advisory board meeting by presenting the extrapolation curves along with predicted mean, median and % survival at key timepoints (e.g., 5 years, 10 years).(11) The exponential model for OS also proved to be the best fit with the lowest AIC and BIC fit statistics. Nevertheless, results from the DSA and scenario analyses revealed OS to be one of the main drivers of the ICERs (Section B.3.11). As discussed in Section B.3.3.2.3 we have recognised the uncertainty in the long-term extrapolation of OS and noted its impact on the results. Therefore, we have taken a responsible and conservative approach in the base case and adjusted the OS extrapolation in the analysis beyond 5 years to match the BAT risk thereafter. This five-year cut-off corresponds to the limit of the observed belumosudil data which has a maximum follow-up duration of 4.7 years and so the fitted curve to the data can be relied upon to be accurate up to this time point. (Section B.3.3.1.1).

- **Response definitions from REACH-3 and belumosudil trials were not compatible**

Response to treatment was incorporated in our economic evaluation by distributing patients in the FF state across various response levels, and cost and health outcomes were accrued accordingly. An important limitation to note is related to the variation in definition of response across the trials. Although both ROCKstar and REACH-3 build on the 2014 NIH Consensus Criteria definition(81), ROCKstar evaluated best response at any post-baseline assessment, while REACH-3 assessed best overall response at any time up to week 24. As a result, there is uncertainty regarding the comparability of response outcomes across the trials. Scenario analyses demonstrated that excluding response from the cost-effectiveness analysis had only minimal impact on the ICER (i.e., decreased by 1.28%; Section B.3.11.3). All response-related inputs were also included in the DSA and did not appear to be one of the top 20 most influential drivers of the ICER (Section B.3.11.2).

- **Disease management data were based on assumptions**

Data on long-term disease management costs based on failure states or response levels were not available from the clinical trials or published sources (Section B.3.5). As a result, with the exception of recurrent malignancy, disease management costs in the model were estimated based on our study examining the HES data.(45) It is very difficult to categorise patients within the HES database according to response status directly matching the modelled health states and so proxies must be used. In order to do this, we extracted several cohorts of patients with alloHSCT, according to GVHD status and use of high-cost therapy treatment. This does not capture differences directly aligned to the failure states or response levels used in the model but nevertheless parallels can be drawn. For example, it can be assumed that alloHSCT patients without GVHD correspond to complete responder patients in the model. This assumption has face validity because the observed costs reflect lower disease management costs for patients who responded to treatment than those with treatment failure. This was validated with clinical experts.

Various scenarios examining disease management costs were tested based on different assumptions of the cost categories taken from the HES study. The disease management cost of patients in different health states were among the most influential drivers of the ICER based on the DSA results

(Section B.3.11.2). Therefore, several assumptions for disease management costs were explored in deterministic sensitivity analyses resulting in changes to the ICER in both ways, ranging from -61.9% to 23.7% (Section B.3.11.3).

- **Lack of data to support cause of failure in the long-term**

Failure events among patients who were alive can be caused by a recurrent malignancy or a change to a new systemic chronic GVHD treatment. Among incident patients falling into the failure state in each cycle, a proportion was applied to determine the cause of failure. Certain assumptions were made for this input in the base case (Section B.3.3.1) due to a lack of data beyond the trial period. The different assumptions were explored in scenario analyses and appeared to have small impact on the ICER (changes ranged from 1% to 17%; Section B.3.11.3).

- **Lack of information on distribution of subsequent treatment**

Evidence and clinical guidelines on treatments for chronic GVHD in fourth-line therapy and beyond were inconclusive (Section B.1.3.2). The treatment pathway is highly patient/case-dependent, and patients also cycle through multiple therapies, but could also stop receiving treatment altogether. Therefore, in the base case a uniform distribution of subsequent treatments were assumed (except for rituximab), with an additional assumption that patients spend 60% of their remaining lifetime on treatment. This assumption was applied across all initial treatments in the model. Different sensitivity analyses were run to test these assumptions. Changing the distribution to 100% (i.e., assuming that patients remain on one of the treatment options for their rest of their life) decreased the ICER by 61.9%. An alternative was also tested, where each subsequent treatment is assumed to be received by patients for a specific duration and the corresponding subsequent treatment costs are applied as a one-off cost at the time of entering the Failure - new chronic GVHD systemic therapy health state. This approach increased the ICER to £23,734 with a severity modifier (QALY weight of 1.2) and to £28,481 without a severity modifier (Section B.3.11.3).

- **Clinical inputs for BAT were based on treatment composition in the REACH-3 trial**

The composition of treatments within the BAT basket was adjusted to be aligned with treatment practices in England (Section B.3.2.3) as it included treatments not available in England (e.g., ibrutinib). The adjustment reflected the higher utilisation of ECP in England compared to the proportion of patients receiving ECP in the BAT arm of REACH-3. The adjusted composition of treatments was aligned with feedback from NICE advisory board.(11) Given the lack of clinical data to inform the efficacy of the individual BAT treatments, efficacy was assumed to be similar across all BAT treatments used in REACH-3 and no adjustments were made to the clinical data as this was not feasible. This was accepted by experts consulted at our advisory board.

- **Indirect economic impact was not considered**

The GBMA indication for belumosudil includes patients aged 12 years and older. For patients aged less than 18 years, it is likely there will be additional economic burden on caregivers that is not captured in the presented model. This may be counted, in terms of indirect costs due to productivity losses due to caring responsibilities. Chronic GVHD is a very rare disease with a high economic and clinical burden per patient affecting a very limited number of patients in England.

Given this context for an ultra-orphan disease with limited treatment options available, the low base case ICER and sensitivity analyses which show that the uncertainty is mitigated by the PAS price (the large majority of SAs fall below the threshold with or without application of the severity modifier of 1.2), belumosudil represents an effective use of NHS resources.

We are hopeful that the committee will recognise the PAS price has been set for this submission at a responsible level to mitigate against the uncertainty inherent in the assessment of this ultra-orphan disease with a limited evidence base. It is critical that rapid access to belumosudil is gained for the very small population of patients suffering with chronic GVHD at third-line and beyond for whom there are no NICE-recommended licensed options at this point in the pathway.

As a clinician told us:

> ***"These patients have wide ranging and profound medical, psychological and social needs. They use an amount of resource that is undefined, and certainly grossly under-appreciated... [there are also] indirect losses to society with the morbidity these patients suffer and the care they require... this is a multi-system disease that results in pan-system effects."***

References

1. Medicines & Healthcare products Regulatory Agency (MHRA). Rezurock Summary of Product Characteristics 2022 [Available from: https://mhraproducts4853.blob.core.windows.net/docs/01da0fb17239cd9bc996b6a30821bc09e588ca75.
2. Zanin-Zhorov A, Blazar BR. ROCK2, a critical regulator of immune modulation and fibrosis has emerged as a therapeutic target in chronic graft-versus-host disease. Clin Immunol. 2021;230:108823.
3. Medicines & Healthcare products Regulatory Agency (MHRA). Orphan Register 2022 [Available from: https://www.gov.uk/government/publications/orphan-registered-medicinal-products/orphan-register.
4. Medicines & Healthcare products Regulatory Agency (MHRA). Guidance - Project Orbis 2022 [Available from: https://www.gov.uk/guidance/guidance-on-project-orbis.
5. MacDonald KP, Hill GR, Blazar BR. Chronic graft-versus-host disease: biological insights from preclinical and clinical studies. Blood, The Journal of the American Society of Hematology. 2017;129(1):13-21.
6. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2015;21(3):389-401.e1.
7. Mawardi H, Hashmi SK, Elad S, Aljurf M, Treister N. Chronic graft-versus-host disease: Current management paradigm and future perspectives. Oral Dis. 2019;25(4):931-48.
8. Finazzi MC, Boschini C, Craddock C, Rambaldi A, Ward J, Malladi RK. Characteristics of graft-versus-host disease occurring after alemtuzumab-containing allogeneic stem cell transplants: incidence, organ involvement, risk factors and survival. British journal of haematology. 2020;188(4):550-9.
9. British Society of Blood and Marrow Transplantation and Cellular Therapy (BSBMTCT). Annual activity 2020 2020 [Available from: https://bsbmtct.org/activity/2020/.
10. Cooke KR, Luznik L, Sarantopoulos S, Hakim FT, Jagasia M, Fowler DH, et al. The Biology of Chronic Graft-versus-Host Disease: A Task Force Report from the National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2017;23(2):211-34.
11. Sanofi. Data on File. NICE Belumosudil Virtual Advisory Board Report. 2023.
12. British Society of Blood and Marrow Transplantation and Cellular Therapy (BSBMTCT). The outcome of haematopoietic stem cell transplantation: an analysis of registry data for UK transplants performed 2009-2014 inclusive 2017 [Available from: https://www.uhs.nhs.uk/Media/SUHTExtranet/Services/WessexBloodAndMarrowTransplantationService/BSBMT-8th-Report-To-Commissioners.pdf.
13. AnthonyNolan.org. GvHD of the eye: patient experiences and a potential new treatment 2021 [Available from: https://www.anthonynolan.org/blog/2021/09/23/gvhd-eye-patient-experiences-and-a-potential-new-treatment.
14. Salhotra A, Eiznhamer D, Hennegan K, Dehipawala S, Aggarwal SK. PRESENTATION AND MANAGEMENT OF CHRONIC GRAFT-VERSUS-HOST DISEASE IN REAL-WORLD CLINICAL PRACTICE: A MEDICAL CHART AUDIT. EHA Virtual Congress2020.
15. Arora M, Cutler CS, Jagasia MH, Pidala J, Chai X, Martin PJ, et al. Late Acute and Chronic Graft-versus-Host Disease after Allogeneic Hematopoietic Cell Transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2016;22(3):449-55.
16. Sanofi. Data on File. Patient Video Transcript - QOL Impact. 2022.
17. Styczyński J, Tridello G, Koster L, Iacobelli S, van Biezen A, van der Werf S, et al. Death after hematopoietic stem cell transplantation: changes over calendar year time, infections and associated factors. Bone Marrow Transplant. 2020;55(1):126-36.
18. Agh T, Csanadi M, Voko Z, Webb T, Jeyakumaran D, Trudeau J, et al. Humanistic burden of patients with chronic graft-versus-host disease - systematic literature review of health-related quality of life and functional status. Expert review of hematology. 2019;12(5):295-309.

19. Lee SJ, Onstad L, Chow EJ, Shaw BE, Jim HSL, Syrjala KL, et al. Patient-reported outcomes and health status associated with chronic graft-versus-host disease. Haematologica. 2018;103(9):1535-41.
20. de Vere Hunt I, Kilgour JM, Danby R, Peniket A, Matin RN. “Is this the GVHD?” A qualitative exploration of quality of life issues in individuals with graft-versus-host disease following allogeneic stem cell transplant and their experiences of a specialist multidisciplinary bone marrow transplant service. Health and Quality of Life Outcomes. 2021;19(1):11.
21. Sanofi. Data on File. Patient Video Transcript - Psychological Impact. 2022.
22. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2005;11(12):945-56.
23. Pidala J, Kurland B, Chai X, Majhail N, Weisdorf DJ, Pavletic S, et al. Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. Blood. 2011;117(17):4651-7.
24. Sanofi. Data on File. Patient Video Transcript - Bringing Clarity. 2022.
25. Sanofi. SLR Technical Report: Health Related Quality of Life (HRQoL) and Health Utility Values in Patients with Chronic Graft-Versus-Host Disease. (data on file). 2022.
26. Sanofi. RZR Patient Video Clips (data on file). 2022.
27. National Institute for Health and Care Excellence (NICE). NICE Scientific Advice: Advice Report August 2022. (data on file). 2022.
28. Sanofi. GVHD Disease Specific Programmes (DSP) Analysis Slide Deck. (data on file). 2022.
29. Anderson P, Benford M, Harris N, Karavali M, Piercy J. Real-world physician and patient behaviour across countries: Disease-Specific Programmes - a means to understand. Current medical research and opinion. 2008;24(11):3063-72.
30. Yu JT, J.; Stewart, SK.; Wartman, L.; Williams, P.; Amoloja, T.; et al., editors. . Impact of Chronic Graft-Versus-Host Disease on Patients’ Quality of Life, Employment, Productivity, Income, and Caregivers. Abstract presented at 2021 TCT Transplantation & Cellular Therapy. Meetings of ASTCT and CIBMTR TCT Meetings; February 8, 20212021.
31. Sanofi. GVHD Disease Specific Programmes (DSP) Analysis Tables. (data on file). 2022.
32. Matza LS, Paramore LC, Stewart KD, Karn H, Jobanputra M, Dietz AC. Health state utilities associated with treatment for transfusion-dependent β-thalassemia. The European Journal of Health Economics. 2020;21:397-407.
33. Sanofi. Data on File. Utilities in chronic GVHD Draft Report - Estimating the quality of life burden and impact of chronic graft versus host disease. 2023.
34. Ubel PA, Peeters Y, Smith D. Abandoning the language of “response shift”: a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. Quality of Life Research. 2010;19:465-71.
35. Jansen SJ, Stiggelbout AM, Wakker PP, Nooij MA, Noordijk EM, Kievit J. Unstable preferences: a shift in valuation or an effect of the elicitation procedure? Medical Decision Making. 2000;20(1):62-71.
36. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. Quality of life Research. 2003;12:599-607.
37. Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. The economic journal. 2008;118(525):215-34.
38. Ludwig K, Ramos-Goñi JM, Oppe M, Kreimeier S, Greiner W. To what extent do patient preferences differ from general population preferences? Value in Health. 2021;24(9):1343-9.
39. Goodwin E, Green C, Hawton A. What difference does it make? A comparison of health state preferences elicited from the general population and from people with multiple sclerosis. Value in Health. 2020;23(2):242-50.
40. Jamani K, Onstad LE, Bar M, Carpenter PA, Krakow EF, Salit RB, et al. Quality of life of caregivers of hematopoietic cell transplant recipients. Biology of Blood and Marrow Transplantation. 2018;24(11):2271-6.
41. Acaster S, Perard R, Chauhan D, Lloyd AJ. A forgotten aspect of the NICE reference case: an observational study of the health related quality of life impact on caregivers of people with multiple sclerosis. BMC Health Serv Res. 2013;13:346.

42. Liu D, Ahmet A, Ward L, Krishnamoorthy P, Mandelcorn ED, Leigh R, et al. A practical guide to the monitoring and management of the complications of systemic corticosteroid therapy. Allergy Asthma Clin Immunol. 2013;9(1):30.
43. Retzler J, Smith A, Reaney M, Rout R, Hudson R. Process utilities for topical treatment in atopic dermatitis. Qual Life Res. 2019;28(9):2373-81.
44. Lachance S, Hamad N, de Courcy J, Gibson G, Zuurman M, Mohty M, editors. Impact of Chronic GVHD Severity And Steroid Response on The Quality of Life In Patients Following Allogeneic Stem Cell Transplantation: Findings From A Real-World Study. BONE MARROW TRANSPLANTATION; 2021: SPRINGERNATURE CAMPUS, 4 CRINAN ST, LONDON, N1 9XW, ENGLAND.
45. Sanofi. Data on File. HES Report - The clinical and economic burden associated with chronic graft-versus-host disease following allogeneic haematopoietic stem cell transplant in England; Version No.: 01.00. 2022.
46. Sanofi. Data on File. HES Excel Report v.01.02. 2022.
47. Dignan FL, Amrolia P, Clark A, Cornish J, Jackson G, Mahendra P, et al. Diagnosis and management of chronic graft-versus-host disease. British journal of haematology. 2012;158(1):46-61.
48. NHS England. Clinical Commissioning Policy: Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation 2017 [Available from: https://www.england.nhs.uk/wp-content/uploads/2017/03/gvhd-heamatopoietic-stem-cell.pdf.
49. Martin PJ, Lee SJ, Przepiorka D, Horowitz MM, Koreth J, Vogelsang GB, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. The 2014 Clinical Trial Design Working Group Report. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2015;21(8):1343-59.
50. Kadmon Pharmaceuticals. Clinical Study Protocol: A Phase 2, Randomized, Multicenter Study to Evaluate the Efficacy and Safety of KD025 in Subjects with Chronic Graft Versus Host Disease (cGVHD) After At Least 2 Prior Lines of Systemic Therapy (The ROCKstar Study); Amendment 1. 2019 26-June-2019.
51. Alfred A, Taylor PC, Dignan F, El-Ghariani K, Griffin J, Gennery AR, et al. The role of extracorporeal photopheresis in the management of cutaneous T-cell lymphoma, graft-versus-host disease and organ transplant rejection: a consensus statement update from the UK Photopheresis Society. British journal of haematology. 2017;177(2):287-310.
52. Button PM, Child F, Mootien M, Saglam S, Magsino J, Dhanoia H. Service improvement in a nurse-led clinic for extracorporeal photopheresis. British Journal of Nursing. 2021;30(17):S10-S4.
53. Roche. Rituximab (MabThera) Summary of Product Characteristics. 2022.
54. Novartis Pharmaceuticals UK Ltd. Imatinib (Glivec 100 mg film-coated tablets) Summary of Product Characteristics. 2022.
55. Pfizer. Sirolimus (Rapamune) Summary of Product Characteristics. 2022.
56. Roche. Mycopholate Mofetil (CellCept 1g/5ml powder for oral suspension) Summary of Product Characteristics. 2022.
57. Novartis Europharm Limited. Novartis receives European Commission approval for Jakavi® to be the first post-steroid treatment for acute and chronic graft-versus-host disease 2022 [Available from: https://www.novartis.com/news/media-releases/novartis-receives-european-commission-approval-jakavi-be-first-post-steroid-treatment-acute-and-chronic-graft-versus-host-disease.
58. Zeiser R, Polverelli N, Ram R, Hashmi SK, Chakraverty R, Middeke JM, et al. Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. N Engl J Med. 2021;385(3):228-38.
59. NHS England. Ruxolitinib for chronic Graft versus Host Disease (RPS 2010) 2021 [updated 30-Jul-2021. Available from: https://www.england.nhs.uk/coronavirus/documents/c1365-ruxolitinib-for-chronic-graft-versus-host-disease-rps-2010/.
60. AnthonyNolan.org. Graft versus host disease (GvHD) 2021 [Available from: https://www.anthonynolan.org/patients-and-families/recovery-body/graft-versus-host-disease-gvhd#what-is-extracorporeal-photopheresis-(ECP)?
61. Malik MI, Litzow M, Hogan W, Patnaik M, Murad MH, Prokop LJ, et al. Extracorporeal photopheresis for chronic graft-versus-host disease: a systematic review and meta-analysis. Blood Res. 2014;49(2):100-6.

62. NHS Blood and Transplant (NHSBT). TAS units 2022 [Available from: https://hospital.blood.co.uk/patient-services/therapeutic-apheresis-services/tas-units/.
63. Matza LS, Cong Z, Chung K, Stopeck A, Tonkin K, Brown J, et al. Utilities associated with subcutaneous injections and intravenous infusions for treatment of patients with bone metastases. Patient preference and adherence. 2013:855-65.
64. National Institute for Health and Care Excellence (NICE). Highly Specialised Technology Evaluation: Migalastat for treating Fabry disease [ID 868] - Evaluation Report 2016 [Available from: https://www.nice.org.uk/guidance/hst4/documents/committee-papers.
65. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care 2021 2021 [Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-of-health-and-social-care-2021/.
66. National Institute for Health and Care Excellence (NICE). Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies - Technology appraisal guidance [TA559]. 2019.
67. Macmillan Cancer Support. Rituximab 2021 [Available from: https://www.macmillan.org.uk/cancer-information-and-support/treatments-and-drugs/rituximab.
68. Wong FL, Francisco L, Togawa K, Bosworth A, Gonzales M, Hanby C, et al. Long-term recovery after hematopoietic cell transplantation: predictors of quality-of-life concerns. Blood, The Journal of the American Society of Hematology. 2010;115(12):2508-19.
69. Bachier CR, Aggarwal SK, Hennegan K, Milgroom A, Francis K, Dehipawala S, et al. Epidemiology and Treatment of Chronic Graft-versus-Host Disease Post-Allogeneic Hematopoietic Cell Transplantation: A US Claims Analysis. Transplantation and cellular therapy. 2021;27(6):504.e1-.e6.
70. Penack O, Marchetti M, Ruutu T, Aljurf M, Bacigalupo A, Bonifazi F, et al. Prophylaxis and management of graft versus host disease after stem-cell transplantation for haematological malignancies: updated consensus recommendations of the European Society for Blood and Marrow Transplantation. Lancet Haematol. 2020;7(2):e157-e67.
71. Anthony Nolan.org. Facts and Stats. https://www.anthonynolan.org/front/facts-and-stats Last accessed 20 Dec 2022. 2022 [
72. Smith ZI, Rosenstein RK, Banerjee S, Pichard DC, Pavletic SZ, Cowen EW. Quality of Life in Patients With Skin of Color and Chronic Graft-vs-Host Disease. JAMA Dermatol. 2020;156(5):589-90.
73. Cutler C, Lee SJ, Arai S, Rotta M, Zoghi B, Lazaryan A, et al. Belumosudil for chronic graft-versus-host disease after 2 or more prior lines of therapy: the ROCKstar Study. Blood. 2021;138(22):2278-89.
74. Jagasia M, Lazaryan A, Bachier CR, Salhotra A, Weisdorf DJ, Zoghi B, et al. ROCK2 Inhibition With Belumosudil (KD025) for the Treatment of Chronic Graft-Versus-Host Disease. Journal of Clinical Oncology. 2021;39(17):1888-98.
75. Jagasia M, Lazaryan A, Bachier C, Salhotra A. Supplemental Information: Clinical Study Protocol Redaction - A Phase 2a, Dose-Escalation, Open-Label Study to Evaluate the Safety, Tolerability, and Activity of KD025 in Subjects with Chronic Graft Versus Host Disease. American Society of Clinical Oncology. 2021.
76. Kadmon Pharmaceuticals L. Data on File (2.5 Clinical Overview).
77. Kadmon Pharmaceuticals. CLINICAL STUDY REPORT A Phase 2, Randomized, Multicenter Study to Evaluate the Efficacy and Safety of KD025 in Subjects With Chronic Graft Versus Host Disease (cGVHD) After at Least 2 Prior Lines of Systemic Therapy (The ROCKstar Study). 2020.
78. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of Epidemiology & Community Health. 1998;52(6):377-84.
79. Hooper P, Jutai JW, Strong G, Russell-Minda E. Age-related macular degeneration and low-vision rehabilitation: a systematic review. Can J Ophthalmol. 2008;43(2):180-7.
80. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Hematopoietic Cell Transplantation. 2022;1.2022.
81. Lee SJ, Wolff D, Kitko C, Koreth J, Inamoto Y, Jagasia M, et al. Measuring therapeutic response in chronic graft-versus-host disease. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: IV. The 2014 Response Criteria Working Group report. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation. 2015;21(6):984-99.

82. Jacobsohn DA. FFS: an end(point) to our problems in chronic GVHD trials? Blood. 2014;124(8):1216-7.
83. Kadmon Pharmaceuticals. KD025-213 Data Version: (Cutoff Date August 19, 2021). 2022.
84. Kadmon International Ltd. Data on File. KD025-213_two year cut_figures. 2022.
85. Kadmon Pharmaceuticals. Clinical Study Reports: 60-Day Efficacy Update Report Belumosudil in cGVHD. 2022.
86. Revicki DA, Kawata AK, Harnam N, Chen WH, Hays RD, Cella D. Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. Qual Life Res. 2009;18(6):783-91.
87. Thompson NR, Lapin BR, Katzan IL. Mapping PROMIS Global Health Items to EuroQol (EQ-5D) Utility Scores Using Linear and Equipercentile Equating. Pharmacoeconomics. 2017;35(11):1167-76.
88. Kadmon Pharmaceuticals. KD025-ISE Final Run (cutoff date 19Aug2021). 2022.
89. Kadmon Pharmaceuticals. Overall response rate (mITT population who had prior ECP, phase 2 analysis group). 2022.
90. Lee SJ, Cutler C, Blazar BR, Tu A, Yang Z, Pavletic SZ. Correlation of Patient-Reported Outcomes with Clinical Organ Responses: Data from the Belumosudil Chronic Graft-versus-Host Disease Studies. Transplantation and cellular therapy. 2022.
91. Kadmon Pharmaceuticals. Clinical Study Protocol: A Phase 2a, Dose-Escalation, Open-Label Study to Evaluate the Safety, Tolerability, and Activity of KD025 in Subjects with Chronic Graft Versus Host Disease. 2016 22-Feb-2016.
92. Sanofi. Data on File. RWE STUDY SUPPORT MATERIALS. Comparative Effectiveness of Belumosudil in Phase 2 Trial Programme vs. an External Control Arm of Physician's Choice of Best Available Therapy for Patients with Chronic Graft vs. Host Disease. RWE Observational Study SAP 2022.
93. Cutler CL, SJ; Pavletic, St.;. Belumosudil for Chronic Graft-Versus-Host-Disease [cGVHD] After 2 or More Prior Lines of Systemic Therapy [LOTs]: 2-year Safety Results from the Pivotal Phase 2 ROCKstar Study (KD025-213). 2022.
94. Kadmon Pharmaceuticals. KD025-ISS Final Run (cutoff date 19Aug2021). 2022.
95. Kadmon Pharmaceuticals. Data on File. KD025_Grade 3 or higher AEs (cutoff date 19Aug2021). 2022.
96. NHS Blood and Transplant (NHSBT). Therapeutic Apheresis Treatments 2022 [Available from: https://www.nhsbt.nhs.uk/what-we-do/diagnostic-and-therapeutic-services/therapeutic-apheresis/therapeutic-apheresis-treatments/.
97. Flynn R, Paz K, Du J, Reichenbach DK, Taylor PA, Panoskaltsis-Mortari A, et al. Targeted Rho-associated kinase 2 inhibition suppresses murine and human chronic GVHD through a Stat3-dependent mechanism. Blood. 2016;127(17):2144-54.
98. Crespo C, Pérez-Simón JA, Rodríguez JM, Sierra J, Brosa M. Development of a population-based cost-effectiveness model of chronic graft-versus-host disease in Spain. Clin Ther. 2012;34(8):1774-87.
99. CADTH. CADTH reimbursement review: Ruxolitinib (Jakavi). . Canadian Journal of Health Technologies. 2022;10:1-220.
100. HAS. Jakavi (ruxolitinib) - Acute (GvHa) and chronic (GvHc) graft versus host disease when there is an inadequate response to corticosteroids or other therapies. Assess HEALTH TECHNOLOGIES. 2022.
101. Okumura L, Holanda P, Riveros B, Lucchetta R, Rosim M, Dierick K, et al. PRO4 COST-EFFECTIVENESS ANALYSIS OF AN INTEGRATED EXTRACORPOREAL PHOTOPHERESIS SYSTEM VERSUS SELECTED TREATMENTS FOR REFRACTORY CHRONIC GRAFT VERSUS HOST DISEASE. Value in Health Regional Issues. 2019;19:S74.
102. Yalniz FF, Murad MH, Lee SJ, Pavletic SZ, Khera N, Shah ND, et al. Steroid Refractory Chronic Graft-Versus-Host Disease: Cost-Effectiveness Analysis. Biology of Blood and Marrow Transplantation. 2018;24(9):1920-7.
103. Peacock A DF, Taylor C, et al. . P615 Cost-effectiveness of extracorporeal photopheresis for the treatment of chronic graft versus host disease in the australian setting. The 48th Annual Meeting of the European Society for Blood and Marrow Transplantation: Patient Advocacy – Poster Session (P613-P616). Bone Marrow Transplantation. 2022;57(1):468-70.
104. De Waure C, Capri S, Veneziano MA, Specchia ML, Cadeddu C, Di Nardo F, et al. Extracorporeal photopheresis for second-line treatment of chronic graft-versus-host

diseases: results from a health technology assessment in Italy. Value in Health. 2015;18(4):457-66.
105. National Institute of Health and Care Excellence. NICE DSU technical support document 19: partitioned survival analysis as a decision modelling tool in health care: a critical review. 2017. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf. (Accesssed 22 April 2020).
106. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9.
107. Ishak KJ, Kreif N, Benedict A, Muszbek N. Overview of parametric survival analysis for health-economic applications. Pharmacoeconomics. 2013;31(8):663-75.
108. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.; 2011.
109. Sanofi. Data on File. KD025-ISE Final Run (cutoff date 19Aug2021). Additional Analyses for FFS, OS, QDOR, TTD and TR. 2023.
110. Le RQ, Wang X, Zhang H, Li H, Przepiorka D, Vallejo J, et al. FDA Approval Summary: Ruxolitinib for Treatment of Chronic Graft-Versus-Host Disease after Failure of One or Two Lines of Systemic Therapy. The Oncologist. 2022;27(6):493-500.
111. Cutler C, Miklos D, Kim HT, Treister N, Woo SB, Bienfang D, et al. Rituximab for steroid-refractory chronic graft-versus-host disease. Blood. 2006;108(2):756-62.
112. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia [ID1484]. Committee Papers. TA642. 2020.
113. Joshi N, Hensen M, Patel S, Xu W, Lasch K, Stolk E. Health state utilities for acute myeloid leukaemia: a time trade-off study. Pharmacoeconomics. 2019;37(1):85-92.
114. Aristides M, Barlev A, Barber B, Gijsen M, Quinn C. Population preference values for health states in relapsed or refractory B-precursor acute lymphoblastic leukemia in the United Kingdom. Health and quality of life outcomes. 2015;13(1):1-7.
115. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal. Asciminib for treating chronic myeloid leukaemia after 2 or more tyrosine kinase inhibitors [ID3813]. Committee Papers. TA813. 2022.
116. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal. Ponatinib for treating chronic myeloid leukaemia and acute lymphoblastic leukaemia. Committee Papers. TA451. 2017.
117. Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. Health and quality of life outcomes. 2010;8(1):1-9.
118. National Institute for Health Care and Excellence (NICE). Single technology appraisal. Idelalisib for treating chronic lymphocytic leukaemia [ID764]. TA359. 2015.
119. National Institute for Health Care and Excellence (NICE). Single Technology Appraisal. Acalabrutinib for untreated and treated chronic lymphocytic leukaemia [ID1613]. Committee Papers. TA689. 2020.
120. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. 2022.
121. NICE Decision Support Unit. Estimating EQ-5D by age and sex for the UK - Expected EQ-5D-3L by age and sex and covariance matrices of the models using the HSE 2014 dataset (XLS, 58KB) 2022 [Available from: https://www.sheffield.ac.uk/nice-dsu/methods-development/estimating-eq-5d.
122. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. PharmacoEconomics. 2022:1-9.
123. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT) 2021 [Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
124. Royal Pharmaceutical Society. British National Formulary (BNF) by Medicines Complete 2023 [Available from: https://about.medicinescomplete.com/publication/british-national-formulary/.
125. NHS Business Services Authority (NHSBSA). Electronic Drug Tariff Database. 2022.

126. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal. Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374] Committee Papers TA627. 2020.
127. NHS England. National Schedule of NHS Costs 2020/1 2022 [Available from: https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/.
128. Schain F, Batyrbekova N, Liwing J, Baculea S, Webb T, Remberger M, et al. Real-world study of direct medical and indirect costs and time spent in healthcare in patients with chronic graft versus host disease. The European journal of health economics : HEPAC : health economics in prevention and care. 2021;22(1):169-80.
129. Manoukian S, Stewart S, Graves N, Mason H, Robertson C, Kennedy S, et al. Bed-days and costs associated with the inpatient burden of healthcare-associated infection in the UK. Journal of Hospital Infection. 2021;114:43-50.
130. Office for National Statistics (ONS). National life tables: UK 2021 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables. .

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Summary of Information for Patients (SIP)

April 2023

File name	Version	Contains confidential information	Date
ID4021_Belumosudil_SIP_20042023	2.0	No	06.04.2023

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Active ingredient: Belumosudil **Brand name:** REZUROCK®

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Patients aged 12 years and older with chronic graft-versus-host disease (GVHD) who have received at least two prior lines of systemic therapy.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Marketing authorisation (MA) for belumosudil was received in July 2022. Belumosudil is indicated for the treatment of patients aged 12 years and older with chronic GVHD who have received at least two prior lines of systemic therapy. The link for the approved Summary of Product Characteristics (SmPC) can be found here. The Medicines and Healthcare products Regulatory Agency (MHRA) have granted an Innovation Passport to belumosudil through the Innovative Licensing and Access Pathway (ILAP). ILAP is a regulatory process designed to speed up access to new, innovative medicines for patients. Belumosudil was granted orphan drug designation by the MHRA and the European Commission. Orphan drug status is only given to medicines for rare, serious conditions where current treatment options are limited. On this basis, belumosudil has been validated under the ultra orphan process for the Scottish Medicines Consortium (SMC).

Belumosudil is currently licensed in a number of countries, including the USA, Canada and Australia.

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

- Sanofi has made a financial contribution of £10,000 to the Cancer52 Corporate Supporters Programme in 2023.
- Sanofi global CSU have commissioned Rare Disease Research Partners (MPS Commercial) to support development of clinical trials materials for patients. This was to the value of £4,860
- Sanofi provided a financial donation of £24,490 as a pharmaceutical partner of Myeloma UK's London to Paris Bike Ride 2022.

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

What is the main condition that the medicine is planned to treat?
Rezurock (belumosudil) is indicated for the treatment of patients aged 12 years and older with chronic graft-versus-host disease (chronic GVHD) who have received at least two prior lines of systemic therapy.

What is chronic GVHD?
Stem cell transplants are an important treatment option for people with a range of different blood cancers and other diseases which affect the bone marrow. Allogeneic haematopoietic stem cell transplantation (AlloHSCT) is where healthy stem cells are transferred from a donor to treat the recipient patient's disease (21).

GVHD is one of the possible complications of AlloHSCT (21). This is where the donor stem cells recognise the recipient's cells as foreign and begin to attack them. GVHD is classified as either acute or chronic (6). This document only focuses on the chronic form of GVHD, which can affect a wide range of organs and may persist for several months or years.

How many people get chronic GVHD?
In 2019 there were 1,506 allogeneic (donor stem cell) transplants in England. For cancerous indications, approximately 33% of adult and 16% of paediatric donor stem cell transplant recipients developed chronic GVHD. For non- cancerous indications the corresponding figure was 23% in adult and 12% in paediatric stem cell transplant recipients. Between 5 and 11% of stem cell transplant recipients are expected to develop extensive chronic GVHD that may require second or subsequent lines of therapy. Between 2016 and 2020, 713 people were diagnosed with extensive chronic GVHD in England (23).

People who receive a transplant from a mismatched unrelated donor are more at risk of developing chronic GVHD. People from a non-white family background are less likely to find a related donor match which results in people from ethnic minority family background being at increased risk of chronic GVHD. (23)

Symptoms of chronic GVHD:
Chronic GVHD causes inflammation (swelling) and fibrosis (scarring or hardening) which can affect almost any organ in the body. The skin, mouth, and eyes are the most common organs involved, with 86% of patients showing involvement in at least one of these organs at diagnosis. This also includes further symptoms linked to fibrosis, such as fasciitis (inflammation of connective tissues), cutaneous sclerosis (hardening of the skin), and bronchiolitis obliterans (inflammation of the lung's airways) (12) (26).

Table 1: Some Clinical Presentations of chronic GVHD based on 2014 National Institute of Health (NIH) Consensus Criteria (13):

Lung	Shortness of breath, chronic coughing or changes seen on the chest X-ray (symptoms of bronchiolitis obliterans and bronchiectasis)
Nails	Nail loss, ridging, brittleness or splitting
Scalp and body hair	Lesions, scarring and alopecia (after recovery from chemoradiotherapy)
Genitalia	Scarring, lesions, ulcers, and conditions such as clitoral/labial adhesion (known as agglutination) or tightening of the foreskin (phimosis)
Eyes	Dry eyes or vision changes
Skin	Rash, lesions, raised or discoloured areas
Gastro-Intestinal Tract	Pain or difficulty swallowing, weight loss
Muscles, Joints and Fascia (connective tissue)	Muscle cramps, joint stiffness or contractures or pain (symptoms associated to fasciitis)
Mouth	Dry mouth, white patches, ulcers and pain or sensitivity

How does chronic GVHD affect the quality of life of patients and their families/caregivers?
By the time patients undergo stem cell transplantation, they will most likely have already undergone intensive clinical treatment and interventions and have had to deal with the trauma of being diagnosed with a life-threatening condition, such as blood cancer and certain blood, immune system and metabolic disorders. Then, they develop chronic GVHD as a result of their stem cell transplant, which they received to treat their initial illness. Some patients with chronic GVHD experience significant depression or anxiety symptoms (11)(25). In patient interviews conducted in the US, a patient stated: "*If they told me, 'By the way, after this, you're going to experience another nightmare,' I don't know if I would have wanted to know. I don't know if I would have had the strength to fight the first battle if I knew another battle was right behind it*" (30)

Chronic GVHD can cause severe disabilities such as joint contractures (deformities caused by joint tightening), loss of sight, and end-stage lung disease which can severely limit activities of daily living. The immune system can be suppressed, leading to repeated and life-threatening infections, which can limit patients' ability to go out in public, socialise or perform normal tasks (12).

Some studies have shown that patients with chronic GVHD are less likely to return to work within a two-year period following their stem cell transplant than those without chronic GVHD (37) (38).

Patients with chronic GVHD are usually prescribed corticosteroids at some point. These medicines can cause a range of undesirable side effects, such as osteoporosis and fractures, heart disease, reduced immunity, disorders of the skeletal muscles (known as myopathy), weight gain, cataracts and glaucoma, high blood sugar and diabetes, psychiatric disturbances, as well as other gastrointestinal (digestive organs) and dermatological (skin) effects (19).

GVHD is a second most common cause of death (after the cancer itself returning) and leading cause of severe illness in patients who undergo stem cell transplantation, often as a result of organ failure or infection (8) (34). In the UK, between 2009-2014, GVHD was the cause of 26.9% of non-cancer deaths in patients who received a donor stem cell transplant (4).

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

There are two distinct forms of GVHD: acute and chronic GVHD. Diagnosis of chronic GVHD is complex, because it can affect so many different organs and impact people in different ways. Chronic GVHD and acute GVHD can be present at the same time (sometimes referred to as "overlap GVHD") (13).

The chronic GVHD international diagnosis guidelines were last updated in 2014. Diagnosis of chronic GVHD can involve biopsy, laboratory tests, or radiography. There must also be at least one specific diagnostic clinical feature, which can be identified on the skin, mouth, genitalia, lungs, gastro-intestinal (digestive) tract, or musculoskeletal system. However, other organs may be affected. These diagnostic clinical features include distinctive plaques or lesions (tissue damage), or bronchiolitis obliterans syndrome (a rare condition involving inflammation of the airways) (13). Patients may be reviewed by a specialist such as an ophthalmologist or gynaecologist.

No additional diagnostic tests are required with the new treatment.

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

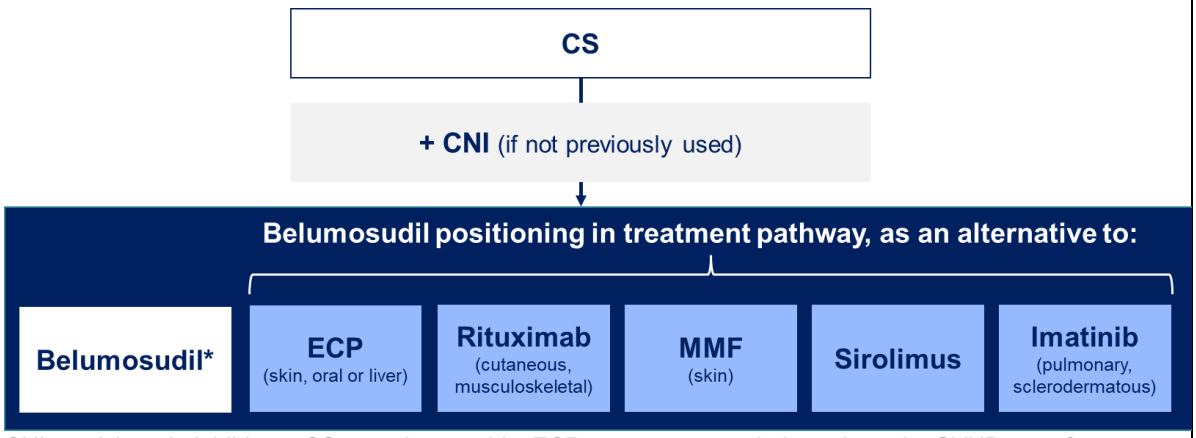
For patients with mild chronic GVHD, symptoms can sometimes be managed using topical (skin) treatments such as creams and ointments. However, most patients with chronic GVHD are prescribed oral corticosteroids (sometimes with calcineurin inhibitors such as cyclosporin) first. These act by suppressing the immune system response. However, corticosteroids can also cause a range of undesirable side effects. They are often not effective enough to resolve the disease symptoms (10).

Several other treatments are available for doctors to prescribe patients with chronic GVHD if corticosteroids are not sufficient (for example, if they are not controlling disease or are causing too many side effects). These include extracorporeal photopheresis (ECP), sirolimus, imatinib, mycophenolate mofetil and rituximab. Treatment guidelines were last published by the British Society for Bone Marrow Transplantation and Cellular Therapies (BSBMTCT) and the British Committee for Standards in Haematology (BCSH) in 2012. The guidelines note that there is not much evidence to support the use of many of the other available treatment options listed above (10).

According to UK clinicians contacted by Sanofi, one of the most commonly prescribed treatments (after corticosteroids, with or without calcineurin inhibitors) is ECP (31).

ECP is a procedure which involves collecting blood from the patient and destroying the white blood cells which cause GVHD by exposing them to ultraviolet light and a medication called methoxypsoralen (8-MOP) (2). The treated blood is immediately returned to the patient's bloodstream. ECP appointments can take 2 hours at a time and patients may need to attend multiple times, with two consecutive days of treatment per fortnight. This may include an overnight stay. Treatment can last 6 months before there is any improvement, but in some cases can last 12-18 months (22). Not all treatment centres have access to ECP nearby so it could be an inconvenient option for patients, particularly those living in remote areas or those who cannot regularly take time out of their day.

Figure 1. Belumosudil expected place in treatment pathway



CNI = calcineurin inhibitors; CS = corticosteroids; ECP = extracorporeal photopheresis; GVHD = graft-versus-host disease; MMF = mycophenolate mofetil
*Only after two systemic treatments
SOURCE: Adapted from Dignan 2012(10) and Sanofi 2023(31)

Unmet need:
There are few treatment options once a patient has tried two or more courses of treatment. Besides corticosteroids, only ruxolitinib and belumosudil are licensed for chronic GVHD in Great Britain; however, neither are currently recommended by NICE.

2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Patients who develop chronic GVHD face a multifaceted burden driven by disease severity and the involvement of multiple organs (1) (18) which exacerbates the distress of already being diagnosed with a life-threatening illness that requires a stem cell transplant (9).

Overview
A systematic review of evidence published between 2007 and 2017 concluded that the most important factors impacting health-related quality of life (HRQoL) in patients with chronic GVHD were disease severity and type of organ involvement (with skin, GI, lung and joint/fascia (connective tissue) manifestations having the greatest negative impacts) (1). In addition to these factors, a further systematic literature review conducted by Sanofi in December 2022, noted that fatigue, depression, anxiety, financial burden and malnourishment reduced the HRQoL of patients with chronic GVHD (27).

In patient interviews conducted in the US, key themes around managing pain and fatigue also affect patients' quality of life. One patient stated, *"chronic GVHD has impacted every facet of my life. I can't go out in the sun. 85% of my skin is affected by chronic GVHD. I can't be a mom to my children or do outdoor activities. My mobility is limited by joint stiffness and pain. I am fatigued, and I sleep a lot."* (30) In the same interview series, a patient with fibrosis stated: *"Because of fibrosis, there are definitely things I can't do now that I used to be able to do. For example, fully lifting my arms, exercising outside, running or lifting heavy items. It's like a rubber band that's been tightened."* (30)

In another interview, when asked to describe the impact of chronic GVHD on quality of life, a US patient stated: *"One symptom that I find harder to deal with than others is the pain. There were sores because the T-cells were attacking the inside of my mouth. So pain with eating and discomfort. That is a low-grade constant pain, but then when you're eating, it's sharp. But eating is an important part of getting well and recovery. When you're growing a new immune system, it takes a huge number of calories."* (28)

Disease Severity and Corticosteroid Use

The health-related quality of life of patients with chronic GVHD is negatively affected by increasing the severity of the disease and a lack of response to corticosteroids (see section 2c for information on corticosteroid response) (16). In an international cross-sectional survey of patient records and patient-completed forms (n=143) that included UK patient data, disease severity was the driving factor in overall symptom burden and number of symptoms.(16) Based on a EuroQol-5 Dimension-5 Level questionnaire (EQ-5D-5L), reported quality of life was lower in patients with severe disease (0.58), than in patients with moderate (0.69) or mild (0.82) disease. (16) Reported quality of life from EQ-5D-5L VAS scores (a self-reported score of heath state) was reported to be lower in patients who relied on steroids to manage symptoms (steroid-dependent) or those who did not respond to steroid treatment (steroid-refractory) compared with steroid-responsive patients. (16) Of the steroid-refractory/dependent patients, 44.0% reported poor HRQoL, compared with 32.3% of steroid-responsive patients. (16)

Psychological Impacts

In addition to functional impairments, patients with chronic GVHD experience significant psychological distress, including depression and anxiety symptoms (9).

An interview study with eight patients with acute or chronic GVHD in England found that patients often feel restricted in what activities they can do, anxiety regarding the unpredictable manifestations of GVHD and inability to plan for the future, inability to go out in public or see family due to the risk of infection, and difficulty adapting to life as a "sick person" with multiple medications and frequent appointments, all of which can lead to depression (9).

When describing the process of getting ready for daily activities, a US patient stated: *"It used to take me about maybe 30 minutes to prepare to go out or to go to work. Now it takes me about 3 hours. I have to make sure that I have all of the medications, creams, eye drops, and mouthwash done, or else I will pay for it if I did not.*" (28)

When describing the impact of chronic GVHD after surviving the initial disease that led to alloHSCT, a US patient in a Sanofi interview stated: "*When I have to be hospitalised for anything regarding GVHD, it's traumatic. It brings up a lot of posttraumatic stress disorder —the beeps of the machines, the IV lines, the nurses and the smells. When you've gone through something like I have, and the primary disease, thank goodness, has not returned, and you're in the hospital for the secondary disease, it's a mindset. It messes with your mind. It's like, wait, I fought this major*

disease, and now I'm in a hospital for another disease that's related to the disease. It's absolutely terrible on the mind. It's terrible." (28)

Economic Impacts

Patients with chronic GVHD face considerable economic pressures due to lost wages and employment changes resulting from the difficulties associated with chronic GVHD (32) (39). Of the patients with chronic GVHD who progressed after two prior lines of therapy included in a multi-national observational study, 39% were on long-term sick leave, retired or unemployed as a result of chronic GVHD. (32)

The most common GVHD symptoms prompting people to stop working were muscle weakness (89%), fatigue (78%), skin problems (78%), shortness of breath (67%) and eye or vision difficulties (67%). (39)

In the UK, 61% of patients reported that they had to be hospitalised in the past 12 months, with infection listed as the predominant reason for hospitalisation (82%). (32)

Impact on caregivers

Carers/families of patients with chronic GVHD are also impacted by the disease as they may have to work fewer hours, leave their job, or retire earlier than planned to fulfil their caregiving requirements. (39) At the time of the multi-national observational survey, 52% of patients had a non-professional caregiver, meaning that the burden of care was likely placed on family members (32). Physicians in the EU-4 and UK reported that the average caregiver of a patient with chronic GVHD who has progressed after two prior lines of therapy spends an average (mean) of 35.2 hours (range: 2 to 160 hours) per week providing care for the patient (32). Partners and spouses spend even more time per week on providing care (mean: 73.5 hours, range: 10 to 160 hours) (32). The main reasons why caregivers changed or reduced their work hours was the patient's depression/anxiety or loss of capability to complete daily tasks (e.g., washing, dressing). (32)

Studies in the US have shown that caregivers of people who have received a donor stem cell transplant often have worse quality of life when compared to the general population. (15)

Depression and sleep disorders are also more likely to occur in these caregivers. (15)

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

Belumosudil is an oral tablet, taken once a day, which has been shown to be effective in patients with chronic GVHD who have received multiple courses of treatment. The link for the approved Summary of Product Characteristics (SmPC) can be found here. The link for the approved Patient Information Leaflet (PIL) can be found here.

Belumosudil targets a protein called ROCK2 (rho associated coiled coil containing protein kinase 2) which influences inflammation and fibrosis. Unlike most other treatments for chronic GVHD, by targeting the ROCK-2 receptor belumosudil rebalances the immune system response to reduce inflammation and fibrosis. (7) This is important as patients who might be eligible for belumosudil are already likely to have problems with their immune system, either from the disease itself, or from exposure to chronic GVHD treatments such as corticosteroids. (6)

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes / No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

Belumosudil is not intended to be used in combination with another medicine.

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

Belumosudil is usually taken once a day as an oral 200mg tablet (patients treated with proton-pump inhibitors (PPIs) or strong CYP3A enzyme-inducing medicines should take one tablet twice a day). It should be taken at approximately the same time, with a meal. Treatment should continue until disease progression or until it is no longer tolerated by the patient.

Belumosudil can be taken at home, so patients do not need to go to a hospital to have it administered. This may be of benefit where patients would otherwise be receiving ECP, which requires frequent hospital appointments and is not always easily accessible to patients. (10)

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

Belumosudil has been studied in two main clinical trials to date. This included a Phase 2a study (KD025-208) and a Phase 2b study (KD025-213 / ROCKstar), in a total of 186 patients, who had previously received between 1 to 5 prior lines of therapy. This means the population was representative of patients with advanced disease and not many treatment options left available to them. Both studies were conducted in the United States. Further details are provided below.

Phase 2a study (KD025-208)	
Official Title:	A Phase 2a, Dose-Escalation, Open-Label Study to Evaluate the Safety, Tolerability, and Activity of KD025 in Subjects With Chronic Graft Versus Host Disease
Number of patients:	54
Study Start Date:	September 2016
Primary Completion Date:	April 2022
Study Completion Date:	May 2022
Clinicaltrials.gov identifier:	NCT02841995
Publication:	Jagasia M, et al. ROCK2 Inhibition With Belumosudil (KD025) for the Treatment of Chronic Graft-Versus-Host Disease. J Clin Oncol. 2021 Jun 10;39(17):1888-1898. (Jagasia, M. H., Lazaryan A., Bachier C., et al (2021))

Phase 2b study (KD025-213 / ROCKstar)	
Official Title:	A Phase 2, Randomized, Multicenter Study to Evaluate the Efficacy and Safety of KD025 in Subjects With cGVHD After At Least 2 Prior Lines of Systemic Therapy (The ROCKstar Study)
Number of patients observed from latest data cut-off	132
Study Start Date:	October 2018
Estimated Primary Completion Date:	December 2024
Estimated Study Completion Date:	December 2024
Clinicaltrials.gov identifier:	NCT03640481
Publication:	Cutler, C., Lee S., Arai S et al , Belumosudil for chronic graft-versus-host disease after 2 or more prior lines of therapy: the ROCKstar Study. Blood. 2021 Dec 2;138(22):2278-2289. (Cutler, C., Lee S., Arai S et al 2021)

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

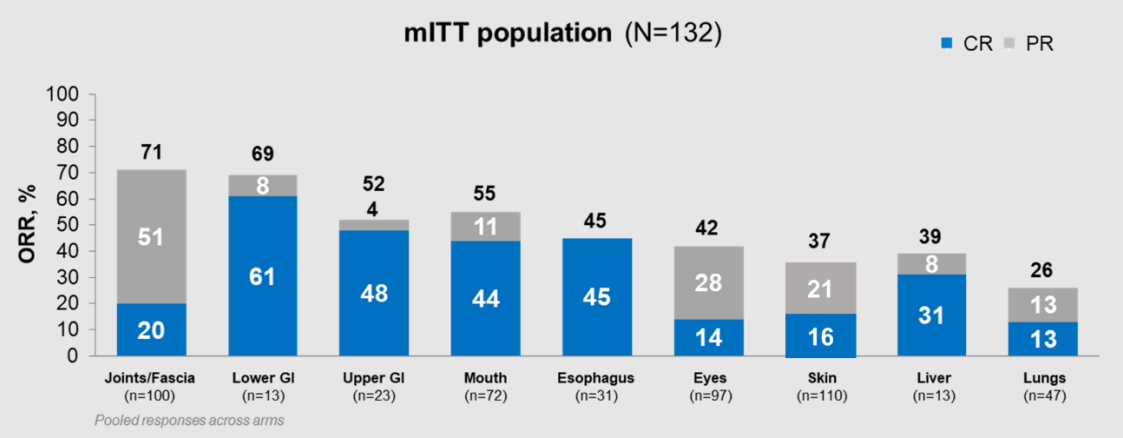
In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

Overall response rate (ORR):
ORR is a measure of efficacy of the treatment. It is the proportion of patients who achieve either a partial response or a complete response to a treatment. (36)

In the ROCKstar study, patients were split into two groups, with 66 patients receiving 200mg belumosudil once per day and 66 receiving it twice per day. After 12 months, the ORR was 74% (95% CI, 62-84) for patients in the group receiving it once daily, and 77% (95% CI, 65-87) in the group receiving it twice daily. Seven patients achieved a complete response in all affected organs. The average (median) duration of response was 54 weeks. (7)

In patients treated with belumosudil, responses were observed regardless of the severity of the disease or number of organs involved. Responses were seen in all organs, including the lungs, liver, and skin, which can be especially difficult to treat. They were also consistent between patients who had started 4 or more types of treatment prior to belumosudil (ORR=74) and those who had started less than 4 types of treatment (ORR=78). (7)

Figure 3. ORR by organ system in the ROCKstar study (7)



Abbreviations: CR = complete response; mITT = modified intention-to-treat; ORR = overall response rate; PR = partial response.
Organ-specific analyses in the mITT population demonstrated ORRs in the skin, eyes, mouth, liver, lungs, joints/fascia, upper GI tract, lower GI tract, and oesophagus. CR was seen across all affected organs.

In the Phase 2a study (54 patients), ORR was achieved in 65% (95% CI, 51-77) of patients. (14)

Failure free survival (FFS):
FFS is an alternative measure of the efficacy in chronic GVHD treatments used in the clinical trials. It is defined as the time between starting treatment (belumosudil) and either starting a new chronic GVHD treatment, a return of cancer symptoms (known as relapse), or dying of something other than relapse. (24)

In the ROCKstar Phase 2b study, the FFS was 75% (95% CI, 66-81) at 6 months and 56% (95% CI, 47-64) at 12 months. The most common reason for failure was starting a new chronic GVHD treatment (38%). 89% of patients were still alive after 2 years of treatment. (7)

Corticosteroid dose reduction:
Reducing the dose of corticosteroids is an important treatment goal in chronic GVHD, as these medicines are associated with a range of unpleasant side effects. (19)

In the ROCKstar study, 65% of patients reduced their corticosteroid dose during treatment with belumosudil. The mean corticosteroid dose was reduced by 45% in all patients, with a mean dose reduction of 54% in patients who responded. 21% of patients stopped taking corticosteroids altogether. In addition, 22% of those patients stopped taking calcineurin inhibitors, 20% stopped taking sirolimus, and 21% stopped taking mycophenolate therapy. (7)

Results were similar in the Phase 2a study, with 67% of patients reducing their corticosteroid dose while being treated with belumosudil, and 19% stopping corticosteroid treatment. (14)

Limitations
Both the ROCKstar and the Phase 2a trials were 'single-arm'. This means that the efficacy of belumosudil was assessed after an average (median) of at least 2 prior lines of therapy, but not in direct comparison to another therapy. In addition, both trials were run in the United States, where some of the initial treatments used in chronic GVHD patients are different to what is prescribed in Great Britain.

Comparators

There is a lack of robust clinical data for the treatment of chronic GVHD in the literature. The ROCKstar study of belumosudil provides some of the best available clinical efficacy and safety evidence for a treatment in this therapy area and is the data source from which the belumosudil marketing authorisation was granted. However, as a single-arm study in heavily pre-treated (at least two prior systemic therapies) patients, it does not enable direct comparison with other treatment options.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

Clinical trial patient reported outcomes:
The Lee Symptom Scale (LSS), is a measure of chronic GVHD symptom severity. It covers a wide range of relevant symptoms including itchy skin, dry eyes, mouth ulcers, depression, and anxiety. Once the presence of symptoms is established, patients report how "bothered" they feel about each symptom over the previous month using a five-point scale from "not at all" to "extremely". This tool has been used to show that symptoms of chronic GVHD (particularly moderate to severe

chronic GVHD) can have a major negative impact on patients. An improvement of 7 points or more is considered to be clinically meaningful (17). The LSS score was one of the outcomes measured in the ROCKstar study.

In the ROCKstar study, 59% of the population receiving one belumosudil tablet per day, and 62% of the population receiving two tablets per day, had a clinically meaningful improvement on the LSS score of 7 points or more. This was higher in the populations who responded to treatment (69% in the once-daily group and 71% in the twice-daily group). (7)

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Clinical trial data

In the ROCKstar study, the main side effects of belumosudil were consistent with those expected for chronic GVHD patients taking corticosteroids and other immunosuppressant medicines. The most common side effects were fatigue (38%), diarrhoea (33%), nausea (31%), cough (28%), and upper respiratory tract infection (27%). 38% of subjects had one or more serious side effects, with the most common of these being pneumonia (8%). The safety overview and list of most common side effects are reported in Tables 1 and 2 below (7):

Table 1: Belumosudil safety overview from ROCKstar Phase 2b study (7)

Safety overview	REZUROCK 200mg QD (n=66)	REZUROCK 200mg BID (n=66)	Overall (n=132)
Median duration of treatment, mo	9.4	11.8	10.4
Any AE, n (%)	65 (99)	66 (100)	131 (99)
Grade ≥ 3 AEs, n (%)	37 (56)	34 (52)	71 (54)
SAEs, n (%)	27 (41)	23 (35)	50 (38)
Drug-related AEs, n (%)			
Any related AE	49 (74)	40 (61)	89 (67)
Related SAEs	5 (8)	2 (3)	7 (5)
Deaths[a], n (%)	8 (12)	6 (9)	14 (11)

[a] Six subjects died during long-term follow-up (>28 days after last dose)

Abbreviations: AE = adverse event; AML = acute myeloid leukaemia; BID = twice daily; MODS = multiple organ dysfunction syndrome; SAE, serious adverse event; QD = once daily.

Table 2: Commonly reported side effects from ROCKstar Phase 2b study (7)

Commonly reported AEs, n (%)	REZUROCK 200mg QD (n=66)	REZUROCK 200mg BID (n=66)	Overall (N=132)
All grades in ≥20% of patients			
Fatigue (tiredness)	30 (46)	20 (30)	50 (38)
Diarrhea	23 (35)	21 (32)	44 (33)
Nausea (feeling sick)	23 (35)	18 (27)	41 (31)
Cough	20 (30)	17 (26)	37 (28)
Upper respiratory tract infection	17 (26)	18 (27)	35 (27)
Dyspnea (shortness of breath)	21 (32)	12 (18)	33 (25)
Headache	13 (20)	18 (27)	31 (24)
Liver-related AEs	12 (18)	19 (29)	31 (24)
Peripheral edema (swelling due to fluid accumulation in the lower limb)	17 (26)	13 (20)	30 (23)
Vomiting	18 (27)	10 (15)	28 (21)
Muscle spasms	13 (20)	13 (20)	26 (20)
Grade ≥3 in ≥5% of patients			
Pneumonia	6 (9)	4 (6)	10 (8)
Hypertension (high blood pressure)	4 (6)	4 (6)	8 (6)
Hyperglycaemia (high blood sugar)	3 (5)	3 (5)	6 (5)

Abbreviations: AE = adverse event; BID = twice daily; QD = once daily.

As these side effects are in line with those of the current treatments, the risk benefit ratio appears favourable.

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration

Due to this being an oral medication, convenience remains a key benefit as some of the alternative treatments require intravenous administration, and require an overnight hospital stay or other overnight accommodation. An oral treatment is beneficial for patients and caregivers, in particular for patients with low mobility due to chronic GVHD.

From an efficacy point of view, patients in this stage of the chronic GVHD journey often run out of options for management of their condition since most of the treatments have been tried with a relative small number of licensed interventions. For some patients, discontinuation of

corticosteroids or calcineurin inhibitors (due to starting belumosudil) can be a relief from the adverse effects resulting from the long term use of these agents.

Evidence from the ROCKstar study shows low rates of viral reactivations that are often feared after such treatments. The safety profile is well described in the tables above. In the ROCKstar study, there was one case of Epstein-Barr virus (a cause of glandular fever) and one case of reactivation of cytomegalovirus (a common viral infection related to the herpes virus).

Since belumosudil targets the fibrotic pathways, evidence from the ROCKstar study has shown that responses have been obtained in patients with multiorgan and fibrotic manifestations such as the skin. Responses have also been obtained the difficult to treat organs like the lungs and liver.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

Until now, there is a relative lack of experience in the Great Britain as the belumosudil trials were conducted in the United States.

The side effect profiles are described in Tables 1 and 2. These describe the side effects (adverse events) that have been observed in the patients who have received belumosudil after 2 lines of therapy.

The belumosudil data that we have gathered so far have been in trials that did not have a comparator (a so-called 'non-controlled trial'). This means that it is difficult to provide a reference point for the disadvantages as compared to the other current methods of treatment.

3i) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?)
- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?

- How the condition, taking the new treatment compared with current treatments affects your quality of life.

How the model reflects the condition

The economic model submitted to NICE aims to evaluate the value for money (or cost-effectiveness) of belumosudil. It does this by estimating the costs and health benefits over the course of a treatment journey for patienta with chronic GVHD receiving belumosudil compared to a patients receiving alternative treatments. The model represents the progressive course of disease from the point of receiving a third therapy for chronic GVHD. Patients move through a series of health states which mimic those evaluated in the clinical trials (for example, failure-free survival with partial response to treatment, or failure due to recurrence of cancer). Healthcare costs and quality of life status (utility) are then assigned to each health state and summed together at the end of the model.

Modelling how much the treatment improves quality of life

Treatments improve quality of life in the model by keeping patients in the failure-free health states (which include complete, partial, or lack of response). Treatment failure or relapse of cancer are associated with a reduction in quality of life (utility). Where possible data from patients in the clinical trials was used to inform the utility of patients in the model. However, this data was not available for patients whose treatment had failed, or whose cancer had relapsed. Therefore, we had to use other sources including previous NICE submissions and clinical expert opinion. The model also accounts for disutility (reduction in quality of life) associated with treatment side effects and intravenous (IV) infusions. The quality of life of caregivers was included to account for the significant impact chronic GVHD can have on carers or family members. However, since data was lacking in this population, we applied quality of life data from caregivers of patients with multiple sclerosis instead, as advised by clinicians.(31)

Modelling costs of treatment

As well as drug costs, the model includes costs of hospital care for chronic GVHD (based on NHS hospital data), costs of administering medicines, and costs of adverse events from treatments.

Belumosudil is a tablet which can be taken at home. This is a benefit compared to treatments such as ECP which is expensive to administer and requires healthcare professionals in a hospital setting.

We assumed that healthcare costs reduce over time for a patient responding to treatment, since they are likely to require fewer hospital appointments as time goes on.

Uncertainty

Several assumptions had to be made in the model because there is limited data available for treatments in this very rare disease.

The Phase 2 clinical trials did not include a comparator treatment. Therefore, we have used clinical data from a different trial (REACH-3) to model the impact of best available therapy. However, this data included patients with chronic GVHD who had received only one previous treatment, meaning they were at an earlier stage in the disease pathway.

Chronic GVHD is a complicated disease which can present in many different ways. It was not possible to model all the different organs that could be affected, so we used the best available data to provide a simplified approach which represented the overall patient population.
As described above, there was also uncertainty related to the health-related quality of life of patients in the model. The quality of life data used from the ROCKstar trial was from a measure

(PROMIS-GH) which was not designed to inform cost-effectiveness modelling. These may not fully reflect the burden of disease in the population.

There was also uncertainty related to the costs of managing the disease. Although the data came from NHS hospital database, it was not possible to match this accurately to the patient population in the model.

We ran several analyses to test the impact of the main uncertainties on the model results. Despite the uncertainty, the majority of scenarios tested were found to be cost-effective.

Additional considerations

Chronic GVHD is a severe disease with a major impact on quality of life for patients, particularly those who have not responded to early treatments. To reflect this, we have made the case for a severity modifier of 1.2 to be applied to the economic modelling.

Given the capacity challenges within the NHS, belumosudil may help to reduce pressure on the healthcare system by keeping patients out of hospital for longer. This is particularly true when comparing to ECP, which can only be given in a hospital setting. This system benefit has not been formally included in the model analysis.

Summary

Based on the evidence available and all the above considerations, the economic analysis shows belumosudil to be a good use of NHS resources as an oral treatment option for patients with chronic GVHD.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

Belumosudil is innovative in its mechanism of action, which impacts both inflammation and fibrosis. As an oral therapy licensed in chronic GVHD who have failed previous treatments, it represents a step change in management for these patients.

It was licensed under the Project Orbis programme (20), which was set up to review promising cancer drugs to help patients access treatments faster. Belumosudil also received an Innovation Passport by the MHRA in April 2021 (ILAP reference number ILAP/IP/21/53904/01).

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here
People who receive a transplant from a mismatched unrelated donor are more at risk of developing chronic GVHD. Currently, only 37% of patients from a minority ethnic background can find the best possible stem cell match from a stranger (compared to 72% of patients from White backgrounds (3). As people from a non-white family background are less likely to find a related donor match, people from an ethnic minority family background are at an increased risk of developing chronic GVHD. (23) The current lack of effective treatments with a favourable safety profile may disadvantage these populations. Skin manifestations are some of the most common and major complications of chronic GVHD, and skin assessments are required for disease diagnosis, yet current physician- and patient-reported outcome measures may not adequately capture the subtle changes in patients with non-white skin, leading to potential errors or delays in diagnosis for such patients (34). Minority ethnic patients may also experience increased stigma around skin changes (hyper- or hypopigmentation, chronic skin shedding) and bowel urgency, which may lead to a greater degree of social isolation and an increased psychological impact. Geographical access to ECP services and specialist blood and marrow transplant clinics can be a barrier to people in lower socioeconomic groups who may be unable to take time off work or afford to travel to appointments. Patients with lower socioeconomic status may have to decline ECP if they fall outside of the travel distance requirements that would grant them free accommodation between the two therapy days. Having the option of an oral, at home treatment alternative could be particularly beneficial in these groups.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc. Where possible, please provide open access materials or provide copies that patients can access.
Further information on NICE and the role of patients**:** • Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE

- NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE
- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

Adverse event	An adverse event is the occurrence of an undesirable event during or following the exposure to the drug, but not necessarily caused by the drug itself.
Licensed	A medicine that has received a 'market authorisation' by the Medicines and Healthcare products Regulatory Agency (MHRA)
Fibrosis	The scarring and thickening of any affected tissue as wound healing response which interferes with normal organ function.
95% CI (Confidence Interval)	This is the range of values that is likely to include 95% of all the observed data
ECP / extracorporeal photopheresis	A common form of therapy for chronic GVHD which involves exposing white blood cells in the patient's body to ultraviolet light and a medication called methoxypsoralen (8-MOP)
FFS / failure free survival	An alternate measure of efficacy used in clinical trials for chronic GVHD which measures the time between starting treatment and either switching therapy, relapsing, or dying for a reason other than disease relapse.
MHRA / Medicines and Healthcare products Regulatory Agency	The regulatory body which grants marketing authorisation for medicines to be used in the UK
ORR / Overall Response Rate	An outcome used in clinical trials for chronic GVHD which measures the proportion of patients who either have a partial or complete response to treatment

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

1. Agh T, Csanadi M, Voko Z, Webb T, Jeyakumaran D, Trudeau J, et al. Humanistic burden of patients with chronic graft-versus-host disease - systematic literature review of health-related quality of life and functional status. Expert review of hematology. 2019;12(5):295-309.
2. Anthony Nolan (2021), Graft versus host disease (GvHD), https://www.anthonynolan.org/patients-and-families/recovering-a-stem-cell-transplant/graft-versus-host-disease-gvhd, [Date Accessed 15/03/2023 0832]
3. Anthony Nolan (2022) Facts and Stats. https://www.anthonynolan.org/front/facts-and-stats [Date accessed 20 Dec 2022]
4. British Society of Blood and Marrow Transplantation and Cellular Therapy (BSBMTCT) (2015). *The outcome of haematopoetic stem cell transplantation: an analysis of registry data for UK transplants performed 2009-2014 inclusive*. Retrieved from https://www.uhs.nhs.uk/Media/SUHTExtranet/Services/WessexBloodAndMarrowTransplantationService/BSBMT-8th-Report-To-Commissioners.pdf [Accessed 09/03/2023 0914]
5. BSBMTCT. (2020). *British Society of Blood and Marrow Transplantation and Cellular Therapy Annual activity 2020*. Retrieved from https://bsbmtct.org/activity/2020/ [Accessed 09/03/2023 1302]
6. Choi S.W. and Runaas L. (2023), Graft-versus-host disease, BMJ Best Practice, https://bestpractice.bmj.com/topics/en-gb/946 [Date Accessed 15/03/2023, 0836]
7. Cutler, C., Lee S., Arai S et al (2021). Belumosudil for chronic Graft-Versus Host Disease (cGVHD) after 2 or more lines of therapy: The ROCKstar Study. *Blood*, 138: 2278-2289
8. DeFilipp, Z., Alousi, A., Pidala, J. et al (2021). Nonrelapse mortality among patients diagnosed with chronic GVHD: an updated analysis from the Chronic GVHD Consortium. *Blood Advances*, 5(20):4278-4284.
9. de Vere Hunt I, Kilgour JM, Danby R, Peniket A, Matin RN. "Is this the GVHD?" A qualitative exploration of quality of life issues in individuals with graft-versus-host disease following allogeneic stem cell transplant and their experiences of a specialist multidisciplinary bone marrow transplant service. Health and Quality of Life Outcomes. 2021;19(1):11.
10. Dignan, F., Amrolia, P., Clark, A. et al (2012). Diagnosis and management of chronic graft-versus-host disease. *British Journal of Haematology*, 158,46–61.
11. El-Jawahri, A., Pidala, J., Khera, N. et al (2018). Impact of psychological distress on quality of life, functional status, and survival in patients with chronic graft-versus-host disease. *Biology of Blood and Marrow Transplantation, 24*(11), 2285-2292.
12. Filipovich, A., Weisdorf, D., Pavletic, S. et al (2005). National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biology of Blood & Marrow Transpl, 11*(12), 945-956.
13. Jagasia, M. H., Greinix H., Arora M. et al (2015). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. *Biology of Blood Marrow Transplant*, 21(3):389-401.e381.
14. Jagasia, M. H., Lazaryan A., Bachier C., et al (2021). ROCK2 Inhibition With Belumosudil (KD025) for the Treatment of Chronic Graft-Versus-Host Disease. *J Clin Oncol*, 39(17):1888-1898.

15. Jamani K, Onstad LE, Bar M, Carpenter PA, Krakow EF, Salit RB, et al. Quality of life of caregivers of hematopoietic cell transplant recipients. *Biology of Blood and Marrow Transplantation*. 2018;24(11):2271-6
16. Lachance S, Hamad N, de Courcy J, Gibson G, Zuurman M, Mohty M, editors. Impact of Chronic GVHD Severity And Steroid Response on The Quality of Life In Patients Following Allogeneic Stem Cell Transplantation: Findings From A Real-World Study. Bone Marrow Transplantation; 2021
17. Lee, S., Cook, E., Soiffer, R. et al (2002). Development and validation of a scale to measure symptoms of chronic graft-versus-host disease. *Biology of Blood Marrow Transplant*, 8(8):444-452.
18. Lee SJ, Onstad L, Chow EJ, Shaw BE, Jim HSL, Syrjala KL, et al. Patient-reported outcomes and health status associated with chronic graft-versus-host disease. *Haematologica*. 2018;103(9):1535-41.
19. Liu, D., Ahmet, A., Ward, L. et al (2013). A practical guide to the monitoring and management of the complications of systemic corticosteroid therapy. *Allergy, Asthma & Clinical Immunology*, 9:30, 1-25
20. MHRA Guidance – Project Orbis (2023), https://www.gov.uk/guidance/guidance-on-project-orbis (Date Accessed 05/04/2023 1753)
21. NHS (2022), Overview - Stem Cell and Bone Marrow Transplants, https://www.nhs.uk/conditions/stem-cell-transplant/ [Date Accessed 15/03/2023 0839]
22. NHS Blood and Transplant (2014), Patient Information Leaflet - Extra Corporeal Photopheresis, https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/1945/extracorporeal-photopheresis.pdf [Date Accessed 15/03/2023 0843]
23. NICE 2023, Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021], Final Scope, https://www.nice.org.uk/guidance/gid-ta11136/documents/final-scope [Date Accessed 09/03/2023 1258]
24. Palmer J., Chai., X, Martin P. et al (2015). Failure-free survival in a prospective cohort of patients with chronic graft-versus-host disease, *Haematologica*, 100 (5), 690-695
25. Pidala, J., Kurland, B., Xiaoyu, C. et al (2011). Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. *Blood*, 117(17):4651-4657.
26. Salhotra, A., Eizenhammer D., Hennegan, K. et al (2020). Presentation and management of chronic graft-versus-host disease in real-world clinical practice: A medical chart audit. *EHA Virtual Congress.*
27. Sanofi. SLR Technical Report: Health Related Quality of Life (HRQoL) and Health Utility Values in Patients with Chronic Graft-Versus-Host Disease. (data on file). 2022
28. Sanofi. RZR Patient Video Clips (data on file). 2022.
29. Sanofi. Data on File. Patient Video Transcript - Bringing Clarity. 2022.
30. Sanofi. Data on File. Patient Video Transcript - QOL Impact. 2022
31. Sanofi. Data on File. NICE Belumosudil Virtual Advisory Board Report. 2023.
32. Sanofi. GVHD Disease Specific Programmes (DSP) Analysis Slide Deck. (data on file). 2022.
33. Sanofi. GVHD Disease Specific Programmes (DSP) Analysis Tables. (data on file). 2022.
34. Smith ZI, Rosenstein RK, Banerjee S, Pichard DC, Pavletic SZ, Cowen EW (2020). Quality of Life in Patients With Skin of Color and Chronic Graft-vs-Host Disease. *JAMA Dermatol.* 156(5):589-90.
35. Styczynski, J., Tridello, G., Koster, L. et al (2020). Death after hematopoietic stem cell transplantation: changes over calendar year time, infections and associated factors. *Bone Marrow Transplant, 55(1)*, 126-136.
36. Villaruz L. and Socinski M., The Clinical Viewpoint: Definitions, Limitations of RECIST, Practical Considerations of Measurement, *Clin Cancer Res.,* (2013) May 15; 19(10): 2629-2636

37. Wong, F., Francisco, L., Togowa, K. et al (2010). Longterm recovery after hematopoietic cell transplantation: predictors of quality-of-life concerns. *Blood*, 115(12):2508-2518.
38. Worel, N., Biener, D., Kalhs, P. et al (2002). Long-term outcome and quality of life of patients who are alive and in complete remission more than two years after allogeneic and syngeneic stem cell transplantation. *Bone Marrow Transplantation*, 30:619-626.
39. Yu JT, J.; Stewart, SK.; Wartman, L.; Williams, P.; Amoloja, T.; et al., editors. . Impact of Chronic Graft-Versus-Host Disease on Patients' Quality of Life, Employment, Productivity, Income, and Caregivers. Abstract presented at 2021 TCT Transplantation & Cellular Therapy. Meetings of ASTCT and CIBMTR TCT Meetings; February 8, 2021

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Clarification questions

April 2023

File name	Version	Contains confidential information	Date
ID4021 Clarification questions_SANOFI_05042023_FINAL [ACIC]	**FINAL**	**Yes** Academic in confidence; commercial in confidence	**25/04/2023**

Section A: Clarification on effectiveness data

Data cut

A1. Priority question: In the company submission (CS), the company stated a later data cut of ROCKstar was available in late 2022, but the economic model uses data from August 2021. Please update the pooled analysis for the ≥2 prior lines of therapy subgroup, and thus update the economic model, with the 2022 data cut.

We have updated the model with the 2022 data cut comprising the pooled analysis for the ≥2 prior lines of therapy subgroup and licensed dose (See ID4021_Belumosudil_CEM_25042023_ACIC.xlsm attached). The updated base case cost-effectiveness results are presented below.

All of the curves have been implemented in the model (and can be visualised as before) but there has been insufficient time within the period available for return of the clarification responses to fully develop a complete narrative description of the new data set. We will follow up with this as soon as we are able. We would appreciate guidance from the EAG about what might be considered critical deliverables and an acceptable timeframe for their delivery.

Results from the updated model

The cost-effectiveness probabilistic and deterministic estimates are provided below taken from the updated model including changes to the drug acquisition costs taken from eMIT, as recommended in questions C4 and C5, but otherwise use the original base case settings (Table 3 and Table 4). These correspond to tables 61 to 64 in the CS. The updated one-way deterministic sensitivity analysis (DSA) and ICERs for selected sensitivity analyses, with and without the application of the severity modifier are presented in Appendix A. The recalculation of the original base case results includes the correction to the mycophenolate mofetil 250 mg price and to the imatinib 100 mg and 400 mg prices from eMIT noted in questions C4 and C5. It also includes the updated list of modelled Grade ≥3 AEs based on the new data cut.

There has been a significant drop in the probabilistic ICER based on the 2022 data cut, to £3,046 (new) vs. £15,032 (old) (without severity modifier) or £2,539 (new) vs. £12,526 (old) (with severity modifier) in the original CS. We investigated the reasons behind this change,

in order to evaluate the stability of the model both in terms of incremental quality of life benefit and costs. These are discussed in more detail below.

Quality of life benefit

We found that the model provided a stable probabilistic result in terms of QALYs compared to the previous analysis (https://sanofi.sharepoint.com/sites/UK-Ireland-Medical/HealthOutcomes/Therapy/Genmed/rezurock/NICE/Clarification%20questions/ID4021%20belumosudil%20additional%20clarification%20questions_12052023_Sanofi%20response%20%5bACIC%20redacted%5d.docx?web=1 (new) vs. [redacted] (old) incremental QALYs, without severity modifier or [redacted] (new) vs. [redacted] (old) incremental QALYs, without severity modifier). This marginal improvement in QoL was due to the improved FFS observed in the later data cut (median FFS: [redacted] for the 19th August 2021 and 30th September 2022 data cuts respectively, pooling both belumosudil arms).

The Generalized Gamma curve fit was chosen for FFS for the new data cut. This was consistent with the previous iteration of the model based on AIC / BIC. Table 1 provides the updated AIC/BIC rankings.

Table 1. Information criteria for the FFS fits (30th September 2022 data cut)

	AIC	BIC	Rank AIC	Rank BIC
Exponential	391.6361	397.9771	7	6
Weibull (PH)	388.4621	397.9735	5	5
Gompertz	372.7933	382.3048	3	3
Log-logistic	375.4974	385.0088	4	4
Log-normal	369.8527	379.3642	2	2
Gamma	390.9984	400.5099	6	7
Generalized Gamma	363.9921	376.6741	1	1

Costs

Recalculation of the base case with the 2022 data cut produces a significant decrease in the ICER due to changes in the estimated cost of treatment with belumosudil. This is driven mainly by application of the full data set for time on treatment.

The KD025-208 and ROCKstar (adult cohort) studies had completed by the time of the 30th of September 2022 data cut, but ROCKStar continues to recruit adolescent patients. This means that adult patients in these studies were characterised as having discontinued treatment while for the 19th of August 2021 data cut some patients were still receiving

treatment. Table 2 provides the reasons listed for treatment discontinuation in the 30th of September 2022 data cut and Figure 1 overleaf compares the KM data for the 2 data cuts.

Table 2. Reasons for treatment discontinuation in the updated data cut of the ROCKStar and KD025-208 studies (≥2 prior lines of therapy subgroup))

Reason for discontinuation	Number of patients
Adverse event	[redacted]
Death	[redacted]
Disease relapse	[redacted]
Failure to meet continuation criteria	[redacted]
Non-compliance with study drug	[redacted]
Non-compliance with protocol	[redacted]
Other	[redacted]
Physician decision	[redacted]
Progressive disease	[redacted]
Study terminated by sponsor	[redacted]
Withdrawal by subject	[redacted]
Total	176

Following investigation of the reason for treatment discontinuation, we decided to carry out a revised analysis which characterised patients who discontinued treatment because of study termination as censored at the time they received their last dose of study drug. This recognises that these patients would have likely received treatment for longer should the study have continued, or had they been treated in real world clinical practice.

We have provided an overlay of these KM curves in Figure 1 below, for TTD with the 19th of August 2021 data cut originally used in the CS and with two options for the KM curve from the 30th of September 2022 data cut (these include censoring of patients who discontinued treatment because of study termination and no censoring). With the updated data cut, there is no plateau of the time on treatment curve evident around 12% for the 19th of August 2021 data. This means that the parametric survival estimates obtained using the 30th of September 2022 data cut are inevitably shorter than the ones used in the original CS. We note that the time on treatment for all therapies is truncated at 5 years in the model so a very long tail is not implicated but nonetheless this change to the 2022 data cut introduces a reduction in overall belumosudil cost.

Figure 1 [redacted]

There is no substantive difference in the curve fits for overall survival and duration of response between the 19th of August 2021 data cut and the 30th of September 2022 data cut.

Results from the updated model

The cost-effectiveness estimates are provided overleaf.

Table 3. Base-case probabilistic results (PAS, WITHOUT and WITH severity modifier)

	WITHOUT severity modifier		WITH severity modifier	
Outcome	**Belumosudil**	**BAT**	**Belumosudil**	**BAT**
Total costs	[redacted]	£ 250,314	[redacted]	£ 250,314
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]		[redacted]	
Incremental LYs	[redacted]		[redacted]	
Incremental QALYs	[redacted]		[redacted]	
ICER (£/QALY)	£ 3,046		£ 2,539	
INHB (£20,000/QALY)	[redacted]		[redacted]	
INHB (£30,000/QALY)	[redacted]		[redacted]	
INMB (£20,000/QALY)	[redacted]		[redacted]	
INMB (£30,000/QALY)	[redacted]		[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

Table 4. Base-case deterministic results (PAS, WITHOUT and WITH severity modifier)

	WITHOUT severity modifier		WITH severity modifier	
Outcome	**Belumosudil**	**BAT**	**Belumosudil**	**BAT**
Total costs	[redacted]	£ 248,736	[redacted]	£ 248,736
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]		[redacted]	
Incremental LYs	[redacted]		[redacted]	
Incremental QALYs	[redacted]		[redacted]	
ICER (£/QALY)	£ 3,571		£ 2,976	
INHB (£20,000/QALY)	[redacted]		[redacted]	
INHB (£30,000/QALY)	[redacted]		[redacted]	
INMB (£20,000/QALY)	[redacted]		[redacted]	
INMB (£30,000/QALY)	[redacted]		[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

The results for the probabilistic analysis (without severity modifier) are plotted on the cost-effectiveness plane for belumosudil compared to BAT over 5,000 iterations in Figure 2 below (corresponding to Figure 35 in the CS). The majority of iterations (58%) fell in the north-east quadrant, indicating that treatment with belumosudil was consistently more effective and more costly than BAT. The remainder (42%) of the iterations fell in the south-east quadrant

which indicates belumosudil being a dominant treatment option, providing better health outcomes at lower costs compared to BAT.

Figure 2. [redacted]

[redacted]

The cost-effectiveness acceptability curves are presented in Figure 3 below (corresponding to Figure 36 in the CS). Belumosudil is more likely to be the optimal treatment choice vs. BAT at WTP thresholds above £5,200/QALY. The probability of belumosudil being the optimal treatment choice vs. BAT at a £30,000/QALY WTP threshold is 86.2%.

Figure 3. [redacted]

[redacted]

Further updates to the economic model.

We have carefully considered all of the comments made by the EAG and present an alternative scenario in Appendix B which includes plausible adjustments to the original analysis.

Population

A2. Priority question: Belumosudil has been given marketing authorisation for the treatment of patients aged 12 years and older with chronic graft-versus-host disease (chronic GvHD) who have received at least two prior lines of systemic therapy. However, neither ROCKstar or KD025-208 recruited any participants under 18 years of age in the latest data cut presented.

a) **Please explain the basis on which belumosudil has been granted marketing authorisation in children aged 12 years and older?**

b) **Please provide a clinical rationale why the dosing regimen, developed in an adult population, will be effective and safe in a population of children aged 12 years and older.**

c) **Are there ongoing trials of belumosudil that include children aged 12 years and older with chronic graft-versus-host disease (cGvHD) who have received at least two prior lines of systemic therapy?**

Answer to A2a

The aetiology, pathomechanism and risk factors for outcomes of cGVHD are the same in adolescents as in adults (Jacobsohn, 2010; Jacobsohn et al., 2011; Baird et al., 2010). Furthermore, based on the available literature, clinical response to treatment and the pharmacokinetic profile of belumosudil are expected to be the same in adolescents and adult populations. Both the United States Prescribing Information (USPI), and the Australian Product Information, acknowledge that the course of the disease is sufficiently similar in adult and paediatric patients to allow extrapolation of data (REZUROCK USPI, 2021; RHOLISTIQ Australian PI, 2021).

Although the number of patients in the belumosudil pivotal studies in this rare disease was not large and originally no adolescents were included in the studies, the efficacy and

favourable adverse event profile, of belumosudil demonstrated in the adult cohort, combined with the observations above indicate that adolescent patients with cGVHD are likely benefit from treatment with belumosudil. A protocol amendment made to study KD025-213 (ROCKstar) to enrol adolescent subjects was approved on 01 June 2020, and recruitment efforts continue.

It is likely that the benefit of treatment outweighs the risks within the adolescent population because there is significant lack of treatment options for this patient group diagnosed with cGvHD. An important consideration is the fact that cGvHD is rare in adolescents, particularly those with late-stage disease who have progressed on multiple prior lines of therapies. This severely limits the opportunity for enrolment into clinical trials. Therefore, there remains a need to extrapolate from pharmacokinetic studies as well as to transpose available adult data to adolescent patients where applicable.

In addition to the disease pathophysiology being unchanged between adolescents and adults, it was deemed appropriate to include adolescent patients in the indication for the following reasons:

No specific risks relevant to adolescent patients are expected.

There are no risks specific to the adolescent population with belumosudil. Important identified and important potential risks established for the adult population are equally applicable to the adolescent population; however, these risks are adequately managed through routine pharmacovigilance activities and information in the Summary of Product Characteristics (SmPC).

When considering adverse effects that may be of particular concern to adolescent patients, such as infections and gastrointestinal (GI) adverse events (AEs) such as diarrhoea (Jacobsohn, 2010), belumosudil presents a favourable safety profile. Mycophenolate mofetil (MMF) is associated with significant diarrhoea (Jacobsohn, 2010). With pentostatin therapy, the main toxicity is infection, with studies showing Grade 3 and 4 infectious events in up to 20% of patients (Jacobsohn, 2010). In comparison to these widely used therapies, with belumosudil most infections in the pooled cGVHD safety analysis group were mild or moderate, non-serious and assessed as not related to belumosudil.

Belumosudil has not demonstrated any overt impact on growth and development in the Good Laboratory Practice (GLP) general toxicology studies in adolescent/adult animals at clinically relevant exposures. Impact on male fertility (GLP rat fertility study) in rats has generally been reversible and has been observed at exposure levels higher than expected

clinically. In addition, belumosudil is not brain penetrant based on the rat quantitative whole-body autoradiography study, had no pharmacologically relevant in vitro activity in a panel of central nervous system (CNS) targets, and had no CNS-related findings in the rat CNS safety pharmacology/general toxicology studies. Therefore, it is expected that belumosudil has low potential for impact on CNS development in adolescents.

Clinical response to treatment is likely to be the same in adults and adolescents.

The treatment of cGVHD in adolescents is mostly extrapolated from the experience in adults (Jacobsohn, 2010). The immune system at the time of adolescence is relatively mature and comparable to adults (Georgountzou & Papadopoulos, 2017), and the standard of care for the treatment of cGVHD in adolescents is the same as for adults comprising immunosuppressive therapies.

Regardless of the choice of treatment, adolescent patients with cGVHD follow much the same treatment pathway as adults, cycling through various therapies in an attempt to control symptoms (Penack et al., 2020). Whilst there is a scarcity of published trials, what literature there is suggests similar responses to second- or subsequent-line treatments in adolescents, as in adults. Invariably with these later lines of therapy, complete response (CR) rates are low (Takahashi et al, 2021; Zeiser, 2015; Yang, 2021). However, partial response (PR) and individual organ CR rates (combined with reductions in steroid use) provide clinically meaningful benefit.

Belumosudil can reduce the long-term complications associated with steroid use in adolescents.

The treatment of cGVHD in adolescents must also include consideration of the possible impact any therapy will have on growth, nutrition, organ function, psychosocial functioning and immune reconstitution. As steroids remain the foundation of cGVHD therapy the consequences of long-term steroid use in children are well described and long-term deleterious effects on growth and bone density persist even after discontinuation of therapy (Ward, 2020). The beneficial effects of belumosudil as demonstrated in preclinical and clinical studies demonstrate that belumosudil is effective in treating cGVHD and in reducing corticosteroid therapy, which would benefit the patients, particularly adolescent patients.

Reflective of the response rate, and acknowledging that most patients were already receiving corticosteroids, there were significant reductions in corticosteroid dose (prednisone equivalent) in patients treated with belumosudil. Across the two studies (KD025-208 and

ROCKstar), approximately 64% of patients overall were able to reduce their corticosteroid dose during the study and 36 patients (19.4%) were able to discontinue the use of corticosteroids altogether. Belumosudil also supported tapering and discontinuation of other anti cGVHD systemic therapies such as calcineurin inhibitors (CNIs), MMF and sirolimus.

Taking this into consideration, the clinical response to belumosudil, both in terms of Overall Response Rate (ORR) and reduction in use of corticosteroids, is expected to be similar in adolescents as in adults, conferring a meaningful and significant clinical benefit to this underserved patient population.

No expected pharmacokinetic (PK) differences between adults and adolescents.

Population pharmacokinetic (popPK) analyses were conducted incorporating subjects from 7 clinical trials. A total of 174 healthy subjects and 178 patients with cGVHD were included in the analysis. The primary objectives of the study were to determine the effect of various covariates on the PK of belumosudil, and the model was devised as a 2-compartment model with first order absorption and lag time in absorption. This model demonstrated no effect of bodyweight on clearance down to 38.6kg and up to 143kg. Considering average bodyweight of UK boys and girls aged 12 years old is 38kg and 40kg respectively (50th percentile, UK-WHO, 2012), this supports extrapolation of the efficacy shown in adults by belumosudil to adolescent patients.

The existing popPK model was updated to include allometric scaling components on clearance and compartment volumes. The steady-state AUC0-24 for adults with a weight of 80.65 kg (median weight from the popPK model) and adolescent patients uniformly distributed between the ages of 12 and 18 years were simulated. The body weights for adolescents were sampled from the CDC age-weight chart and ranged from 45.9 kg to 74.2 kg. Both age groups consisted exclusively of 1000 cGVHD patients with a 1:1 male to female ratio. The simulated patients were administered a 200mg tablet of belumosudil either QD or twice daily (BID) and were not administered any concomitant medications. There is significant overlap between the AUCs of the adult and the adolescent population with the geometric mean of the steady-state AUC0-24 after QD administration in adults of 20500 ng*h/mL compared to 25900 ng*h/mL in adolescents.

Belumosudil exposure-response and exposure-safety relationships were examined across the range of Cmax values from 143-5780 ng/mL and range of AUC0-24 at steady state values from 2780-83800 h*ng/mL. Exposure-safety relationships were evaluated for

headache, fatigue, abnormal liver function, nausea and diarrhoea. Both exposure-response and exposure-safety relationships were flat, indicating a wide range of tolerable exposures, a lack of correlation between exposure and AEs, and a maximum effect at the lowest tested dose. These findings support the choice of dose of 200mg QD for adolescent patients.

In summary, the manageable safety profile, in particular in relation to adverse reactions of concern to adolescent patients, can reliably be expected to be the same in adolescents as in adults. This is due to the similarity of the disease pathophysiology, general response to treatment, PK modelling and flat exposure-safety relationship. In addition to this, the efficacy demonstrated in adults, not just the ORR but the reduction of steroid dose, indicates that adolescent patients would benefit from treatment with belumosudil. The choice of dose is supported by the popPK modelling showing large overlap between adult and adolescent AUCs, as well as the flat exposure-response and exposure-safety relationships. Therefore, the benefit:risk balance is positive for belumosudil in the treatment of cGVHD in adolescent patients at a dose of 200mg QD.

Answer to A2b

The starting dose of belumosudil in paediatric patients is proposed based on modelling and simulation, using the adult population PK model and allometric scaling of PK parameters based on body weight as described above. Drug metabolising enzymes reach maturity by 2 years of age and no differences in the ADME properties is expected in this paediatric population. The proposed dose is anticipated to match the exposure in adult patients with cGVHD following a daily dose of 200 mg belumosudil in tablet form. Considering the similarities in the pathophysiology and responses to therapy of cGVHD between adults and the paediatric patients, we expect that the proposed doses, based on modelling and simulation from adult patients, will produce comparable responses in adolescent paediatric participants.

Although there have not been any non-clinical studies of belumosudil in juvenile animals, existing clinical/non-clinical absorption, distribution, metabolism, and excretion (ADME)/PK and safety data would also support the low-risk nature of this dosing regimen in children aged 12 years and older.

Answer to A2c

A protocol amendment made to ROCKstar to enrol adolescent subjects was approved on 01 June 2020, and recruitment efforts continue. However, recruitment may be limited by the rarity of adolescent patients with cGvHD who have late-stage disease and have progressed

on multiple prior lines of therapy. Currently there are 2 adolescents patients enrolled in the ROCKstar study. The KD025-218 study is open for recruitment of adolescent patients but to date none have been enrolled.

A3. The SmPC states that strong CYP3A4 inducers and proton pump inhibitors (PPIs) decrease the exposure of belumosudil. Please substantiate the claim that patients with co-administered CYP3A4 inducers or PPIs need double the dose of belumosudil to obtain the same benefit?

The SmPC statement requiring a doubling of the dose of belumosudil when co-administered with strong CYP3A4 inducers or PPIs, is based on the systemic exposure of belumosudil being approximately half that of patients not taking strong CYP3A4 inducers or PPIs. This was not based on efficacy or safety data recorded in the belumosudil studies.

We conducted clinical drug-drug interaction (DDI) studies and population PK (PopPK) analyses to understand the effect of proton pump inhibitors (PPI) and CYP3A4 inducers on belumosudil systemic exposure.

***PPI*:** Approximately [redacted] of subjects in studies KD025-208 and ROCKstar were taking concomitant PPIs, which resulted in a decrease in exposure of both belumosudil and metabolites. Concomitant PPI use and the subsequent decrease in belumosudil exposure in these cGVHD patients did not affect efficacy or safety; however, there is potential for diminished efficacy. A DDI study was conducted (KD025-107) and demonstrated a >80% decrease in belumosudil exposure when co-administered with the strong proton pump inhibitor (PPI), rabeprazole and a 50% reduction in belumosudil exposure when co-administered with the moderate PPI inhibitor, omeprazole. Population PK analyses incorporated subjects from 7 clinical trials (174 healthy subjects and 178 patients with cGVHD) and revealed a significant PPI effect on relative bioavailability across healthy and cGVHD subjects, after accounting for variabilities in bioavailability by parameters such as food status, concomitant medications (CYP3A4 inducers and inhibitors) and disease. These analyses showed that concomitant PPI administration reduced the bioavailability of belumosudil by approximately 48% compared to no PPI administration.

CYP3A4 inducers: Similarly, data from one subject with cGVHD in the popPK subset and a DDI study to evaluate potential CYP3A4 induction (KD025-107) showed that when belumosudil was co-administered with the strong CYP3A inducer rifampicin, the C_{max} (maximum concentration) and overall exposure of belumosudil was reduced by ~60% and ~72% respectively. Moderate CYP3A4 inducer coadministration is predicted to decrease belumosudil C_{max} by 32% and AUC (area under the curve; exposure) by 35% in healthy

subjects. Population PK analyses estimated a significant effect of concomitant strong CYP3A4 inducers on apparent clearance (CL/F) of belumosudil by 2.88-fold.

To mitigate the risks of reduced belumosudil exposure and efficacy in patients taking concomitant PPI or strong CYP3A4 inducers, an increase in dose from 200mg QD to 200mg BID or 400mg QD was considered. Since the $AUC0_{-24,ss}$ (exposure at steady state) with a 200mg BID dosing regimen is closer to that of a 200mg QD dosing regimen than the 400mg QD dosing regimen, and to limit the potential for AEs, it was decided to increase the dose of belumosudil to 200mg BID in these patients. This dose prevents drug exposures from dropping below the $AUC0_{-24,ss}$ observed with 200mg QD dose administration in the pivotal studies and is expected to provide patients with the same benefits.

Indeed, while the study was not powered for this purpose, a subgroup analysis of overall response rate (ORR) in the ROCKstar mITT population showed no apparent correlation between ORR and concomitant PPI usage for either of the treatment arms (Table 5).

Table 5. Subgroup analysis of ORR for patients receiving concomitant PPI in ROCKstar (mITT, 2022 data cut)

Treatment arm	**200 mg QD (n=77)**		**200 mg BID (n=75)**	
Concomitant PPI?	Yes (n=39)	No (n=38)	Yes (n=35)	No (n=40)
ORR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
CR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
PR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

A4. Please fill the table below detailing the number of patients who co-administered CYP3A4 inducers and/or PPIs with belumosudil in each treatment arm in ROCKstar and KD025-208.

The SmPC requires the dose of belumosudil to be increased to 200mg twice daily only when co-administered with PPIs or *strong* CYP3A inducers. Concomitant CYP3A4 inducer co-administration data were not available from the ROCKstar and KD025-208 studies as this analysis was not conducted. However, since use of strong CYP3A4 inducers was prohibited in the ROCKstar protocol and other CYP3A4 inhibitors / inducers were to be used with caution, co-administration of strong CYP3A inducers for patients in trial is expected to be very low. Whilst we are unable to provide CYP3A usage data, co-administration data for PPIs is presented in the amended Table 6 below.

Table 6. Co-administration data for PPIs from pooled analysis (2022 data cut)

Trial / treatment arm	Co-administered Proton Pump Inhibitors while on-treatment, n (%)
KD025-208, ≥2 prior lines of therapy subgroup	
200 mg once daily (N=15)	[redacted]
200 mg twice daily (N=9)	[redacted]
400 mg once daily (N=14)	[redacted]
ROCKstar, August 2021 data cut	
200 mg once daily (N=66)	33 (50.0%)
200 mg twice daily (N=66)	33 (50.0%)
ROCKstar, 2022 data cut	
200 mg once daily (N=77)	[redacted]
200 mg twice daily (N=75)	[redacted]

A5. Please fill the table below detailing the baseline characteristics of the KD025-208 subgroups who had ≥2 prior lines of therapy.

Table 7. Baseline characteristics of KD025-208 subgroups who had ≥2 prior lines of therapy

Baseline characteristic	200 mg once daily (n=15)	200 mg twice daily (n=9)	Combined 200 mg (n=24)
Median age (range), years	57.0 (20, 63)	58.0 (42, 68)	57.0 (20, 68)
Males, n (%)	11 (73.3%)	5 (55.6%)	16 (66.7%)
HLA matching of donor/recipient, n (%) Matched Partially matched Unknown Missing	 13 (86.7%) 2 (13.3%) 0	 7 (77.8%) 2 (22.2%) 0	 20 (83.3%) 4 (16.7%) 0
Time from chronic GvHD diagnosis to enrolment, median (range), months	32.72 (6.5, 130.7)	25.13 (10.3, 69.9)	29.57 (6.5, 130.7)
NIH chronic GvHD severity,[a] n (%) Severe Moderate Mild	 10 (66.7%) 5 (33.3%) 0	 8 (88.9%) 1 (11.1%) 0	 18 (75.0%) 6 (25.0%) 0

Organ involvement, n (%)			
No. of organs involved, median (range)	4.0 (2, 6)	4.0 (1, 5)	4.0 (1, 6)
≥4 organs involved	9 (60.0%)	5 (55.6%)	14 (58.3%)
Eyes	12 (80.0%)	7 (77.8%)	19 (79.2%)
Skin	12 (80.0%)	6 (66.7%)	18 (75.0%)
Mouth	11 (73.3%)	5 (55.6%)	16 (66.7%)
Joints/fascia	11 (73.3%)	6 (66.7%)	17 (70.8%)
Lungs	5 (33.3%)	2 (22.2%)	7 (29.2%)
Upper GI	2 (13.3%)	1 (11.1%)	3 (12.5%)
Oesophagus	2 (13.3%)	0	2 (8.3%)
Lower GI	1 (6.7%)	0	1 (4.2%)
Liver	0	1 (11.1%)	1 (4.2%)
Prior therapy characteristics, n (%)			
Median prior LOTs, n	3	3	3
≥2 prior LOTs	15 (100.0%)	9 (100.0%)	24 (100.0%)
≥4 prior LOTs	2 (13.3%)	0	2 (8.3%)
≥6 prior LOTs	0	0	0
Refractory to prior LOT	10 (66.7%)	4 (44.4%)	14(58.3%)
Concomitant systemic chronic GvHD therapies, n (%)			
CS	15 (100.0%)	9 (100.0%)	24 (100.0%)
CNI	6 (40.0%)	4 (44.4%)	10 (41.7%)
ECP	4 (26.7%)	4 (44.4%)	8 (33.3%)

Outcomes

A6. Priority question: ROCKstar and KD025-208 define failure-free survival (FFS) as an absence of chronic GvHD treatment change, non-relapse mortality, and recurrent malignancy. REACH-3 defines FFS as time to relapse or recurrence of underlying disease, addition or initiation of new systemic treatment for chronic GvHD, or death due to underlying disease or non-relapse mortality, whichever came first.

a) Please clarify if ROCKstar and KD025-208 use mortality linked to relapse as a failure event within FFS.
b) REACH-3 specifies the addition of another systemic therapy for cGvHD as a failure event. Please clarify if the addition of a new systemic treatment for cGvHD was a failure event in the ROCKstar and KD025-208 analysis?
c) In ROCKstar and KD025-208, was the removal of a systemic therapy considered a failure event?
d) Please provide data on FFS using the same definition as REACH-3 to aid comparability.

Answer to A6a

Yes. Mortality linked to relapse would be captured as a failure event within FFS for both the ROCKstar and KD025-208 trials. ‘Mortality linked to relapse’ would occur as a result of the underlying malignancy recurring, and therefore would be categorised as ‘recurrent malignancy’. Recurrent malignancy included haematologic relapse, any unplanned intervention to prevent progression of malignancy or any other evidence of malignant disease after transplantation.

Answer to A6b

Yes. Our SAP states “*FFS is defined as the absence of **new** cGVHD systemic therapy, non-relapse mortality and recurrent malignancy* (i.e. underlying disease), and therefore any change to or introduction of new systemic therapy for cGVHD would be classed as a failure event.

Answer to A6c

No. The removal of a systemic therapy was not considered a failure event in the ROCKstar and KD025-208 studies. The removal or reduction in dose of certain systemic medications, namely, corticosteroids and calcineurin inhibitors, would be seen as positive outcome (i.e., a ‘steroid sparing’ effect).

Answer to A6d

The components of the FFS are the same in REACH-3 and the ROCKstar and KD025-208 trials.

We acknowledge there are slight differences in the wording for death as a failure event. In REACH-3 the definition is ‘*death due to underlying disease or non-relapse mortality'*. In the belumosudil studies ‘*Death to any reason*’ was included as a failure event. We believe for the purposes of the studies these can be considered comparable definitions and so we have not provided additional analysis to present an alternative view of the FFS events between the belumosudil and REACH-3 studies.

A7. Priority question: From the clinical study report (CSR) for ROCKstar, lack of response (LR) was defined as mixed (LR-M), unchanged (LR-U) or

progression (LR-P). Please explain why in the current ROCKstar/KD025-208 analysis, lack of response did not lead to a change of cGvHD treatment failure event for the FFS outcome.

According to the study protocols, subjects received belumosudil treatment until clinically significant progression of cGvHD (defined as progression that required the addition of a new systemic therapy for cGvHD, histologic recurrence of underlying malignancy, unacceptable toxicity, investigator decision, subject preference/withdrawn of consent, loss of follow-up, sponsor decision, or death [whichever occurred first]).

In this context a patient could be characterised as in LR but for the LR-M or LR-U states but their clinician may have considered them to be stable on treatment (or at least not getting worse or progressing). This may not warrant a change to their therapy. For patients at this stage in their journey, remaining without progression, whilst not a desirable outcome, could be considered to be at least acceptable.

Some participants who experience cGVHD progression as defined by NIH criteria may be considered to be LR-P but no new systemic therapy may be planned at that point and so they may continue to receive belumosudil and be assessed again at their next cycle.

In the model movement to the failure state was driven by the FFS outcome and whilst related, was not directly linked to progression status. The choice to switch to a new treatment was based on clinical judgement (See paragraph above).

A8. Priority question: The EAG's clinical experts stated that patients who did not respond to treatment would likely change systemic treatment and a change of systemic treatment would be a failure event for the FFS outcome. However, it appears a significant proportion of patients in the ROCKstar and KD025-208 trials have both LR and FFS.

a) **Please clarify if there were protocols used in the trial that specified a person's treatment pathway after LR?**

b) **Would a patient's treatment pathway be linked to whether their LR was a mixed (LR-M), unchanged (LR-U), or progression (LR-P) response?**

Answer to A8a

No specific protocols were used in belumosudil trials that specified a person's treatment pathway after LR. Participants who experience cGVHD progression as defined by NIH

criteria but for whom no new systemic therapy is planned may continue to receive belumosudil and be assessed again at their next cycle. (See answer to A7 above.) If progression according to the NIH criteria is not confirmed or no new systemic therapy is planned, participants may continue on belumosudil per investigator discretion until they fulfil one of the criteria requiring discontinuation of study drug. Subjects with a Lack of Response-Mixed (LR-M) response assessment may continue treatment with belumosudil. Subjects who have not achieved a response after 12 cycles of belumosudil should be withdrawn if in the Investigator's judgment there is no evidence of clinical benefit.

Answer to A8b

No. Same applied as above. This is typically a clinician decision as there are no clinical guidelines providing a concrete answer. Overall, lack of response at a certain timepoint would warrant a new therapy but this is based on the clinician's discretion and would be included as a failure event (hence would be captured in the economic model).

A9. Priority question: Overall partial response is defined in two different ways in the CS. Table 16 defines partial response as, "Improvement in at least 1 organ or site without progression in any other organ or site". Table 1 in Appendix M is the NIH definition of partial response and defines global partial response as, "Clinician overall severity score decreases by 2 or more points on a 0–10 scale".

a) **Please clarify if both criteria were utilised to categorise people as partial responders to belumosudil treatment?**

b) **If so, please provide full outcome data using each definition of partial response.**

c) **If only a single definition was used, please offer reasoning why this was chosen.**

d) **The REACH-3 study defined partial response according to 2014 NIH consensus criteria. If these criteria were not used in the ROCKstar and KD025-208, please justify why this is a fair comparison.**

e) **The EAG's clinical experts defined partial response as being strongly linked to improvement in a patient's "main organ". The main organ was**

decided on an individual basis but was more likely to be the liver, gut, or lungs, rather than skin. How generalisable is the trial definition of partial response to that used in clinical practice?

Answer to A9a

The 2014 National Institute of Health (NIH) Consensus Development Project includes both definitions of partial response (see https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4744804/). The NIH describe objective measures (*partial response is defined as improvement in at least one organ or site without progression in any other organ or site),* as well as the clinical assessment (c*linician overall severity score decreases by 2 or more points on a 0–10 scale),* used in combination to determine a partial response. The cGVHD clinician assessment sheet (Form A in the publication) shows that the objective organ scores are always accompanied by a subjective clinician severity assessment score between 0 and 10. These were the criteria used to categorise partial responders in the ROCKstar and KD025-208 trials.

Scoring of partial responses were done in combination and were not exclusive of one another as has been described above. Therefore, the figure for partial responses involves the investigator making a decision based on both the subjective and objective aspects of the response.

Answer to A9b

As both measurements were used in concert to categorise patients, we do not have analyses considering full outcome data using each definition of partial response separately.

Answer to A9c

Both measurements were used to categorise patients as discussed above.

Answer to A9d

Responses in the ROCKstar and KD025-208 studies were defined by the 2014 National Institute of Health (NIH) Consensus Development Project on clinical trials in cGVHD.

Answer to A9e

Whilst defining ‘partial response’ as linked to an improvement in the patient’s main organ would be a meaningful way to measure the response, it would also require subjective

assessment from the clinician as to which organ is classed as the ‘main’ organ. Due to the complexity of cGVHD and the variety of organs involved there is likely to be variation in the choice of the ‘main organ’ among clinicians.

According to the NIH criteria, total of 9 organs including skin, mouth, liver, upper and lower GI, oesophagus, lung, eye, and joint/fascia are considered when evaluating overall response and the weight for each organ in the analysis is considered equal. There is no definition of ‘main’ organ for response assessment in clinical practice or in our studies. (It is worth noting that the NIH criteria are well established, and are widely used in the clinical trial setting i.e. REACH-3).

In our January advisory board, clinical experts stated that the ROCKstar study population is generalisable to the UK population. Whilst the NIH criteria may not be used in day-to-day practice by clinicians, they will be familiar with its interpretation from UK trials.

Analysis

A10. Priority question: The ROCKstar study compares belumosudil 200 mg once daily to 200 mg twice daily and patients were randomised to each dosing regimen. In the CS, the company provides analysis where the treatment arms in ROCKstar and KD025-208 are “pooled” within their dosing regimen. This is the analysis used in the model. The “pooled” analysis, to which the majority of patients were randomly allocated, shows similar efficacy between the two dose arms for overall response rate (ORR), duration of response (DoR), FFS, and overall survival (OS). The EAG consider this to be credible evidence of once daily and twice daily dosing having the same efficacy for treating cGvHD, and as such, could be analysed together. The company also provides a “pooled and combined” analysis.

Additionally, the EAG’s clinical experts stated that it was unclear whether a comparison of the combined belumosudil arms with the control arm from REACH-3 was a conservative approach. There are baseline characteristics that suggest the REACH-3 best available treatment (BAT) population may have a better prognosis than ROCKstar and KD025-208, but also other factors, such as specific organ involvement, that indicate the REACH 3 arm is more complex to treat. Based on the “pooled and combined” analysis of ROCKstar and

KD025-208 (≥2 prior lines of therapy subgroup), and using the 2022 data cut (if available), please conduct a matching-adjusted indirect comparison (MAIC) for the following outcomes:

a) Overall response at week 24;

b) Failure-free survival at week 24;

c) Overall survival at week 24;

d) Duration of Response at week 24.

Recognising that the Phase II belumosudil studies were uncontrolled, our original intention was to identify evidence that could be used to conduct a MAIC. With this in mind, an SLR was performed which complied with the requirements for use within a NICE technology appraisal submission and in accordance with the Cochrane Handbook for Systematic Reviews of Interventions and the PRISMA guidelines. This is reported in the CS.

Published evidence is sparse and the only relevant study identified by the SLR was the REACH-3 clinical trial which evaluated the efficacy and safety of ruxolitinib compared with the investigator's choice of therapy in patients with moderate or severe glucocorticoid-refractory or -dependent chronic GVHD.

A feasibility assessment (FA) was done to determine whether the REACH-3 comparator arm composed of Best Available Therapy (BAT) could be considered for inclusion in an indirect treatment comparison (ITC) with belumosudil for CGVHD patients for whom two or more prior lines of therapy had failed. This FA aimed to determine how closely the two study populations matched one another and whether a population-adjusted indirect comparison (PAIC) could ultimately be achieved.

Unfortunately, critical differences in study inclusion criteria and populations prevented us from conducting a robust PAIC.

ROCKstar included subjects with cGvHD who had received 2 to 5 prior lines of therapy while in REACH-3 this cohort of patients were ineligible. Hence the population of exclusively second line patients studied in REACH-3 were less advanced along the treatment pathway compared to belumosudil eligible patients with a median of 3 prior lines of systemic therapy. This difference in the pathway characteristics of the patients is manifested in the higher proportion of severe patients treated with belumosudil ([redacted]) compared to BAT in REACH-3 (54.3%).

Table 8 below compares key risk factors between the REACH-3 and belumosudil pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) data. It clearly illustrates why conducting a matched adjusted comparison would be very difficult.

It is reported in the literature, that high-risk disease at transplantation, lower GI involvement at second-line treatment, and severe NIH global score at second-line treatment are associated with increased risks of treatment failure for second-line treatment. (Yoshihiro, 2013). The proportions of patients with severe cGvHD and lower GI involvement are higher in belumosudil treated patients than those in REACH-3. Importantly whilst organ involvement is reported for the BAT in REACH-3, the number of involved organs is not reported, and so a fair (matched) comparison cannot be made.

Table 8. Comparison of key prognostic baseline characteristics from REACH-3 and the pooled belumosudil studies (≥ 2 lines of prior therapy subgroup) (mITT, 2022 data cut).

cGVHD therapeutic	BAT (N=164)	Belumosudil 200mg once daily (N=92)	Belumosudil 200mg twice daily (N=84)
Previous aGVHD no.(%)	88 (53.7%)	[redacted]	[redacted]
cGVHD severity no.(%)	Moderate 74 (45.1%) Severe 89 (54.3%)	[redacted]	[redacted]
Median Line of prior Tx	Not reported	3	3
Number of involved organs	Not reported	[redacted]	[redacted]
Baseline organ involvement			
Skin	110 (67.1%)	[redacted]	[redacted]
Eye	93 (56.7%)	[redacted]	[redacted]
Mouth	99 (60.4%)	[redacted]	[redacted]
Esophagus	17 (10.4%)	[redacted]	[redacted]
Upper GI tract	21 (12.8%)	[redacted]	[redacted]
Lower GI tract	10 (6.1%)	[redacted]	[redacted]
Liver	83 (50.6%)	[redacted]	[redacted]
Lung	49 (29.9%)	[redacted]	[redacted]
Joint and fascia	44 (26.8%)	[redacted]	[redacted]

aGVHD=acute graft-versus-host disease; GI=gastro-intestinal; Tx=treatment

Whilst we acknowledge that the distribution of organ involvement in the REACH-3 study is different from that of ROCKstar, the current available data does not support the suggestion made by the clinical advisors to the EAG that patients at 2nd line in REACH-3 may be more burdened or be more complex to treat, than patients who might be encountered at 3rd line

and beyond. It is clear from Table 8 that organ involvement expressed as the proportion of patients is higher in the belumosudil studies.

Beyond these clinical considerations which suggest the REACH 3 patients are not more complex to treat, matching REACH-3 and the ROCKstar/KD025-208 (≥2 lines of prior systemic therapy subgroup) studies on organs affected would require potentially matching 9 organs combined with 4 levels of severity (0 to 3 score severity) amounting to over 200 000 combinations.

With these limitations in mind, we reported in the CS how we had discussed the relevance of the REACH-3 data with clinicians at our recent advisory board and in several face to face conversations with individual clinicians. The clinical experts agreed with us that a comparison with the REACH-3 BAT would likely be conservative due to the earlier treatment line but in the absence of other data it represented the best available evidence.

Results

A11. Priority question: Best response at any post-baseline assessment was reported in both ROCKstar/KD025-208 and REACH-3. However, REACH-3 reported the outcome up to assessments at week 24 and the combined ROCKstar and KD025-208 analysis included assessments after 24 weeks. The table below is adapted from Table 40 in the CS, please fill this table using the pooled, and the pooled and combined, analysis limited to best response up to week 24 assessments.

The best response up to the week 24 assessment point for the ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) studies are compared with the REACH-3 BAT in Table 9 below.

Table 9. Best response up to week 24 assessment for the ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) studies compared with the REACH-3 BAT (2022 data cut).

Week 24	ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup)			REACH-3
	Belumosudil 200 mg once daily (n=92)	Belumosudil 200 mg twice daily (n=84)	Belumosudil doses combined (n=176)	BAT (n=164)
Number of patients with response (%)	[redacted]	[redacted]	[redacted]	99 (60.4%)
CR (%)	[redacted]	[redacted]	[redacted]	11 (6.7%)

PR (%)	[redacted]	[redacted]	[redacted]	88 (53.7%)
Number of patients with no response (%)	[redacted]	[redacted]	[redacted]	65 (39.6%)

CR=complete response; PR=partial response.

A12. Priority question: Participants in ROCKstar and KD025-208, who were on-treatment, were assessed for response in 28-day treatment cycles. Based on the pooled and combined analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup), and using the 2022 data cut (if available), please provide details of the number of patients with FFS at defined timepoints during the study and the response or lack response to treatment within the FFS group. This is the response at the specified timepoint rather than best response at any prior timepoint.

The number of patients with FFS at six monthly timepoints during the studies and the response to treatment status within the FFS group are presented in Table 10 below for the pooled data and for the once and twice daily cohorts in Table 11 and Table 12.

Table 10. Pooled and combined ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) (2022 data cut)

	Pooled and combined ROCKstar and KD025-208 belumosudil (N=176)						
	Patients with FFS, n (%)	**Responders, n (%)**	**CR, n (%)**	**PR, n (%)**	**LR-U, n (%)**	**LR-M, n (%)**	**LR-P, n (%)**
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 11. Pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) belumosudil 200 mg once daily (2022 data cut)

	Pooled ROCKstar and KD025-208 belumosudil 200 mg once daily (n=92)						
	Patients with FFS, n (%)	Responders, n (%)	CR, n (%)	PR, n (%)	LR-U, n (%)	LR-M, n (%)	LR-P, n (%)
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 12. Pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) belumosudil 200 mg twice daily (2022 data cut)

	Pooled ROCKstar and KD025-208 belumosudil 200 mg twice daily (n=84)						
	Patients with FFS, n (%)	Responders, n (%)	CR, n (%)	PR, n (%)	LR-U, n (%)	LR-M, n (%)	LR-P, n (%)
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

A13. Please provide cumulative response rate as presented in Figure 12 in the CS, based on both the pooled and the pooled and combined analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup), and using the 2022 data cut (if available).

An updated version of figure 12 from the CS is presented in Figure 4 below.

Figure 4

A14. Priority question: Based on the pooled and combined analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup), and using the 2022 data cut (if available), please provide details of the response to belumosudil by organ system.

The response from baseline of organ involvement for the pooled and combined ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) is presented in Table 13 to Table 15 below and overleaf.

Table 13. Response from baseline of organ involvement in the pooled and combined ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) (2022 data cut).

Organ	Pooled and combined ROCKstar and KD025-208 (n=176)				
	Number affected, n	Responders, n (%)	CR, n (%)	PR, n (%)	LR, n (%)
Skin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Joints/fascia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eyes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Mouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lungs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oesophagus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upper GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lower GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Liver	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; GI=gastro-intestinal; PR=partial response; LR =lack of response

Table 14. Response from baseline of organ involvement in the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) belumosudil 200 mg once daily (2022 data cut).

Organ	Pooled ROCKstar and KD025-208 belumosudil 200 mg once daily (n=92)				
	Number affected, n	Responders, n (%)	CR, n (%)	PR, n (%)	LR, n (%)
Skin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Joints/fascia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eyes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lungs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oesophagus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upper GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lower GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Liver	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; GI=gastro-intestinal; PR=partial response; LR =lack of response

Table 15. Response from baseline of organ involvement in the pooled ROCKstar and KD025-208 ((≥2 prior lines of therapy subgroup) belumosudil 200 mg twice daily (2022 data cut).

Organ	Pooled ROCKstar and KD025-208 belumosudil 200 mg twice daily (n=84)				
	Number affected, n	Responders, n (%)	CR, n (%)	PR, n (%)	LR, n (%)
Skin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Joints/fascia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eyes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Lungs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oesophagus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upper GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lower GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Liver	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; GI=gastro-intestinal; PR=partial response; LR =lack of response

[A15. Please provide data from the ROCKstar and KD025-208 analysis (≤2 prior lines of systemic treatment subgroup) for the distribution of failure events for complete and partial responders and for patients who have had a lack of response?

The distribution of failure events for complete and partial responders is provided in Table 16 to Table 18 below and overleaf.

Table 16. Failure events in the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup, 2022 data cut).

Time point	Pooled and combined ROCKstar and KD025-208 belumosudil 200 mg (N=176)						
	Patients without FFS, n (%)	Responders, n (%)	CR, n (%)	PR, n (%)	LR-U, n (%)	LR-M, n (%)	LR-P, n (%)
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 17 Failure events in the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup; belumosudil 200 mg once daily, 2022 data cut).

	Pooled and combined ROCKstar and KD025-208 belumosudil 200 mg once daily (N=92)

Time point	Patients without FFS, n (%)	Responders, n (%)	CR, n (%)	PR, n (%)	LR-U, n (%)	LR-M, n (%)	LR-P, n (%)
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 18. Failure events in the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup; belumosudil 200 mg twice daily, 2022 data cut).

Time point	Pooled and combined ROCKstar and KD025-208 belumosudil 200 mg twice daily (N=84)						
	Patients without FFS, n (%)	Responders, n (%)	CR, n (%)	PR, n (%)	LR-U, n (%)	LR-M, n (%)	LR-P, n (%)
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
30 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
42 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
48 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Interpretation

A16. Does the company consider belumosudil to be a life-extending drug for people with cGvHD and, if so, what is the clinical rationale for this?

Belumosudil has a unique mode of action (selective ROCK2 inhibition) which works through rebalancing of the immune response, reducing inflammation and arresting fibrosis. It therefore improves or stabilises disease. (It is important to note that no other available treatments target **both** the inflammatory and fibrotic pathogenic pathways in cGVHD).

It may be the case that these benefits realised whilst on treatment contribute to an overall extension in survival in the longer term (as well as improving quality of life in the short term). However, belumosudil is not a therapy specifically indicated for the treatment of haematologic cancers or directed against recurrence of malignancy and there is no strong rationale to consider it to be a life-extending drug for people with cGvHD beyond the implied impact that delays to fibrotic or inflammatory processes may have.

Nonetheless naive side by side comparison of belumosudil and BAT arm from REACH-3, suggests an overall survival benefit with belumosudil over BAT albeit in different lines of treatment. This advantage is reflected in the modelled outcomes (See question B5 below) and was also observed in the ECA data where a survival benefit over BAT from real-world was observed after adjustment.

In order to reflect these observations, we have taken a somewhat conservative approach to the modelling, recognising that the long-term relative mortality risk once off treatment is uncertain. We have constrained the model to assume that after 5 years (when all patients remaining in the FFS state are assumed to be off initial treatment) mortality risk is the same in each arm (i.e., same cycle probability of death for the belumosudil arm as the BAT arm post-5 years). The lower risk prior to this point in the belumosudil arm does result in longer overall survival but this is significantly curtailed by the switch to BAT risk. This assumption of equalised mortality risk after 5 years was validated by clinicians at our advisory board and is tested in a sensitivity analysis.

Section B: Clarification on cost-effectiveness data

For any scenarios requested in Section B, please ensure these are implemented as user selectable options in the economic model ("Settings" tab). If scenarios cannot be implemented as user selectable options, please supply instructions on how to replicate the scenario. Furthermore, if the company chooses to update its base case analysis, please ensure that cost-effectiveness results, sensitivity and scenario analyses incorporating the revised base case assumptions are provided with the response along with a log of changes made to the company base case.

We have provided answers based on the original company base case in the following sections. However, we have also carefully considered all of the helpful comments made by the EAG and present an alternative scenario in Appendix B which includes plausible adjustments to the original analysis.

Model approach

B1. Priority question: In Appendix N, the company assessed the proportional hazards (PH) assumptions for trial arms within ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) and for REACH-3 separately. However, given the company has assumed a naive comparison of the pooled data from ROCKstar and KD025-208 and REACH-3, the EAG considers that the PH assumption for belumosudil and BAT could have been assessed based on their Kaplan-Meier (KM) curves directly. As such, please provide an assessment of the PH assumption for belumosudil and BAT based on a comparison of the KM curves from the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup and REACH-3.

The original CS did not use hazard ratios (HRs) to generate the extrapolations for the BAT treatment. Rather, parametric fits directly estimated on reconstructed individual patient-level data obtained from the REACH-3 trial using the Guyot et al. algorithm were implemented for BAT, for OS, FFS and DOR. As a result, no PH assumption testing between the belumosudil arms and the BAT arm were conducted as part of the original CS.

As requested, KM curves and log-cumulative hazard as a function of logarithm of time plots for the belumosudil pooled arms using data from ROCKStar and the KD025-208 studies (≥2

prior lines of therapy subgroup) using the 30th of September 2022 data cut and the BAT arm from the REACH-3 trial are presented in Figure 5. P-values from the Grambsch and Therneau test for PH assumption are provided in Table 19.

Concerns can be raised about the PH assumption in the comparison of belumosudil vs BAT with respect to FFS and so we believe that the use of a HR to summarise the relative efficacy of belumosudil vs BAT on FFS is not appropriate. It is important to note that in the CEM analyses presented as part of the CS the survival extrapolations for BAT were conducted by fitting parametric models directly to the BAT data and were not generated through the use of a HR.

Figure 5 [redacted]

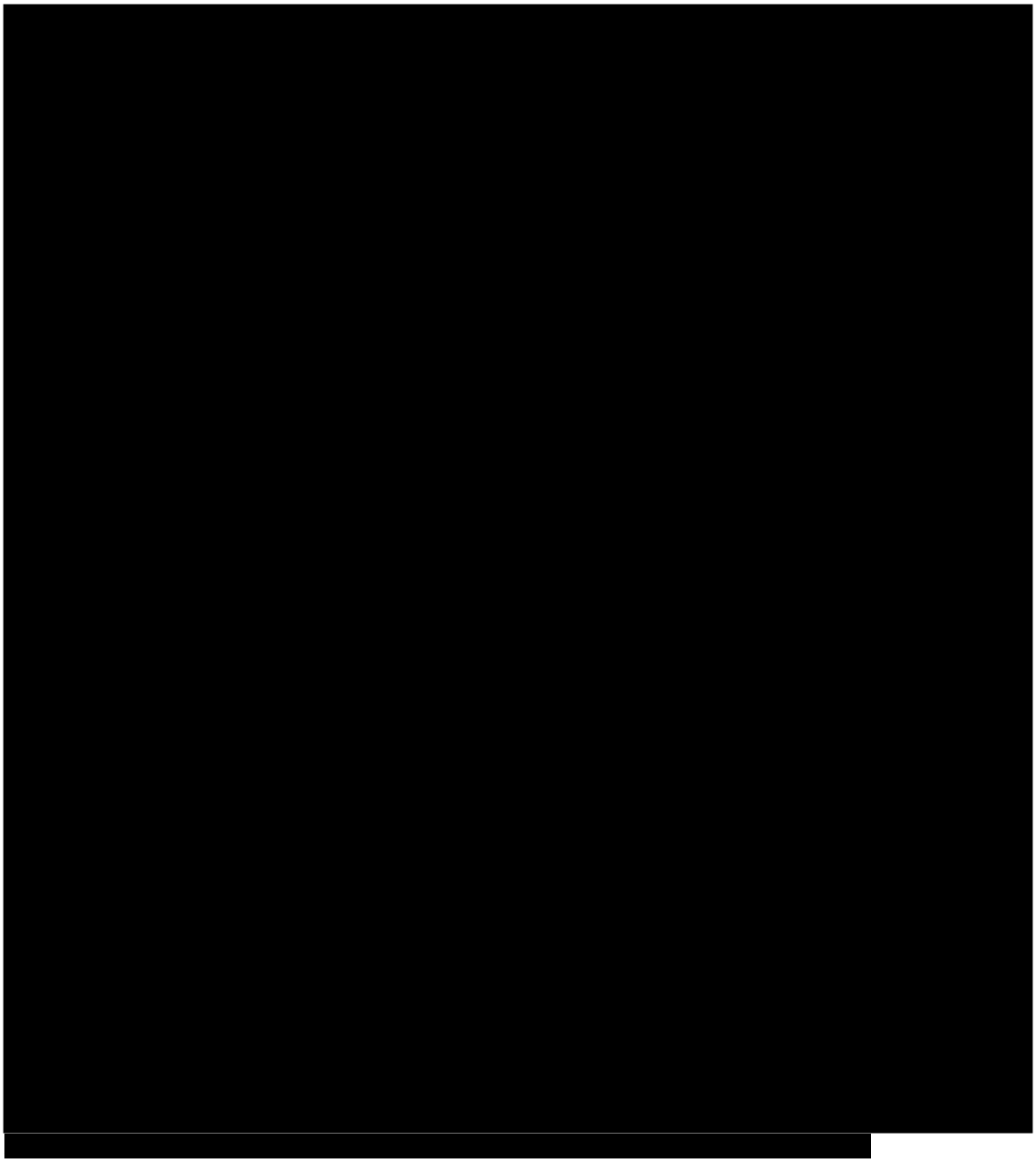


Table 19. P-values for the Grambsch and Therneau tests for the analyses of belumosudil vs BAT on OS, FFS and DOR

	p-Value of Grambsch and Therneau test	
	Time	**Logarithm of time**
OS	[redacted]	[redacted]
FFS	[redacted]	[redacted]
DOR	[redacted]	[redacted]

Abbreviations: DOR = duration of response; FFS = failure-free survival; OS = overall survival

B2. Priority question: Please provide the underlying cumulative hazard function plot for FFS and OS in the model and explain if the base case extrapolations reflect the underlying hazards.

The observed cumulative hazards plot as a function of time using the KM estimates (solid lines) are presented along parametric survival predictions of cumulative hazard (dashed lines) for OS and FFS in the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) studies for belumosudil 200 mg BID and QD and in the REACH-3 study for BAT in Figure 6. Predicted cumulative hazards for OS were obtained using the exponential distribution, while for FFS the generalised gamma distribution was used. 30th of September 2022 data cut was used for the derivation of the belumosudil estimates.

As discussed in the original CS, because of the immaturity of the OS data in both the ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) studies and the REACH-3 trials no definite conclusion can be made on whether extrapolation accurately captures the shape of the true underlying hazards. None of the parametric survival distributions allowed capture of the shape of the FFS curve in the BAT arm.

Figure 6. [redacted]

B3 Priority question: Please clarify why response or FFS outcomes are not modelled for patients in the failure health state whose failure event was initiation of a new cGvHD systemic treatment?

In order to make best use of the available data we have implemented a three-state partitioned survival structure for our model in both a deterministic and probabilistic framework. As is usual with such 3-state models, the 'progressive' state is not further subdivided into tunnel states or split to provide a sequencing structure.

The large majority of entrants into this state are patients whose failure event was initiation of a new chronic GVHD systemic therapy. However, we have differentiated within the state between these patients and those with recurrent malignancy by including different costs of subsequent lines of therapy (acquisition and administration costs) and quality of life outcomes for each group.

Hence this structure does reflect the progressive nature of chronic GVHD because it includes both the switch to new therapy and the potential relapse of malignancy with differentiated payoffs.

The time in the failure state is the difference between the occupancy of FFS and OS. Whilst this is a simplifying calculation it does account for the remaining time until death and with the inclusion of differentiated payoffs is the most appropriate approach to the data available. This means that explicit modelling of FFS in the failure state is neither needed nor would provide more certainty in the model results.

However, to stratify further within the failure state by response (subsequent FFS or failure) is actually not possible within the confines of the available data for a number of reasons:

- The comparator REACH-3 data for FFS post subsequent therapy has not been published. Hence, this approach for the BAT arm cannot be implemented.
- The response/benefit of the new cGVHD treatment are implicitly captured through the OS derived from clinical trials for both belumosudil and BAT. Explicit modelling of response or FFS outcomes due to the new cGVHD systemic treatment may result in some double counting.
- Notwithstanding this paucity of information to support a single split to the failure state, the belumosudil data comprises patients at 3rd line treatment and beyond. To properly implement a structure which accommodates FFS and failure in a post switch state would require sequences of such states to account for 4th, 5th, 6th and so on, lines of therapy. Given the extant data, this level of granularity would necessitate a large number of strong assumptions and would likely introduce significant uncertainty into the modelled outcomes.

We hope the EAG will recall our conversation at the clarification call where this was discussed and recognise that modelling of FFS and failure states for the post switch patient is not possible to fulfil.

B4. Priority question: The EAG considers that the failure health state should be two separate health states as patients who enter the failure health state because of recurrent malignancy are likely to have different survival outcomes compared to patients without recurrent malignancy but have changed cGvHD treatment. As such, please provide a scenario where each type of failure event is a separate health state with overall survival stratified by failure type. Additionally for those patients whose failure event was initiation of a new cGvHD systemic treatment, please consider including FFS (and response outcomes, if available).

The overwhelming number of failure events in ROCKstar or KD025-208 (≥2 prior lines of therapy subgroup) are due to switch to a different treatment. The absolute number of recurrent malignancies in the pooled dataset which trigger a failure event is very low ([redacted] in the 2022 data cut). This means that estimates for time to entry into a differentiated health state made on the basis of recurrent malignancy alone will be uncertain.

Due to the low number of recurrent malignancy events, the OS data for this group of patients is also very limited; only [redacted] of the [redacted] deaths reported in ROCKstar had underlying (recurrent) malignancy as the primary cause of death. Using the OS data would therefore result in a very uncertain choice of OS curve fit (or even fitting not being possible) and the

introduction of further uncertainty. (It is worth noting that OS is estimated from baseline and not from the failure event.)

The equivalent data required for the analysis is not published for the BAT from REACH-3. Indeed, similar considerations to those above exist as well for REACH-3. For example, only ~5% of patients experienced recurrent malignancy in REACH-3.

We hope the EAG is satisfied that stratification by failure event is not possible with the available data from our studies or with the comparator data. However, we do wish to reiterate that differentiation within the failure state between patients with different failure events is included via the implementation of different costs and quality of life outcomes for each group.

Overall survival

B5. Priority question: The EAG's clinical experts considered that the proportion of time patients spend in the failure health state (as measured by life years [LYs]) is not clinically plausible. Specifically, they considered that:

- **For BAT patients whose failure event is a new cGvHD treatment, it is not clinically plausible that their mean survival in this health state would be [redacted]. The EAG's clinical experts considered that this group of patients are representative of those with progressive or refractory cGvHD and mortality would be higher than predicted in the model.**

- **In the belumosudil arm, patients spend only slightly longer in the failure free health state as the failure state ([redacted] and [redacted] respectively). As with the BAT patients, the EAG's clinical experts considered that patients whose failure event was a new cGvHD treatment, it is likely they have progressive or refractory cGvHD and mortality would be higher than predicted in the model.**

In ROCKstar, patients had multiple lines of treatment (49% of patients had four or more lines of prior therapy). Thus, the EAG considers that analysis of overall survival for patients who are on their fourth line (or later) systemic treatment is feasible.

As such, please provide an analysis comparing survival outcomes for patients at their third line of treatment with patients who have had four or more lines of treatment and use these analyses to provide a scenario where overall survival in the model is adjusted to produce more clinically valid estimates of survival for patients in the failure health state (or the new health state for patients who have failed due to starting a new cGvHD treatment, as recommended in B4).

We have implemented the 2022 data in the economic model and the modelled total life years and time in the failure-free and failure states for the updated base case is compared with the original, unchanged estimates for BAT from the CS in Table 20 below.

Table 20. Overall survival and time in states for the pooled belumosudil efficacy analysis (2022 data cut) and BAT (REACH-3) treated patients (Discounted outcomes).

	Total life years*	Mean time in Failure-free	Mean time in Failure
Belumosudil	[redacted]	[redacted]	[redacted]
BAT**	[redacted]	[redacted]	[redacted]

*Median OS was not reached in either the belumosudil trials nor REACH-3, which had maximum durations of follow up of 5.4 and 2.4 years, respectively. **REACH-3, unchanged from original CS.

Inspection of Table 20 shows that implementation of the updated data set provides similar results to the original model.

We have discussed earlier in our response to Question A16 how belumosudil is not explicitly expected to be a life extending therapy but the impact of reduction on fibrosis and inflammation whilst on treatment may be implicated in survival benefit. This is evident when a naïve comparison of the studies is made (bearing in mind that the REACH-3 cohort was treated at 2nd line and those patients might be therefore expected to live for a shorter time from 3rd line onwards). However, we have taken a parsimonious approach to the OS extrapolation for the belumosudil arm. The shorter time in the failure state for belumosudil in the model base case relative to BAT is a consequence of the adoption of the BAT OS risk post 5 years for belumosudil patients. This recognises the lack of long-term evidence and attempts to normalise between the 2 arms by reducing the OS for belumosudil treated patients relative to BAT.

It is important to recognise that the nature of a 3-state partitioned survival model is not to explicitly model each partition separately but rather to estimate overall survival and capture the time in the failure state as the difference, in this case, between the FFS and OS. Following this methodology, we have estimated OS based on the best available data from

the studies and in both cases, it is around [redacted] years. In this regard we believe that we have accurately modelled the overall lifespan of patients who may be eligible for belumosudil or BAT.

We have not calculated OS as the sum of OS in the FF and failure states. To do so would be technically difficult given the data and would likely produce very similar results to those we have observed. The difference in the modelled time in state for the two arms springs from the data. We present some precedent from the literature and clinical opinion below to support the potential lifespan of patients with cGVHD treated in current clinical practice.

Our approach to the survival analysis was discussed with clinical experts – see below.

Validation of the fitted curves.

Little external data is available to help validate the parametric extrapolations used in the economic model and so we consulted with more than 10 clinicians and two health economists in several individual conversations and two advisory boards.

We have noted in the CS that the fits implemented in the base case were selected based on the lowest AIC and BIC fit statistics. We believe the choice of the extrapolations for all the fitted data used in the base case is robust. This is because they were based on the lowest AIC/BIC combinations and in most cases the data sets were more than 50% complete providing confidence that the fit statistic information criteria can be used reliably.

However, the OS datasets are less complete (see the discussion of literature precedent below) and the plausibility of these fits was discussed with clinicians. For example, this was accomplished by the provision of stimulus material at the virtual advisory board held between 19th to the 27th of January 2023 in which 9 clinicians and 1 health economist were consulted. The KM data were presented overlaid with the parametric fitted curves. The advisors were asked to consider a landmark analysis for each curve at 5 and 10 years and they told us that they found the landmarks a helpful way to consider which curve to pick based on their experience of patients in their clinical practice.

There were several perspectives on the choice of the key survival curves. For OS, the respondents agreed that the exponential fit which provided the lowest survival should be used but they did discuss whether the fit should be more punitive because at this point in the pathway the expectation is for poor survival. However, they noted that patients on belumosudil would be expected to have a lower risk whilst on treatment compared with other

patients on BAT. On this basis the exponential model was recommended by the advisors with no downward adjustment whilst on treatment.

The clinical advisors agreed that longer term risk once off treatment would revert to BAT levels and so our approach to normalising the risk was agreed.

Recognising the comments, the EAG received during their consultation we have looked again at our survival estimates generated by the model and considered other ways to validate them.

Survival data from the pooled analysis of the 2022 belumosudil data

The AIC / BIC data were not disclosed before the discussions. The updated 2022 data cut for OS does not differ in the extrapolation between the previous and new fits and so there is no reason to believe that the reasons for choosing the original fit would change in this new iteration. See Figure 7 below.

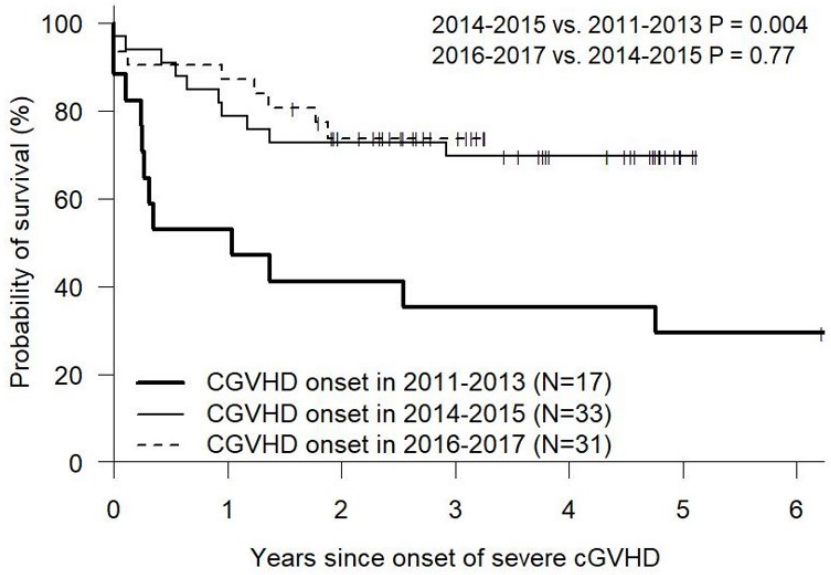
Figure 7. [redacted]

There were very few mortality events over the follow-up period for the belumosudil studies. [redacted] deaths in total were recorded ([redacted] in ROCKstar, 10 in KD025-208). [redacted] were in patients entering the study at 3rd line only and [redacted] in patients at 4th line and beyond. (An additional death was recorded in the 400 mg arm of the study but his is not within the licence dose). The small number of deaths observed means that any long-term estimates for OS split by line would be highly uncertain. Therefore, we do not believe this analysis would be informative and have not carried it out.

Literature precedent

Precedent exists in the literature to suggest that there has been a trend towards improved probability of survival for cGVHD patients over the last 10 to 15 years. In a recent study from the US evaluating the survival of patients diagnosed with severe aGVHD or cGVHD compared to historic controls the authors found that after 5 years of follow up since diagnosis of severe cGVHD (which would likely trigger systemic treatment and could be considered 2nd or 3rd line) median OS was not reached and after 2 years, more than two thirds of patients remained alive. (Bashey, 2019) (Figure 8.

Figure 8. probability of survival after the onset of cGVHD for cohorts diagnosed over the last years. (Reproduced from Bashey 2019)



In their article Bashey et al state that… '*This [*trend towards longer survival*] must be taken into account when evaluating novel therapies for severe GVHD*.' This is important when considering clinical opinion gathered for this appraisal.

A similar pattern of survival at 3 years (65.5% remaining alive) was observed in recently published retrospective health claims data base in Germany. (Scheid, 2022).

These findings concur with the data from the REACH 3 study in which there was a total of 27 deaths in the BAT arm of the REACH-3 study over a maximum follow-up duration of approximately 2.4 years. This represents a probability of survival at the end of the follow up of around 75%.

Our modelling is not inconsistent with the published data and the observation of increased survival for patients over the last 10 – 15 years. At 2 years the OS rate pooling the data from ROCKstar was [redacted] and in the overall pool of ROCKstar and KD025-208 it was [redacted] using the 30th of September 2022 data cut in both cases. Using the most conservative fit for OS (Exponential) the model estimates a mean of [redacted] total life years with current standard of care and up to mean [redacted] years with belumosudil treatment could be achieved. This may reflect continued improvements in survival with new treatments and better standard of care. This assumption was tested with a clinician we spoke to at the EBMT conference (Paris, 23-25/04/2023) who told us that expectations for survival are much longer today than they were even 5 years ago.

Treatment discontinuation

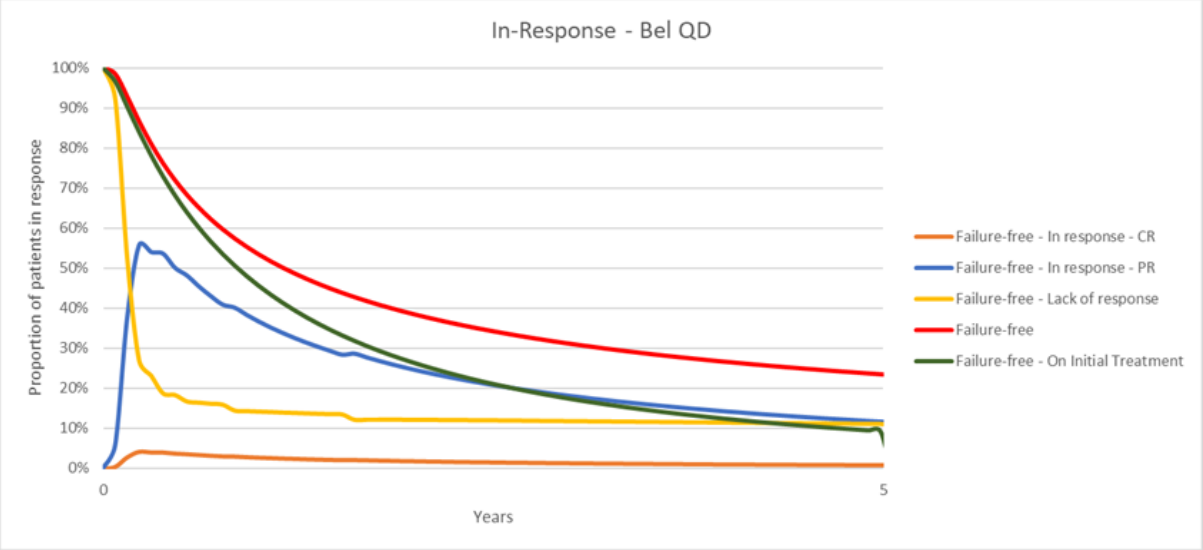
B6. Priority question: The EAG's clinical experts considered that for those patients who had a lack of response, they would not continue on their current treatment, but instead would be given a new treatment (which would be classed as a failure event in the definition of FFS). However, in the model, a significant proportion of patients have a lack of response to treatment, yet remain on treatment (as estimated by the time to treatment discontinuation

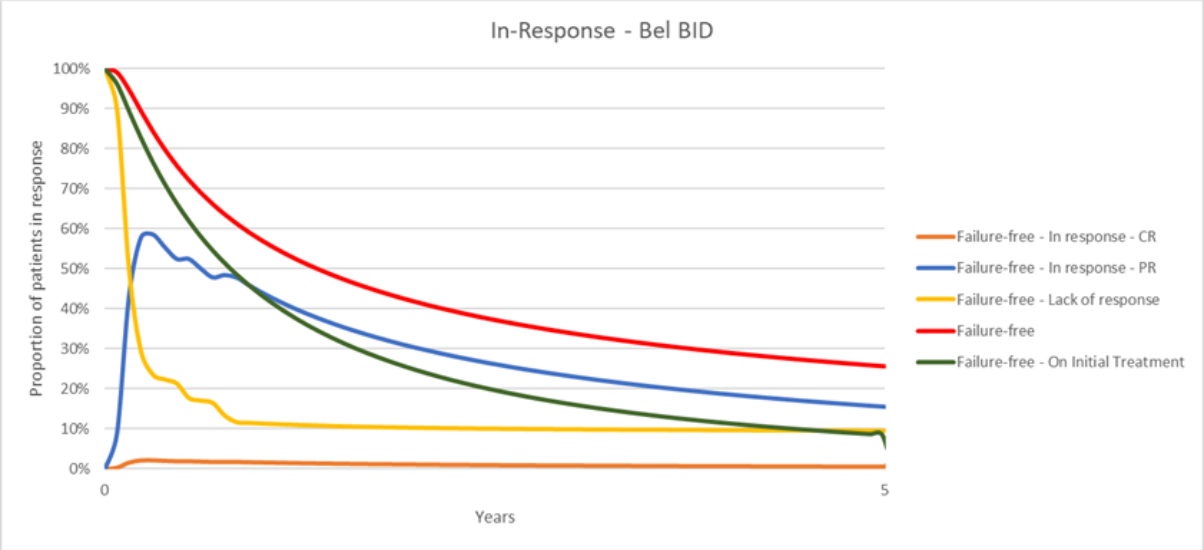
[TTD] curve) and failure-free (as estimated by the FFS curve). Please see the plots below.

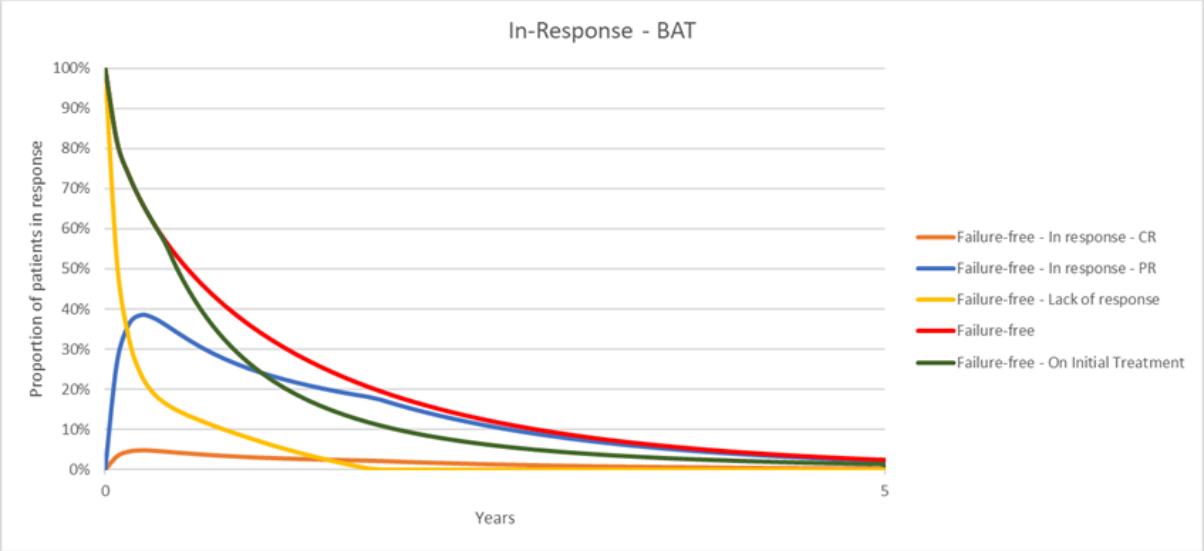
a) **Please discuss the clinical validity of patients remaining in the failure-free health state and on treatment even though they have a lack of response for up to five years.**

b) **Please clarify why lack of response in the model doesn't lead to a change in cGvHD treatment?**

c) **Please discuss the clinical plausibility of patients with a lack of response having a higher quality of life than patients who have started a new cGvHD treatment, when both have not had a recurrence in their malignancy?**







Answer to B6a

We have discussed lack of response in our answers to questions A7 and A8. A lack of response would not necessarily result in a patient being initiated on a new treatment. ‘Lack of response’ (LR) included three categories: mixed, unchanged and progression. Due to the potential severity of cGHVD symptoms after two lines of systemic therapy, a patient being stable and not deteriorating further may be a good reason to keep patients on their current therapy. If the patient deteriorates and falls into the LR-P category, a new treatment could be initiated, and a failure event would be recorded. However, this is at the clinician’s discretion and new treatment may not be initiated immediately. Hence some patients can persist in the FFS with LR-P until the choice is made to initiate new treatment. Due to the limited number

of drugs being available to these patients, if a patient is stable and not deteriorating, they may remain on current treatment, and hence the ‘failure-free’ health state for several years.

Answer to B6b

Please see B6a and A7 and A8 above.

Answer to B6c

Patients who are in the ‘lack of response’ health state may be considered to be in a stable condition. Whilst their cGVHD symptoms are not improving, they are also not deteriorating and, whilst not optimal this may be an acceptable outcome.

The condition of patients who meet the criteria for initiation of a new systemic cGVHD therapy is likely to be poor. Self-evidently they will have deteriorated significantly enough that they have been assessed with current failure of treatment and this deterioration will necessarily lead to a lower QoL score than stable patients. We discussed the move to another treatment option (failure in the model) with clinical experts and they pointed out that the failure of a treatment line necessitating move to a different therapy is often accompanied by a high rate of infections, hospitalisation, reduction in performance status and significant concerns about the risk of death. All of these factors combine to reduce a patient’s quality of life significantly.

It is also important to note that the utility value attracted by the failure state in the model must represent the average HRQoL over the remaining patient journey, from the point of switching 3L+ treatment until death. As a progressive disease with poor outcomes associated with later lines of treatment, it is anticipated that the utility score would decline over time. Clinicians noted that many patients would be initiated on palliative care within this health state when all treatment options had failed.

This progressive decline is reflected in the multi-national observational study of cGVHD patients described in the company submission (Adelphi cGVHD Disease Specific Programme, Data on file). The mean patient-reported EQ-5D was 0.71 for patients on first line of cGVHD therapy (n=74), decreasing to 0.67 and 0.60 for those on second (n=51) and third line or more (n=30), respectively. Although this represents a cohort of patients in an earlier position of therapy, the trend of decreasing HRQoL with progressive treatment failure is evident.

The model is designed to capture all patients with 3rd and later lines of therapy so it must be noted that much more burdened patients at 4L and beyond are included in the estimate.

Therefore, it is therefore critical not to assume that even if the move to a subsequent therapy provided some efficacy benefit resulting in an upswing in QoL that this would persist for the whole of the time spent in the failure health state.

B7. Priority question: The EAG's clinical experts explained that for patients with a complete or partial response, they are likely to be weaned off treatment over time.

a) **Please explain if patients on belumosudil would be weaned off treatment over time and if so, does TTD in the model capture treatment weaning?**

b) **If TTD doesn't capture treatment weaning, please explore a scenario in the model where assumptions around treatment weaning are included for the complete response (CR) and partial response (PR) patients.**

Answer to B7a.

Patients who achieved a complete or partial response would be likely to have their treatment weaned once stable after a period of time. Clinicians in our advisory board stated that it was standard practice to wean patients off any systemic cGVHD treatment if the patient had responded and was stable, to avoid unnecessary exposure.

The ROCKstar study protocol included the following tapering guidance:

Belumosudil was tapered after a sustained response for 6 months and cessation of all other immunosuppressants for at least 3 months. The tapering schedule for belumosudil was as follows:

- *Arm A: 200 mg QD → 200 mg once every other day (QOD) for 2 cycles → discontinue; and*
- *Arm B: 200 mg BID → 200 mg QD for 2 cycles → 200 mg QOD for 2 cycles → discontinue.*

Similarly, subjects whose cGVHD had not progressed or responded at the time of discontinuation of belumosudil treatment and who came off the study for reasons other than

AEs could be tapered off belumosudil by reducing the dose every 2 cycles as described above.

TTD in the model does not capture treatment weaning. It represents the time on treatment regardless of dose changes. This means the costs of belumosudil treatment may be overestimated in the base case results where 100% relative dose intensity (RDI) was used in order to be conservative due to the lack of information about weaning in the studies.

Data describing the number of patients in the clinical studies who were weaned off belumosudil are not available. However, this may be captured to some extent by the RDI of belumosudil ([redacted]% for the once daily arm and [redacted]% for the twice daily arm [ROCKstar 2022 data cut]). This is described in more detail below in B7b.

Answer to B7b

According to the protocol above, patients on QD belumosudil who are eligible for tapering would receive 50% of the total treatment exposure for the final two cycles, whilst those on BID dosing would receive 37.5% of treatment exposure for the final four cycles.

Although we have not provided a pharmacoeconomic scenario analysis with these estimates as this would represent an extreme scenario and is not reflective of the actual use of belumosudil in the studies (which may be expected to drive efficacy), we have considered how this might impact real world use if all responding patients were weaned according to the protocol.

The impact of treatment weaning on RDI is illustrated below (Table 21). This estimates the RDI assuming that patients who responded to belumosudil (CR or PR) would be weaned according to the tapering protocol outlined above.

Table 21. Calculation of RDI for belumosudil tapering

	QD	BID
Proportion of responders (QD)	[redacted]	[redacted]
Median duration of treatment, months (2022 pooled analysis)	[redacted]	[redacted]
Median duration of treatment, days (2022 pooled analysis)	[redacted]	[redacted]
RDI before tapering	[redacted]	[redacted]
RDI during tapering	[redacted]	[redacted]
Tapering duration (days)	[redacted]	[redacted]
Overall RDI of tapered (responder) patient	[redacted]	[redacted]
Overall RDI of non-tapered (non-responder) patient	[redacted]	[redacted]
Weighted overall RDI (total population)	[redacted]	[redacted]

Data sourced from 2022 pooled analysis of patients on ≥2 prior therapies. RDI=relative dose intensity.

To better reflect the impact of tapering (and potentially reduction in dose intensity for other unknown reasons) from the clinical studies, we conducted a scenario analysis using the average overall RDI from ROCKstar (2022 data cut), which was [redacted]% for the once daily (QD) arm and [redacted]% for the twice daily (BID) arm.

The results from this analysis are provided below in Table 22.

This analysis is incorporated into the updated scenario presented in Appendix B.

Table 22. Deterministic scenario analysis with adjusted belumosudil dose intensity (PAS, WITH and WITHOUT severity modifier)

	ICER (£/QALY)		Change from base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£1,598	£1,332	-55.25%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To adjust the dose intensity for belumosudil: Go to **[Costs]** sheet > Change cell **H58** from 100% to [redacted]% and cell **H59** from 100% to [redacted]%.

Failure events

B8. Please describe how failure events were reweighted in the model. Specifically, describe what weight was applied to each category of failure?

A description of how failure events were reweighted was included in Appendix N3 of the CS.

B9. Priority question: Please clarify the assumption that beyond 36 months, all new failure events are due to starting a new cGvHD treatment? Please discuss the clinical plausibility of no risk of recurrent malignancy after 36 months.

a) **Please provide scenario analyses exploring the impact of including risk of recurrent malignancy beyond 36 months for all arms of the model.**

The distribution of the failure events observed in both ROCKstar/KD025-208 (≥2 prior lines of therapy subgroup) and REACH-3 studies was used for the period before 36 months. For the period beyond 36 months, we assumed that all the failure events are due to starting a new therapy. This decision was supported by the fact that **no** recurrent malignancy events were observed in both ROCKstar/KD025-208 (≥ 2 prior lines of therapy subgroup) or

REACH-3 beyond month 24, with the majority of events due to starting a new cGVHD treatment.

After receiving the clarification questions, we consulted further on this topic. The clinical experts we spoke to confirmed that the risk of recurrent malignancy is highest directly after the transplant and gradually reduces with successful GVHD control. Risk would be expected to be close to 0 after 5-7 years. 5 years is the recommended time for follow-up, but clinicians confirmed that they have seen relapses after up to 7 years. They told us that at 36 months after 3L therapy for cGVHD, the rate of recurrent malignancy would be expected to be around 5%-10%. On this basis we have set the figure to 5% in the model to reflect the model distribution for 36 months and onwards.

The results from the scenario are provided below in Table 23.

This analysis is incorporated into the updated scenario presented in Appendix B.

Table 23. Deterministic results with inclusion of ongoing risk of recurrent malignancy beyond 36 months (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,704	£3,087	3.74%

User instructions: To change the distribution of FFS events after 36 months: Go to [**Efficacy**] sheet > Change cells **G67**, **K67** and **S67** to 95%. Change cells **H67**, **L67, P67** and **T67** to 5%.

Response outcomes

B10. Priority question: Please provide a scenario using response at 24 weeks for the ≥2 prior lines of therapy subgroup (as requested in Question A11)?

Using response at 24 weeks is not consistent with the structure of the model because of the way time to response is implemented. However, we have provided a scenario analysis using response at 24 weeks (Table 24).

Table 24. Deterministic results with response at 24 weeks (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,787	£3,156	6.06%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To change the proportions of responders: Go to [**Response**] sheet > Apply inputs according to the screenshot below:

Figure 9. User instructions screenshot for Question B10

B11. Priority question: The EAG explored the inclusion of response outcomes in the model with its clinical experts and while health-related quality of life (HRQoL) and resource use for patients may be different based on response, the company's model does not go far enough to model survival appropriately for the different types of responders as FFS and OS are not stratified by response. Additionally, the type of failure event is not linked to the type of responder (for instance, linking new cGvHD systemic treatment to patients

with a lack of response). As such, building upon the scenario already available in the model which removes response, please:

a) **Maintain the utility value for the failure-free health state (0.735).**

b) **Explore estimating a weighted disease management cost per cycle using the response rates included in the model.**

c) **Based on the response to B7, include the assumption of treatment weaning for the proportion of patients who have a complete and partial response for all treatment arms in the model.**

d) **Depending on the response to B4, implement separate health states for patients whose failure event was initiation of a new cGvHD systemic treatment and patients who have a recurrent malignancy.**

e) **Depending on the response to question A10, implement the results from the MAIC for the ≥2 prior lines of therapy subgroup for FFS and OS in the model, choosing appropriate extrapolations for OS.**

Please note: The three tables for the scenarios requested in parts a, b and c below build on one another to produce a cumulative ICER.

Answer to B11a

We have updated the utility analysis using the 2022 data cut and the utility value for the failure-free health state has been revised from 0.735 to 0.741. (See clarification answer B19). The results of the sensitivity analysis without inclusion of response in the analysis (and therefore maintenance of the utility for the overall failure-free state at 0.741) are provided overleaf in Table 25.

Table 25. Deterministic results with maintenance of utility value for failure-free at 0.741 (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,434	£2,862	-3.82%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Settings** sheet > Select "*No*" for cell **G16**.

Answer to B11b

We have implemented weighted disease management costs taking a simple approach, combining the "naïve" average proportions of responders across the 3 treatment arms (belumosudil QD, belumosudil BID and BAT) and applying that to all arms in the model. For instance, the weight for CR was assumed to be the average of the proportions of CR for belumosudil 200 mg QD, belumosudil 200 mg BID, and BAT: (4.35%+2.38%+6.71%)/3 = 4.48%.The weights are provided in Table 26 below.

Table 26. Calculation of weights for response levels used to subsequently derive weighted average disease management costs for the overall failure-free state (2022 data cut)

Treatment	Number of patients				Proportion		
	CR	PR	LR	Total	CR	PR	LR
Belumosudil 200 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Belumosudil 200 mg BID	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Average					[redacted]	[redacted]	[redacted]

CR: Complete response, PR: Partial response, LR: Lack of response

The results from implementing these weights to calculate disease management costs in the model for the failure-free state are provided below in Table 27. This builds on the answer to B11a above and includes the maintenance of utility value for the failure-free state at 0.741.

Table 27. Deterministic results with weighted average disease management costs and utility value for the overall failure-free state (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,237	£2,697	-9.36%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Settings** sheet > Select "*No*" for cell **G16**. Go to **Costs** sheet > Change cell **G315** to *£22,966.71*, **H315** to *£20,098.00*, **I315** to *£17,229.30*, **J315** to *£14,360.59* and **K315** to *£11,491.88*.

Answer to B11c

Please also see our answer to B7 above.

We have built on the answer to B11b above and applied the relative dose intensity from ROCKstar (2022 data cut; QD: [redacted]%; BID: [redacted]%) to the results as well as including the weighted disease management costs and maintenance of utility value for the failure-free state at 0.741 (Table 28).

Table 28. Deterministic results with relative dose intensity, weighted average disease management costs and maintenance of utility value for the overall failure-free state (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£1,279	£1,066	-64.17%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Settings** sheet > Select "*No*" for cell **G16**. Go to **Costs** sheet > Change cell **G315** to *£22,966.71*, **H315** to *£20,098.00*, **I315** to *£17,229.30*, **J315** to *£14,360.59* and **K315** to *£11,491.88*. Change cell **H58** from *100%* to [redacted]*%* and **H59** from *100%* to [redacted]*%.*

Answer to B11d

In our response to Question B4 we discussed the difficulties with implementing separate health states for patients whose failure event was initiation of a new cGvHD systemic treatment and patients who have a recurrent malignancy. Over and above the paucity of data to populate these states we are concerned that this structure could also result in a violation of OS. This is because the response to, and benefit from the new cGvHD therapy is implicitly captured through OS derived from the clinical trials and explicitly modelling response or FFS outcomes of the new systemic treatment could introduce double counting.

Hence, we have not updated the model to include this split to the failure state.

Answer to B11e

We have explained the reasons for not carrying out a MAIC in our answer to question A10 and so have not updated the model with data related to a matching analysis.

Adverse events

B12. Priority question: The adverse events provided in Tables 47 and 57 of the CS do not match those included in the model. Specifically, neutropenia,

fatigue, sepsis, leukopenia and dyspnoea seem to be missing. Please clarify why costs and disutilities for these events are provided in the CS but proportions are not included in the model.

The model includes Grade ≥3 adverse events (AEs) occurring in at least 5% of patients in either of the treatment arms of the pooled ROCKstar and Phase 2a trials or the BAT arm of the REACH-3 trial. Therefore, Table 43 of the CS incorrectly displayed data for neutropenia; and Tables 47 and 57 incorrectly displayed data for neutropenia, sepsis, leukopenia and dyspnoea. Fatigue was included in the model.

Based on the new data cut, diarrhoea was added to the list of relevant Grade ≥3 AEs, as it occurs in more than 5% of patients in the belumosudil QD arm. In turn, fatigue was removed from the list as it fell below the 5% threshold. The list of modelled Grade ≥3 AEs based on the new data cut is shown in Table 29.

Table 29. Modelled Grade ≥3 AEs occurring in >5%* patients in any treatment arm based on the 2022 data cut

	Pooled ROCKstar and Phase 2a		REACH-3
	Belumosudil 200 mg once daily (n=92)	Belumosudil 200 mg twice daily (n=84)	BAT (n=158)
Pneumonia	[redacted]	[redacted]	9.5%
Hypertension	[redacted]	[redacted]	7.0%
Anaemia	[redacted]	[redacted]	7.6%
Thrombocytopenia and decreased platelet counts**	[redacted]	[redacted]	10.1%
Hyperglycaemia	[redacted]	[redacted]	1.9%
Gamma-glutamyl transferase increased	[redacted]	[redacted]	1.9%
Diarrhoea	[redacted]	[redacted]	1.3%
Central line-related infections	[redacted]	[redacted]	12.9%†

BAT = best available therapy; ECP = extracorporeal photopheresis; N/A = not applicable TEAE = treatment-emergent adverse event
*Other than central line-related infections
** For BAT, thrombocytopenia and decreased platelet count events were reported aggregated. For belumosudil, the category only includes decreased platelet count events as there were no Grade ≥3 TEAEs of thrombocytopenia.
†Calculated value based on the assumptions that 64.6% of patients in the BAT arm are treated with ECP, and approximately 20% of patients have a central line-related infection based on feedback from the NICE advisory board.

a) **In Table 35 of the CS, lung infection met the criteria for inclusion in the model (Grade ≥3 adverse event [AE] occuring in >5% of patients in either**

arm in the pooled analysis [≥2 prior lines of therapy, 2-year analysis]). Please clarify why this was not included in the model.

Lung infection was not included as the definition and distinction from pneumonia (which is included in the model) was unclear, and it was not reported in a similar way across trials. Lung infection is below 5% in both belumosudil arms with the new data cut and so now falls outside the inclusion rule.

Health-related quality of life

B13. Priority question: There is a lack of detail in the CS around the PROMIS-GH data collected in ROCKstar and KD025-208. Please provide details on how PROMIS-GH was measured in the trials (timepoints of measurement, number of responses at each time point, length of follow up, etc.) along with the mean values mapped to EQ-5D-3L at each timepoint.

PROMIS-GH was assessed on D1 to Cycle 1 to Cycle 5 and then on D1 of every other Cycle thereafter (i.e., D1 of C1,2,3,4,5,7,9,11 etc). The mean EQ-5D-3L mapped score per visit, is presented in Figure 10 below.

Figure 10. Mean EQ-5D-3L mapped score per treatment arm and overall in the KD025-213 trial

B14. Please describe how the mapping algorithm for PROMIS-GH to the EQ-5D-3L by Thompson *et al*. was selected.

The mapping algorithm used to transform PROMIS-GH to the EQ-5D-3L by Thompson et al. was identified from inspection of the Oxford Population Health HERC database of mapping studies (HERC database of mapping studies — Health Economics Research Centre (HERC) (ox.ac.uk).

The HERC database included 3 mapping studies from PROMIS to EQ-5D 3L. These were:

- 2 mapping studies from PROMIS-GH to EQ-5D 3L (Thompson N, 2017; Revicki DA, 2009)
- 1 mapping study from PROMIS-29 to EQ-5D 3L (Hartman J, 2018).

The study by Hartman et al. was deemed not suitable as it is specific for PROMIS-29, while in the ROCKstar the PROMIS-GH was used. The more recent study from Thompson et al was considered to estimate EQ-5D-3L utility scores more accurately than the mapping proposed in 2009 by Revicki et al.

B15. Of the six models presented in Table 12 of Appendix N, which model was selected for the final regression? Please provide the goodness of fit statistics for all models assessed.

Information criteria from the six utility models fitted using the 30th of September 2022 data cut of the ROCKstar study are provided in Table 30. The same numbering convention as in Table 12 of Appendix N of the original submission is used. Utility values used in the CEM were based on Model 6 (with response irrespective of depth of response and failure irrespective of cause of failure). It is important to note that model selection was guided by face validity of the estimates obtained from the mixed regression models as well as the CEM model structure more than the information criteria achieved.

Across all models that considered treatment failure as a covariate the utility value obtained for the failure-state was higher than the one obtained in the failure-free state and because of this lack of face validity we chose to use a value from the literature instead of the calculated utility model for this health state.

Table 30. Information criteria achieved by the utility models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AIC	-2,661.509	-2,659.855	-2,622.620	-2,621.580	-2,623.007	-2,622.935

BIC	-2,636.071	-2,629.329	-2,576.999	-2,581.028	-2,587.524	-2,592.521

B16. When applying the scenario to exclude response in the model, a utility value of [redacted] for the failure-free health state is used. Please clarify how this utility value was estimated.

The mean utility value for the failure-free health state (0.735, standard error: 0.007) was estimated with a model investigating the effects of failure in the pooled (QD and BID) treatment arms of the ROCKstar trial (Model 1 in Table 12 of Appendix N.5 in the CS) using the 19th of August 2021 data cut. The updated results of the analysis from the 19th of August 2021 and the 30th of September 2022 data cuts of the ROCKstar study can be found below. (Table 31).

Table 31. Utility analyses for the failure-free health state without response and granularity of health states (fixed effects)

Factor	n patients	n observations	Coefficient	SE	Lower 95% CI	Upper 95% CI	p-value
Analyses in the original CS (19th of August 2021 data cut)							
Intercept	121	982	0.7350	0.0070	0.7211	0.7488	<0.001
Centred baseline utility score	121	982	0.7820	0.0480	0.6874	0.8773	<0.001
Treatment failure	23	52	0.0108	0.0130	-0.0148	0.0363	0.4087
Updated analyses (30th of September 2022 data cut)							
Intercept	140	1,197	0.7406	0.0064	0.7278	0.7532	<0.001
Centred baseline utility score	140	1,197	0.7841	0.0440	0.6976	0.8712	<0.001
Treatment failure	25	74	0.0037	0.0122	-0.0203	0.0278	0.7608

CI = confidence interval; SE = standard error

A summary of model-predicted utility values for each health state is presented in Table 32.

Table 32. Summary of utility scores for the failure-free and failure health states (linear predictions by mixed-effect model)

Health state	Mean	SE
Analyses in the original CS (19th of August 2021 data cut)		
Failure-free	[redacted]	[redacted]
Treatment-failure	[redacted]	[redacted]

Updated analyses (30th of September 2022 data cut)		
Failure-free	[redacted]	[redacted]
Treatment-failure	[redacted]	[redacted]

SE = standard error

B17. Priority question: Please estimate the utility value for a partial responder by excluding data for complete responders from utility regression analysis and compare this to the utility values used in the base case for CR and PR. If the estimated utility value is lower than the base case value, please provide a scenario where the utility value for CR is maintained as per the base case ([redacted]) and the lower utility value is used for PR.

a) **alternatively, the EAG's clinical experts expected that having a complete response would result in at least a 10% improvement in HRQoL compared with a partial response. As such, please explore a scenario where the utility value for CR is 10% higher than the value for PR.**

Regression coefficients estimates of the mixed model that included depth of response and failure as covariates along with predicted utility in the associated health state are presented in Table 33 and Table 34 using the 30th of September 2022 data cut of the ROCKStar study.

Table 33. Regression coefficient estimates from the mixed models including depth of response and treatment-failure as covariates.

Factor	Nr of pat.	Nr of obs.	Coef	Std Err	Low 95%CI	High 95%CI	P-value
Intercept	140	1,175	0.7215	0.0072	0.7073	0.7357	<0.001
Centered baseline utility score	140	1,175	0.7746	0.0424	0.6911	0.8587	<0.001
Treatment-failure	25	74	0.0218	0.0126	-0.0030	0.0465	0.0847
PR	109	725	0.0326	0.0061	0.0205	0.0446	<0.001
CR	6	32	0.0390	0.0174	0.0049	0.0732	0.0250

Table 34. Predicted health state utilities from the mixed models including depth of response and treatment-failure as covariates.

Health states	Mean EQ-5D	Low 95%CI	High 95%CI
Failure-free, no response	[redacted]	[redacted]	[redacted]
Treatment-Failure	[redacted]	[redacted]	[redacted]
Failure-free, PR	[redacted]	[redacted]	[redacted]
Failure-free, CR	[redacted]	[redacted]	[redacted]

As the estimated utility values for CR and PR ([redacted] and [redacted], respectively) from this mixed model are not lower than the utility values used in the base case for CR and PR ([redacted]), the deterministic ICER was calculated for a scenario using the value suggested in point B17 a) above (i.e., utility value for CR is 10% higher than the value for PR, hence utility in CR was applied as 1.1*[redacted] = [redacted]) in the model. See Table 35 below.

Table 35. Deterministic results with updated utility values (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,564	£2,970	-0.18%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Utilities** sheet > Change cell **G10** from [redacted] to [redacted].

B18. Priority question: In Appendix N.5.2 Table 14, a utility value for treatment failure is available, which the EAG considers could be used for patients who have initiated a new cGvHD systemic treatment in the failure health state. Please clarify why a treatment failure utility value was not estimated using the mixed-effects model presented in Table 15?

a) **If a treatment failure utility value can be estimated from the mixed-effects model, please calculate this and explore a scenario analysis where this value is used for patients who have initiated a new cGvHD systemic treatment in the failure health state. Alternatively, use the**

value of 0.712 from the mixed-effects model with the random effect in the requested scenario analysis.

We would like to clarify that Appendix N.5.2 Table 14 does not present the predicted health state value for treatment failure, but the average linear prediction of utility value for patients in the failure state and thus cannot be used to value the failure health state in the model.

Regression coefficients obtained by this model and predicted health state utility values are presented in Table 36 and in Table 37.

Table 36. Regression coefficient estimates from the mixed models including response and treatment-failure as covariates.

Factor	Nr of pat.	Nr of obs.	Coef	Std Err	Low 95%CI	High 95%CI	P-value
Analyses in the original CS (19th of August 2021 data cut)							
Intercept	121	969	0.7176	0.0078	0.7022	0.7330	<0.001
Centered baseline utility score	121	969	0.7730	0.0467	0.6810	0.8657	<0.001
Response	101	622	0.0287	0.0063	0.0164	0.0410	<0.001
Treatment failure	23	52	0.0269	0.0134	0.0006	0.0531	0.0447
Updated analyses (30th of September 2022 data cut)							
Intercept	140	1,175	0.7231	0.0071	0.7090	0.7372	<0.001
Centered baseline utility score	140	1,175	0.7727	0.0427	0.6888	0.8573	<0.001
Response	115	757	0.0291	0.0056	0.0180	0.0401	<0.001
Treatment failure	25	74	0.0200	0.0126	-0.0047	0.0446	0.1119

Table 37. Predicted health state utilities from the mixed models including response and treatment-failure as covariates.

Health states	Mean EQ-5D	Low 95%CI	High 95%CI
Analyses in the original CS (19th of August 2021 data cut)			
Failure-free, lack of response	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]
Treatment-failure	[redacted]	[redacted]	[redacted]
Updated analyses (30th of September 2022 data cut)			
Failure-free, lack of response	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]
Treatment-failure	[redacted]	[redacted]	[redacted]

The higher utility for the treatment failure state compared to the failure-free state is a consequence of the positive regression coefficient estimated for the treatment-failure state.

The total lack of face-validity of the estimates for the treatment failure state is likely driven by the small number of observations in this health state (23 patients reported utility values in a total of 52 visits, in the 19th of August 2021 data cut; 25 patients in 74 visits in the 30th of September 2022 data cut). Hence adverse selection bias might affect the utility estimates for treatment failure: 85 patients in the updated ROCKstar data had a FFS event, but only 25 of them had available utility records: worse-off patients in the treatment-failure state might not have completed the utility questionnaire.

We have noted earlier that the utility value in the failure health state must reflect the remaining course of the patient's life. Therefore, the lack of follow-up and very small number of observations from the study make the use of the observed data problematic and as we stated in the CS, we do not believe that the utilities for the failure state can reasonably be derived from the trial. It is worth comparing this high failure utility value ([redacted]) with the observation from the vignette-based utility study described in Appendix N in which the failure state was valued at only [redacted].

Whilst we feel the results are not informative due to the limitations with the analysis of utility for the failure health state from the studies, we have provided a scenario using the utility value for failure - new treatment estimated from ROCKstar based on mixed models including response and treatment failure as covariates ([redacted]) below in Table 38.

Table 38. Deterministic results with utility value for 'failure - new treatment' estimated from ROCKstar (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£5,186	£4,321	45.23%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Utilities** sheet > Change cell **G14** from *0.479* to [redacted].

B19. Priority question: From the utility models where failure - relapse of malignancy was included as a covariate (Table 12, Appendix N), please provide the estimated utility value from the model?

Regression coefficient estimates from the mixed model that included type of treatment-failure event as a covariate along with predicted utility for the associated health state are presented in Table 39 and Table 40 using the 30th of September 2022 data cut of the ROCKStar study.

Table 39. Regression coefficient estimates from the mixed models including type of treatment-failure event as covariate

Factor	Nr of pat.	Nr of obs.	Coef	Std Err	Low 95%CI	High 95%CI	P-value
Intercept	140	1,197	0.7406	0.0064	0.7278	0.7532	<0.001
Centered baseline utility score	140	1,197	0.7858	0.0439	0.6994	0.8727	<0.001
Failure state - New treatment	22	69	0.0014	0.0128	-0.0239	0.0266	0.9156
Failure state - Relapsed Malignancy	3	5	0.0261	0.0398	-0.0524	0.1045	0.5117

Table 40 Predicted health state utilities from the mixed models including type of treatment-failure event as covariate

Health states	Mean EQ-5D	Low 95%CI	High 95%CI
Failure-free	[redacted]	[redacted]	[redacted]
Treatment-failure - New therapy	[redacted]	[redacted]	[redacted]
Treatment-failure - Relapsed malignancy	[redacted]	[redacted]	[redacted]

As discussed in our answer to B19, the very high 'Failure' utilities lack face validity. For example, in our CS we provided evidence to show that the 'Treatment-failure - Relapsed malignancy' health state should attract a utility score of 0.479 (calculated as a weighted average based on utility values of progression/relapse health state of indications for the most recent transplant: AML, ALL, CML and CLL). This is considerably lower than [redacted] presented above (based on data from only 3 patients and higher than that observed in the failure-free state). Similarly, it is not credible to assume that failure with the addition of a new therapy would have the same quality of life impact as the failure-free health state. We have

discussed above in our answer to B6 how the failure event is likely associated with significant impact on a patient's HRQoL.

B20. Please provide a scenario where the carer disutility associated with failure - new cGvHD systemic therapy is the same as for failure-free PR/LR (-0.045)?

It is important to distinguish between the failure-free and failure health states with respect to caregiver impact. We tested this with the clinical experts we spoke to, and they told us that the chronic, progressive, and disabling nature of cGVHD as well as patients' need for daily assistance would have an increasing burden on caregivers over time, especially for those experiencing treatment failure. They agreed that caregiver disutilities of -0.045 and -0.142 taken from multiple sclerosis studies, were relevant proxies for the two health states respectively.

Therefore, we believe the differential should be maintained in the cost-effectiveness analysis but have provided the deterministic ICER with carer disutility set at the lower value of 0.045 in the failure due to new systemic therapy health state, as well as the failure-free PR/LR health states. The results are presented in Table 41 below.

Table 41. Deterministic results with utility values for carer disutility set to be equal between failure-free and failure - new systemic therapy (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£4,065	£3,388	13.85%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Utilities** sheet > Change cell **G172** from *-0.142* to *-0.045*.

B21. Please clarify why the duration of disutility for intravenous (IV) infusion has been set to 28 days?

As described in the CS, disutility values for intravenous (IV) infusion in the model are based on a time trade-off analysis from the Matza 2013 study. In the study, the participants were asked to consider living in a certain health state for a period of 2 years and were asked to evaluate receiving (a 2-hour) IV treatment once per month. The disutility results reflect both the length of the period and the frequency of treatment. It was assumed that the participants experience this as a constant disutility, therefore the length of disutility was set to be equal to

the length of a model cycle. As described in the study, “a greater treatment frequency would likely be associated with greater inconvenience”, that could influence the disutility value as well. Therefore, in our view it can be argued that receiving ECP treatments with a greater frequency would result in even greater disutility.

a) **Please explore a scenario where the duration of disutility for IV infusion is limited to the day of infusion only.**

We have stated above that the Matza study that was used to describe the disutility of IV infusion captures a single treatment once per month. Therefore, to reduce the disutility to a period of only one day would be to ignore the temporal evidence. We have also discussed how the disutility applied in the model may be conservative due to the frequency of treatments. We do not believe that carrying out this analysis will provide a useful scenario; however, we have included the results below in Table 42. For this scenario, we assumed that the disutility of IV infusion would be applied for 3 days per 4-week cycle (as a conservative assumption, as there are between 3 and 4 ECP procedures per 4-week cycle in the model [depending on the period] and 4 rituximab administrations in the first 4-week cycle of the model).

Table 42. Deterministic results disutility for IV infusion limited to the day of infusion only (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,609	£3,007	1.06%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Utilities** sheet > Change cell **I155** from *28* to *3*.

B22. Priority question: The EAG understands that the quality-adjusted life year (QALY) shortfall calculator (https://shiny.york.ac.uk/shortfall/) has recently been corrected for an error. As such, please ensure that the translation of the calculator in the company model is up to date and producing accurate results. Additionally, please provide a comparison of the model results with the equivalent results obtained from the online calculator as a verification check.

The QALY shortfall calculator in the model is up to date. As a comparison, we have calculated the QALY shortfall using the online calculator and inputs from the pooled clinical studies (i.e. starting age = 54 [rounded up from 53.9 in the study), female population = 42%, remaining QALYs in untreated arm = 2.1 [rounded down from 2.103], discount rate = 3.5%).

The results from the online calculator are presented below alongside those presented in the model. Note the small difference in results is due to rounding limitations in the online calculator.

Table 43. Results of online QALY shortfall calculator vs model

	Online calculator	Company submission model
Remaining QALYs without the disease	[redacted]	[redacted]
Remaining QALYs with the disease	[redacted]	[redacted]
Absolute shortfall	[redacted]	[redacted]
Proportional shortfall	[redacted]	[redacted]
QALY weight	**x 1.2**	**X 1.2**

Resource use and costs

B23. Priority question: In the submission, there is no discussion of concomitant medications and the company's justification for excluding these costs. In ROCKstar and KD025-208, clinical efficacy is based on belumosudil and concomitant medications patients received. Furthermore, in the NICE final scope, it is stated that the intervention is belumosudil with established clinical management.

a) **Please clarify why belumosudil has been modelled as monotherapy and not with established clinical management? The EAG considers that because there are modelled survival advantages with belumosudil, costs of established clinical management could be differential between the treatment arms.**
b) **Please clarify why concomitant drugs costs (based on those received in ROCKstar and the Phase 2a studies) have not been included in the model for the belumosudil arm, even though efficacy is influenced by the inclusion of these medications in the trials?**

c) **Please provide a scenario which included concomitant medication costs for the belumosudil once daily (QD) and twice daily (BID) arms of the model and if appropriate, for the BAT arm.**

Answer to B23a and B23b

Our clinical rationale for not including concomitant medications was based on guidance from our clinical trials team who felt that any 'additional efficacy' gained from belumosudil as an add-on is not likely to be 'boosted' by the pre-existing medications providing 'baseline' efficacy which would become concomitant treatments. Indeed, patients started on belumosudil treatment in the clinical trial because they met the criteria for entry meaning they required additional medication over and above their existing treatments.

BAT in REACH 3 did not include concomitant medications. We felt that in the absence of data to support a basket of third line and beyond comparator concomitant treatments from the study (which would likely be added to the REACH 3 BAT in third line) a fair comparison should not include the extra cost of these in the belumosudil arm only.

However, we do recognise that concomitant medications are used in clinical practice and did feature in the belumosudil studies and that many of these will be common between the two arms in the model. Therefore, we have provided a scenario in our answer to B23c below which includes concomitant medications.

Answer to B23c

We were unable to obtain a full breakdown of concomitant therapies for the pooled 2022 analysis within the timeframe of the clarification. In lieu of this, we have tabulated the therapies recorded in the ROCKstar study (2021 data cut) which are associated directly with systemic treatment for cGVHD overleaf, and which were used in at least 5 clinical trial subjects (Table 44). ECP costs made up 95% of the costs of concomitant treatment in this scenario analysis and so disproportionately effect the belumosudil treatment arm in the model.

This scenario was conducted assuming that patients would receive their concomitant treatment(s) for as long as they receive their main treatment (i.e., belumosudil or BAT), as informed by the time on treatment curves modelled for the latter, with the exception of concomitant ECP (see below).

The scenario including concomitant medication costs is likely to overestimate costs of treatment in the belumosudil arm. While it was not possible to extract the exact number of ECP doses administered as concomitant treatment, analysis of the ROCKstar CSR showed that of the patients receiving concomitant ECP, [redacted] stopped treatment, [redacted] decreased their ECP dose (frequency), [redacted] had no change in frequency, and [redacted] increased the ECP dose between their first and last belumosudil dose.

From the limited available data, we estimate that patients receiving concomitant ECP did so for an average [redacted]% of the time for which they received belumosudil, due mainly to tapering of treatment. This scenario was therefore conducted assuming that belumosudil patients on concomitant ECP received ECP for [redacted] of the time they were on belumosudil (this was implemented in a simplified way by multiplying the proportions of belumosudil patients with concomitant ECP by [redacted]).

Table 44. Concomitant treatment costs based on ROCKstar study (2021 datacut)

Concomitant treatment	Proportion with concomitant treatment (pooled analysis, n=156)		Proportion with concomitant treatment in BAT*	Assumed treatment regimen	Strength	Pack size	Price per pack / ECP procedure	Cost per 28-day cycle	Weighted cost per 28-day cycle		
	QD (n=81)	BID (n=75)							QD	BD	BAT
Prednisolone** n (%)	77 (95.1%)	73 (97.3%)	95.2%	1 mg/kg QOD	25 mg	56	£12.41	£9.73	£9.25	£9.47	£9.26
Tacrolimus n (%)	28 (34.6%)	28 (37.3%)	34.7%	1 mg BID	1 mg	100	£111.36	£62.36	£21.56	£23.28	£21.64
ECP n (%)	20 (24.7%)	26 (34.7%)	0%	3.2 sessions per 28-day cycle	-	-	£1,585.00	£5,072.00	£1,001.88***	£1,406.63***	£0.00
Sirolimus n (%)	17 (21.0%)	18 (24.0%)	21.1%	2 mg QD	2 mg	30	£172.98	£161.45	£33.88	£38.75	£34.13
MMF n (%)	11 (13.6%)	2 (2.7%)	13.0%	1000 mg BID	250 mg	100	£4.32	£9.68	£1.31	£0.26	£1.26
Budesonide n (%)	6 (7.4%)	3 (4.0%)	7.2%	3 mg TDS	3 mg	100	£60.00	£50.40	£3.73	£2.02	£3.65
Montelukast n (%)	4 (4.9%)	4 (5.3%)	5.0%	10 mg QD	10 mg	28	£0.68	£0.68	£0.03	£0.04	£0.03
Azithromycin n (%)	4 (4.9%)	4 (5.3%)	5.0%	250 mg TIW	250 mg	4	£0.44	£1.33	£0.07	£0.07	£0.07
Total	**167 (206.02%)**	**158 (210.07%)**	**181.2%**						**£1,071.72**	**£1,480.52**	**£70.04**

TDS=three times every day; TIW=three times a week; QOD=once every other day

* Proportions of patients receiving concomitant treatments in BAT were assumed based on the data available for belumosudil: they were calculated as a weighted average of the proportions of concomitant treatment use in the belumosudil QD and belumosudil BID arms, assuming 95% of belumosudil QD and 5% of belumosudil BID as per the distribution used in the model's base case. 0% concomitant use of ECP was assumed for the BAT arm.

** Prednisone is not used in the UK. Prednisolone dose is equivalent to prednisone in anti-inflammatory properties (https://cks.nice.org.uk/topics/corticosteroids-oral/background-information/equivalent-anti-inflammatory-doses/)

*** Calculated based on the assumption that patients receive concomitant ECP for an average [redacted]% of the time for which they receive belumosudil, due mainly to tapering of treatment.

The ability to include concomitant medication costs has been included in the economic model. This may be toggled on or off on the Settings sheet in cell [G91].

A scenario including concomitant medication costs is provided below in Table 45. This is also included as part of the alternative scenario presented in the Appendix B.

Table 45. Deterministic results including concomitant medications (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario 1: Assuming that those patients receiving concomitant ECP received it [redacted]% of the time spent on belumosudil according to the observed data in ROCKstar (See Table 44 for cost)	£16,674	£13,895	366.96%
Scenario 2: Assuming that those patients receiving concomitant ECP received it 100% of the time spent on belumosudil	£19,860	£16,550	456.21%
Scenario 3: Assuming that those patients receiving concomitant ECP received it 50% of the time spent on belumosudil	£11,893	£9,911	233.09%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

B24. Priority question: The EAG notes that in the company's NICE advisory board report, participants advised that 11.88% of patients receive tacrolimus and 7.81% receive cyclosporine.

a) **Please clarify why tacrolimus and cyclosporine were excluded from the comparators even though they are listed in the NICE final scope?**

b) **Please provide a scenario where the proportions of BAT reflect the proportions presented in the company's NICE advisory board report.**

c) **The EAG's clinical experts considered that tacrolimus and cyclosporine are used as background therapies in UK clinical practice. As such,**

please provide a scenario where costs of tacrolimus and cyclosporine are considered in the model for both belumosudil and BAT.

Answer to B24a

Clinicians advised that tacrolimus and cyclosporine can be used in practice but only as background therapies (not as standalone treatments). Since usage of these low-cost background therapies was expected to be equivalent for both treatment arms, we did not consider it reflective of clinical practice to incorporate these into the BAT basket. However, we have included tacrolimus as part of the basket of concomitant medications modelled in our answer to B23 for both the belumosudil and BAT arms.

Cyclosporine was used very rarely in the belumosudil studies. Only 1 patient (0.6%) received it in the pooled trial population and so given the intent behind clarification question B23 to ensure that efficacy in the studies was associated with appropriate costs, we have not included cyclosporine in the analysis.

Answer to B24b

Considering our response to B23 and B24a, we have provided a scenario analysis where the proportions of BAT reflect the proportions presented in the company's NICE advisory board report, but without including tacrolimus and cyclosporine. If required by the EAG, tacrolimus can be included along with other concomitant medications by selecting "Yes" on cell **G91** in the *Settings* sheet of the model. The proportions of concomitant medications (including tacrolimus) are adjustable by editing the values in cells **G159:N162** on the *Costs* sheet.

Table 46. Composition of BAT basket based on NICE advisory board report

Treatment	Proportion in alternative BAT
ECP	65.00%
Mycophenolate mofetil	22.50%
Imatinib	6.88%
Sirolimus	6.56%
Rituximab	5.31%

The results are provided overleaf in Table 47.

Table 47. Deterministic results including alternative BAT (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£4,484	£3,736	25.57%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Settings** sheet > Change cells **G64:G75** to align with Table 46.

Answer to B24c

Please see our answer to B24a and B23.

B25. Please explain the different assumptions of vial wastage included as options in the model (Tab 'Costs', cell G10)?

Two approaches are available in the model for the calculation of wastage costs

1) No vial sharing considered with minimum wastage

2) No vial sharing with minimum costs.

If drug costs are calculated using the approach of no vial sharing with minimum wastage (Option 1 above), the different vial sharing combinations are identified and the relevant BSA or weight for the dosage is calculated. The distribution of BSA among the patient population in the ROCKstar/KD025-208 (≥ 2 prior lines therapy subgroup) trials is used to then identify the vial sharing combination which leads to minimum wastage and the weighted cost per administration is calculated.

Similarly, if the approach of no vial sharing with minimum costs (Option 2) is used for calculation of drug costs including wastage, the different vial combinations for the different treatment doses are identified and the costs for these combinations are calculated. The vial combination with the minimum cost for the different dosages are selected and the BSA/weight distribution of patient population in the ROCKstar trial is used to calculate the weighted costs per administration for the patient population, which leads to the minimum costs.

The model uses minimum wastage as the base case.

B26. Priority question: The EAG considers that monitoring costs and resource use associated with follow up should be included in the model.

a) **The SmPC specified monitoring for belumosudil (complete blood cell count and liver function test must be performed before initiating therapy with belumosudil and liver function tests should be performed at least monthly throughout treatment). Please explain why the specified monitoring costs for belumosudil were excluded from the model?**

b) **The EAG's clinical experts considered that for all patients initiating treatment (belumosudil or BAT), complete blood cell count and liver function test would be given. Additionally, routine monitoring tests and specific organ monitoring would be performed. Follow up appointments with a consultant would take place fortnightly for the first 3-4 months of treatment to assess response and then reduce to monthly for the remainder of the first year and subsequent years, follow up would be every 6 to 8 weeks. Please provide a scenario where monitoring and follow up costs are included for all arms of the model based on the company's own clinical expert feedback (providing justification for assumptions) or utilising the EAG's clinical expert feedback. The EAG notes that in the company's NICE advisory board report, some details were given on follow up appointments that can be explored for the scenario.**

Answer to B26a

As described in Section 3.5.3.1 of the CS, we conducted a descriptive retrospective cohort study using Hospital Episode Statistics (HES) data, to quantify the real-world healthcare resource use (HCRU) and costs for cGvHD patients after an alloHSCT in England. HES contains information on reimbursed diagnoses and procedures from all NHS inpatient admissions, outpatient appointments and emergency care attendances. The resulting estimates from this study were considered to be the best available source of disease management costs for use in the model for patients receiving current standard of care in England. This assumption was validated at our advisory board.

The HCRU costs included outpatient (e.g. haematology, respiratory, dermatology and diagnostics services), inpatient, accident and emergency (A&E) attendances, and intensive care unit (ICU) attendances. The exact costs of blood cell counts, liver function tests (LFTs) and other organ-specific tests are not included in the HES study since these are not considered resource-intensive enough to influence the HRGs from which HCRU is calculated. (Note that the more significant costs of the outpatient appointments in which such costs would be incurred, are captured in the model from the HES study.) Any costs associated with testing are therefore assumed to be broadly equivalent between BAT and belumosudil arms because the model is predicated on the same number of patients starting in each arm resulting in a net zero cost differential for initiation of therapy and continued monitoring.

We sought additional validation of this assumption by speaking to UK clinical experts attending the European Society for Blood and Marrow Transplantation Annual Meeting (April 2023). There was unanimous agreement that clinicians would not need to change their current clinical practice to incorporate the monitoring requirements for belumosudil. They also added that, when ruxolitinib was temporarily reimbursed through the interim COVID commissioning policy for cGVHD, there was no change to practice and no additional resource was required. The SmPC for ruxolitinib stipulates that complete blood count should be performed before initiating therapy and monitored every 2-4 weeks until doses are stabilised.

Answer to B26b

The SmPC for belumosudil does specify that a complete blood cell count and liver function test must be performed before initiating therapy, and that liver function tests (LFTs) should be performed at least monthly throughout treatment. We consider this to be consistent with current standard of care for patients at this stage in the disease receiving other treatments. For example, a typical ECP hospital protocol requires platelet count and haematocrit levels to be assessed before each procedure (i.e. multiple times per month) and continued monitoring for anaemia, thrombocytopenia, hypocalcaemia (see https://nssg.oxford-haematology.org.uk/bmt/gvhd/B-74-extracorporeal-photopheresis-pathway-for-acute-and-chronic-gvhd.pdf). The SmPC for mycophenolate mofetil requires regular monitoring (complete blood counts weekly during the first month, twice monthly for the second and third months of treatment, then monthly through the first year) (see also https://www.medicines.org.uk/emc/product/1103/smpc). Similarly, regular full blood tests, liver and renal function monitoring are required prior to and during treatment with imatinib (https://www.medicines.org.uk/emc/product/7779/smpc). Together these three treatments

form 92% of the BAT arm. We consider that the healthcare resource use associated with initiation and monitoring during treatment with belumosudil will be comparable, if not lower than, that of BAT in cGVHD management in the UK.

Across all cGVHD patients evaluated in the HES study which fed into the model (unadjusted for the first year), there were an average of 29 outpatient visits and 10 inpatient visits per patient year. This is significantly higher than the suggestion by the EAG's clinical experts (15-16 appointments in the first year and 6-9 appointments in subsequent years. No information about inpatient visits was offered in the question). The figures suggested by one clinician on the advisory board included ECP appointments and therefore could not be applied to all patients in the model (since ECP resource costs are captured separately). Therefore, despite its limitations, we considered that the HCRU estimates from the HES study which are based on real world English clinical practice are the most appropriate and comprehensive HCRU inputs for the model.

Allogeneic stem cell transplantation and cGVHD are associated with extremely high healthcare resource use, as demonstrated in our HES study, and confirmed by clinicians at the January advisory board. Monitoring and follow up contribute to this resource use.

We have used the best available source of data to capture these costs in the model, of which initiation and monitoring are expected to form a small but balanced proportion.

In order to perform the analysis requested, the outpatient appointments would need to be broken out from the HES data and replaced with the suggestions by the EAG's clinical experts. This would result in lower resource use which we do not believe is reflective of real-world clinical practice and is contrary to the evidence we have collected and discussed above. Therefore, we would prefer the analysis to be maintained with the HES data as the HCRU source of evidence.

B27. Priority question: Based on the pooled analysis of ROCKstar and KD025-208 for the ≥2 prior lines of therapy subgroup, please provide data on the subsequent treatments received by patients whose failure event was initiation of a new cGvHD systemic treatment.

a) **Please explain why subsequent treatments from the pooled analysis of ROCKstar and KD025-208 for the ≥2 prior lines of therapy subgroup**

were not used to inform the basket of subsequent treatments used in the model?

b) **Please provide a scenario using the subsequent treatments and proportions of use from the pooled analysis of ROCKstar and KD025-208 for the ≥2 prior lines of therapy subgroup, making adjustments for treatments not provided in the UK.**

c) **Were any patients in ROCKstar and KD025-208 given a second alloHSCT? If so, please explore a scenario where costs of subsequent alloHSCT are included in the model. If equivalent data are available from REACH-3, please also include this in the scenario for the BAT arm of the model.**

Answer to B27a

No details about the subsequent treatments used beyond FFS were collected in the belumosudil trials. Neither has this information been published for the REACH-3 study. Therefore, we cannot provide a definitive list of subsequent treatments for the belumosudil or BAT arms taken from clinical trial data.

Answer to B27b

Please see our answer to B27a. This scenario cannot be provided as the information was not collected in the trials.

Answer to B27c

No information on a subsequent alloHSCT was collected in the belumosudil trials and this information is not available from the REACH-3 publication. Therefore, this scenario cannot be tested.

B28. Priority question: The EAG's clinical experts consider the basket of subsequent treatments do not reflect UK clinical practice. Instead, they considered that subsequent treatments would consist of those in the below

table. Please provide a scenario implementing the EAG's clinical expert assumptions of subsequent treatment.

Subsequent treatment	Proportion
ECP	14.5%
Mycophenolate mofetil	14.5%
Sirolimus	10%
Rituximab	5%
Imatinib	10%
Cyclosporine	15%
Pulsed methylprednisolone	10%
No further treatment	21%

The subsequent treatment table in the model represents the proportion of time spent on all future subsequent treatments over the remaining course of the treatment pathway (until death).

Pulsed corticosteroids are administered at relatively high doses over a short period of time (e.g. 10 mg/kg per day for 4 days) (Dignan et al., 2012; Akpek et al., 2001). Therefore, even if used in clinical practice, pulsed methylprednisolone is unlikely to represent 10% of subsequent treatment for patients beyond third line therapy.

Pulsed corticosteroids were included as comparators to belumosudil in the NICE draft scoping document based on their inclusion in NHS England's specialised commissioning policy for treatments for cGVHD as a "third line" treatment option. However, clinical experts who responded to the consultation and attended the scoping workshop advised that, like methotrexate, pulsed corticosteroids were rarely used clinical practice in England and were not considered to be a safe and effective therapy. Following these recommendations, pulsed corticosteroids was removed from the final scope. Since the licence for belumosudil covers all lines of therapy beyond third line, we consider the same logic to apply for belumosudil comparators as for the choice of subsequent treatments.

Despite these considerations, we have provided a scenario implementing the EAG's clinical expert assumptions for subsequent treatment. This is provided below in Table 48.

Table 48. Deterministic results including alternate subsequent treatment basket (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)	
	WITHOUT severity modifier	**WITH severity modifier**
Base case	£3,571 (Incremental QALYs: [redacted], Incremental Costs: [redacted])	£2,976 (Incremental QALYs: [redacted], Incremental Costs: [redacted])
Scenario	Belumosudil is Dominant, Incremental QALYs: [redacted], Incremental Costs: [redacted]	Belumosudil is Dominant, Incremental QALYs: [redacted], Incremental Costs: [redacted]

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Subsequent Tmt** sheet > Change cells **G10:J14** to match the proportions of treatment provided in the table above.

B29. Please justify the assumption that for failure patients whose failure event is a new cGvHD treatment, costs of subsequent treatments are incurred for the remaining time horizon even though clinical expert feedback was obtained by the company to estimate the mean durations of each treatment for scenario analysis 28? Please clarify why the clinical expert feedback was deemed not to be suitable to use for the company base case?

The main assumptions behind calculating costs for the remaining time horizon in the failure state reflect the fact that patients cycle through multiple treatments in later lines, that treatment pathways are highly patient/case-dependent and that the pooled trials included patients from various lines of therapy.

Therefore, the intention was to capture a lifecycle of subsequent treatments, allowing for the fact that patients are treated in different lines. The model does account for some 'time off' treatment with an assumption that patients spend only of 60% of their time on treatment during this period.

The alternative approach with a fixed treatment duration (based on the clinical expert feedback) for each of the potential subsequent treatments would only account for a single subsequent line of therapy, calculated as a one-off cost of the basket of treatment options. This would significantly underestimate the cost of treatment in the failure state.

Section C: Textual clarification and additional points

C1. Priority question: Please provide the CSR for the KD025-208 trial.

The CSR for the KD025-208 trial was shared with NICE on 7th April 2023. Please let us know if you did not receive it.

C2. Priority question: Please provide the full-text for the following three documents referenced in the CS:

- Malik MI, Litzow M, Hogan W, Patnaik M, Murad MH, Prokop LJ, et al. Extracorporeal photopheresis for chronic graft-versus-host disease: a systematic review and meta-analysis. Blood Res. 2014;49(2):100-6.
- HAS. Jakavi (ruxolitinib) - Acute (GvHa) and chronic (GvHc) graft versus host disease when there is an inadequate response to corticosteroids or other therapies. Assess HEALTH TECHNOLOGIES. 2022.
- Peacock A DF, Taylor C, et al. . P615 Cost-effectiveness of extracorporeal photopheresis for the treatment of chronic graft versus host disease in the Australian setting. The 48th Annual Meeting of the European Society for Blood and Marrow Transplantation: Patient Advocacy – Poster Session (P613-P616). Bone Marrow Transplantation. 2022;57(1):468-70.

We have provided these articles as part of the package to this response.

C3. In the CS, Table 25 reports identical organ involvement for joints/fascia and eyes. Please provide an updated table if this is an error in reporting.

We have cross-checked Table 25 against the pooled analysis results and can confirm that this information is accurate; no updated table is required.

C4. The EAG could not verify the strength and package size of mycophenolate mofetil included in the model against what is presented in Table 1 of Appendix K of the CS. Additionally, the pack price used in the model and in Table 1 of Appendix K of the CS could not be verified against the price given in eMIT (which is £4.37). Please review and amend the model as necessary.

The package price of £4.88 (250 mg capsules, with package size of 100) was downloaded from eMIT on 3 March 2023. Appendix K incorrectly displays the package strength (500 mg

capsules instead of 250 mg capsules as in the model) and package size (package size of 50 instead of 100 as in the model). A new version of eMIT (updated on 22 March after the submission) includes a £4.32 price for the 250 mg capsules with package size of 100. The model was updated to include this £4.32 price (instead of the £4.37 price proposed by NICE).

All analyses provided in the response to these clarification questions includes this update in the model.

C5. The price of imatinib 100 mg and 400 mg from eMIT is £27.12 and £36.59, respectively. Please update the model with the correct prices.

The package prices for imatinib were downloaded from eMIT on 3 March 2023. The model was updated to include the respective £27.12 and £36.59 package prices proposed by NICE.

All analyses provided in the response to these clarification questions includes this update in the model.

C6. Cell J24 in the 'AEs' tab of the model should be 156. Please clarify if the model is correct and amend as necessary.

The value (158) in the model is correct. 158 patients received control therapy (BAT) in REACH-3 and the safety analyses were conducted on these 158 patients. 158 is the number provided for the control arm in Table 2 in Zeiser et al. 2021, from which the adverse event data used in the model were obtained.

C7. Cell I123 in the 'Utilities' tab of the model does not match the data presented in Table 47 of the CS. Please clarify if the model is correct and amend as necessary.

The model uses 14.7 days (i.e., 2.1 weeks) as the duration for pneumonia. This value was calculated based on Table 73 of TA359, as an average of the durations of pneumonia treated with idelalisib with rituximab (1.6 weeks) and rituximab only (2.6 weeks). Therefore, Table 47 of the CS incorrectly displayed 18.2 days (i.e., 2.6 weeks - the value of pneumonia in the rituximab only arm in TA359) as the duration of pneumonia and not the average.

No amendment to the model is necessary.

C8. Please clarify why in the model, a cycle is assumed to be 28 days (Parameters tab cell K15) rather than 30.44 days (Parameters tab cell K10)?

The model cycle length of 4 weeks (28 days) was chosen as it is short enough to accurately capture differences in cost or health effects between cycles, aligns with the data collection and reporting in both the ROCKstar trial and the pooled ROCKstar and KD025-208 trials, and treatment schedules of comparators can be easily considered. If any of the inputs were only available in months, it was recalculated in the model to be aligned with the cycle length.

C9. Priority question: Please provide instructions on how to run scenarios 7, 12, 21-22, 27-28, 31-32, 34-36, 38, 41-53 and 56. The EAG suggests presenting the guidance in a table with the base case values, the scenario values and then instructions or as a separate appendix.

Table 49 provides explicit instructions about how to update the model for each analysis of interest. In doing so it should be clear to the EAG which units have changed and how they have been varied. Detailed descriptions of the scenarios are included in Appendix N of the CS.

Instructions are provided below in Table 49.

Table 49 Scenario analysis model instructions

#	Scenario instructions
Base case	
7	[redacted]
12	[redacted]
21	[redacted]
22	[redacted]
27	[redacted]
28	[redacted]
31	[redacted]
32	[redacted]

#	Scenario instructions
34	[redacted]
35	[redacted]
36	[redacted]
38	[redacted]
41	[redacted]
42	[redacted]
43	[redacted]
44	[redacted]
45	[redacted]
46	[redacted]
47	[redacted]
48	[redacted]
49	[redacted]
50	[redacted]
51	[redacted]
52	[redacted]
53	[redacted]
56	[redacted]

Appendix A. Disaggregated results and additional sensitivity analysis results

This appendix presents additional deterministic results from the updated model, complementing our answer to clarification question A1.

Disaggregated discounted health benefits by health state for belumosudil and BAT are shown in Table 50 and Table 51 without a severity modifier and with a severity modifier, respectively. Patients treated with belumosudil had an overall increase in health-state related QALYs of [redacted] (without applying a severity modifier) and [redacted] (with a 1.2 QALY weight) compared to BAT. The results show that patients treated with belumosudil have higher QALYs in the failure-free state, and lower QALYs in the 'failure – new chronic GVHD systemic therapy' health state, compared to those treated with BAT.

Table 50. Summary of QALY gain by health state (without severity modifier)

Health state	QALY intervention (belumosudil)	QALY comparator (BAT)	Increment	Absolute increment	% absolute increment
Failure-free - In response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure-free - Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - New chronic GVHD therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - Recurrent malignancy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total health-state related QALYs*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BAT = best available therapy; GVHD = graft-versus-host-disease; QALY = quality-adjusted life year
* Values in this table do not include QALY decrements due to adverse events, decrements associated with IV infusion and decrements related to caregivers' quality of life
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Table 51. Summary of QALY gain by health state (with severity modifier [1.2 QALY weight])

Health state	QALY intervention (belumosudil)	QALY comparator (BAT)	Increment	Absolute increment	% absolute increment
Failure-free - In response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure-free - Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - New chronic GVHD therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - Recurrent malignancy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total health-state related QALYs*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BAT = best available therapy; GVHD = graft-versus-host-disease; QALY = quality-adjusted life year
* Values in this table do not include QALY decrements due to adverse events, decrements associated with IV infusion and decrements related to caregivers' quality of life

Results showing the discounted disaggregated costs by health state for each treatment are shown in Table 52. These summary results by health state include disease management costs and one-off cost related to treatment of recurrent malignancy but do not include drug acquisition and drug administration costs nor costs of managing adverse events.

Table 52. Summary of costs by health state (excluding drug costs and AE costs)

Health state	Cost intervention (belumosudil)	Cost comparator (BAT)	Increment	Absolute increment	% absolute increment
Failure-free - In response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure-free - Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - New chronic GVHD therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Failure - Recurrent malignancy†	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total disease management and one-off recurrent malignancy costs*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BAT = best available therapy; GVHD = graft-versus-host-disease;
* Values in this table do not include drug acquisition and drug administration costs nor costs of managing adverse events
† The values for 'Failure - Recurrent malignancy' include disease management costs in the 'Failure - Recurrent malignancy' health state as well as the one-off cost of treating recurrent malignancy
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Base-case results for resource use by cost categories are shown in Table 53. Patients treated with belumosudil incurred incremental costs of [redacted] per patient compared to patients treated with BAT. Most of the incremental costs of belumosudil are attributable to drug acquisition costs ([redacted]). Considerable cost offsets with belumosudil were achieved in subsequent treatment costs (costs of new chronic GVHD systemic therapy [[redacted]]) and to a smaller extent in disease management costs ([redacted]).

Table 53. Summary of predicted resource use by category of cost (with PAS)

Health state	Cost intervention (belumosudil)	Cost comparator (BAT)	Increment	Absolute increment	% absolute increment
Technology (drug acquisition) costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug Administration Costs*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost of New chronic GVHD Systemic Therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost of Recurring Malignancy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event Management Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	**£ 248,736**	[redacted]	[redacted]	[redacted]

BAT = best available therapy; ECP = extracorporeal photopheresis; GVHD = graft-versus-host-disease;
*For patients receiving ECP within BAT, the row for 'Drug Administration Costs' also includes accommodation costs related to ECP administration.
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Results for resource use by cost categories with list price are displayed in Table 54.

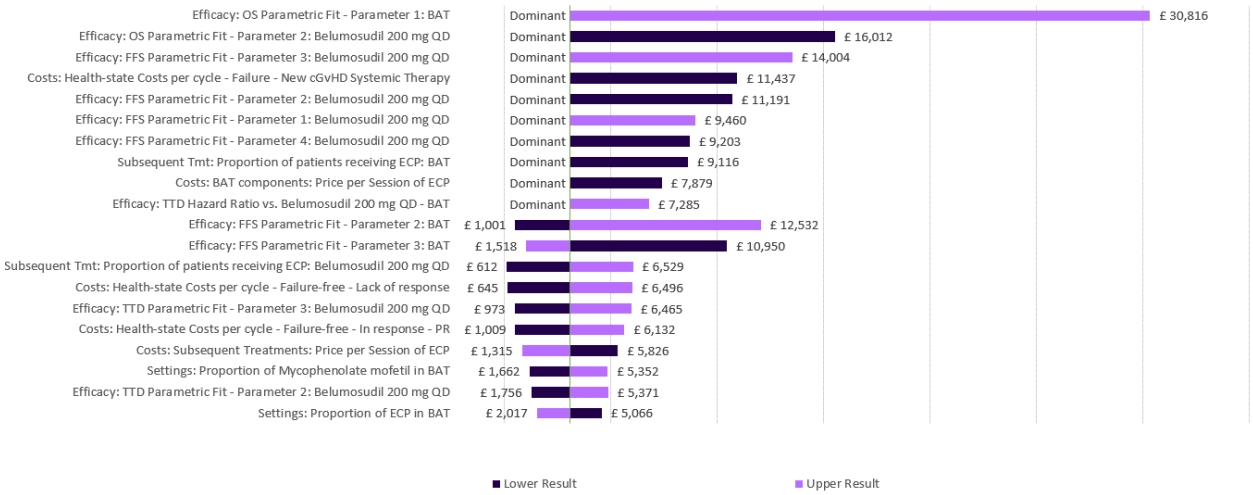
Table 54. Summary of predicted resource use by category of cost (with list price)

Health state	Cost intervention (belumosudil)	Cost comparator (BAT)	Increment	Absolute increment	% absolute increment
Technology (drug acquisition) costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug Administration Costs*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease Management Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost of New chronic GVHD Systemic Therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cost of Recurring Malignancy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event Management Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	**£ 248,736**	[redacted]	[redacted]	[redacted]

BAT = best available therapy; ECP = extracorporeal photopheresis; GVHD = graft-versus-host-disease;
*For patients receiving ECP within BAT, the row for 'Drug Administration Costs' also includes accommodation costs related to ECP administration.
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Updated one-way deterministic sensitivity analyses are presented in Figure 11 below.

Figure 11. Tornado diagram of ICER (incremental cost per QALY gained) for belumosudil vs. BAT (PAS, without severity modifier)



BAT = best available therapy; BID = twice daily; cGvHD = chronic graft-versus-host disease; ECP = extracorporeal photopheresis; FFS = failure-free survival; ICER = incremental cost-effectiveness ratio; OS = overall survival; PAS = patient access scheme; PR = partial response; QALY = quality-adjusted life year; QD = once daily; Tmt = treatment; TTD = time to treatment discontinuation
Notes: FFS Parametric Fit – Parameter 1 = mu (Generalised Gamma distribution); FFS Parametric Fit – Parameter 2 = sigma (Generalised Gamma distribution); FFS Parametric Fit – Parameter 3 = Q (Generalised Gamma distribution); FFS Parametric Fit – Parameter 4 = treatment coefficient; OS Parametric Fit – Parameter 1 = rate (exponential distribution); OS Parametric Fit – Parameter 2 = treatment coefficient; TTD Parametric Fit – Parameter 1 = mean (log-normal distribution); TTD Parametric Fit – Parameter 2 = standard deviation (log-normal distribution); TTD Parametric Fit – Parameter 3 = treatment coefficient

Selected updated sensitivity analyses are presented in Table 55 below. The most impactful scenarios out of those presented in Table 65 of the CS were retained here: all scenarios that resulted in the absolute change from the base-case ICER being ≥ 25% or in belumosudil being dominant in the CS were included here.

Table 55. Sensitivity analyses with percentage change from base-case ICER (PAS)

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
	Base case	£3,571	£2,976	
1	Time horizon: 20 years	£6,436	£5,364	+80.26%
2	Time horizon: 5 years	£53,692	£44,743	+1403.68%
3	Discount rates: 0%	Belumosudil is Dominant, Incremental QALYs: 1.963, Incremental Costs: -£9,561	Belumosudil is Dominant, Incremental QALYs: 2.355, Incremental Costs: -£9,561	

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
4	Discount rates: 1.5%	Belumosudil is Dominant, Incremental QALYs: 1.723, Incremental Costs: -£1,983	Belumosudil is Dominant, Incremental QALYs: 2.068, Incremental Costs: -£1,983	
5	Alternative distribution of BAT components	£8,827	£7,356	+147.21%
6	FFS for all treatments: Joint Fit - Gamma	£64,208	£53,507	+1698.21%
7	FFS for belumosudil QD and BID: Joint Fit - Log-normal; FFS for BAT: Joint Fit - Weibull	£34,893	£29,077	+877.20%
8	FFS for belumosudil QD and BID: Independent Fit - Log-normal for QD and Generalised Gamma for BID; FFS for BAT: Independent Fit - Gamma	£25,163	£20,969	+604.72%
9	OS long-term assumption for belumosudil: Do not assume same probability of death as for BAT after 5 years.	£23,754	£19,795	+565.26%
10	OS for all treatments: Joint Fit - Log-normal	Belumosudil is Dominant, Incremental QALYs: 1.549, Incremental Costs: -£38,544	Belumosudil is Dominant, Incremental QALYs: 1.859, Incremental Costs: -£38,544	
11	OS for all treatments: Joint Fit - Weibull	Belumosudil is Dominant, Incremental QALYs: 1.581, Incremental Costs: -£28,084	Belumosudil is Dominant, Incremental QALYs: 1.898, Incremental Costs: -£28,084	
12	OS for all treatments: Joint Fit - Gamma	Belumosudil is Dominant, Incremental QALYs: 1.552, Incremental Costs: -£22,807	Belumosudil is Dominant, Incremental QALYs: 1.863, Incremental Costs: -£22,807	
13	OS for belumosudil QD and BID: Independent Fit - Gamma for QD; Log-normal for BID; OS for BAT: Independent Fit - Log-logistic	Belumosudil is Dominant, Incremental QALYs: 1.565, Incremental Costs: -£35,967	Belumosudil is Dominant, Incremental QALYs: 1.878, Incremental Costs: -£35,967	

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
14	TTD for BAT: exponential curve fitted to median	£5,174	£4,312	+44.91%
15	Treat until failure (all treatments)	£27,792	£23,160	+678.35%
16	Treat until failure (BAT only)	Belumosudil is Dominant, Incremental QALYs: 1.483, Incremental Costs: -£7,818	Belumosudil is Dominant, Incremental QALYs: 1.779, Incremental Costs: -£7,818	
17	Alternate distribution of subsequent treatments (applied for all initial treatments)	Belumosudil is Dominant, Incremental QALYs: 1.476, Incremental Costs: -£8,674	Belumosudil is Dominant, Incremental QALYs: 1.771, Incremental Costs: -£8,674	
18	Alternate approach to costing of subsequent treatments	£17,158	£14,298	+380.52%
19	Maximum duration of treatment for all treatments (except rituximab): 3 years	Belumosudil is Dominant, Incremental QALYs: 1.475, Incremental Costs: -£2,697	Belumosudil is Dominant, Incremental QALYs: 1.770, Incremental Costs: -£2,697	
20	No maximum duration of treatment for all treatments (except rituximab)	£10,634	£8,862	+197.81%
21	Alternate proportions of responders to ECP assumed for drug cost calculations	£7,742	£6,452	+116.82%
22	Disease management costs for all Failure-free health states follow the decrease observed in Schain et al. 2021	Belumosudil is Dominant, Incremental QALYs: 1.476, Incremental Costs: -£9,553	Belumosudil is Dominant, Incremental QALYs: 1.771, Incremental Costs: -£9,553	
23	Disease management costs for all Failure-free health states reduced in Years 5+	Belumosudil is Dominant, Incremental QALYs: 1.476, Incremental Costs: -£8,680	Belumosudil is Dominant, Incremental QALYs: 1.771, Incremental Costs: -£8,680	
24	Health state utility for Failure - New chronic GVHD Systemic Therapy: Crespo et al. 2012	£4,703	£3,919	+31.71%
25	Value of health state utility for Failure - Recurrent Malignancy and for Failure - New chronic GVHD	£4,586	£3,821	+28.43%

No.	Scenario	ICER without severity modifier	ICER with severity modifier (QALY weight of 1.2)	% change from base-case ICER
	Systemic Therapy: upper bound of range			

BAT = best available therapy; BID = twice daily; ECP = extracorporeal photopheresis; FFS = failure-free survival; HES = Hospital Episode Statistics; ICER = incremental cost-effectiveness ratio; OS = overall survival; QD = once daily; TTD = time to treatment discontinuation

Appendix B. Scenario analysis with consideration of the EAG comments and suggestions

We are grateful to the EAG for their analysis of the original CS and suggestions for changes to the model and inputs.

We have carefully considered all of the comments made by the group and present an alternative scenario below which includes the following plausible adjustments to the original company analysis.

1. Update to the 2022 data-cut for the pooled analysis for the ≥2 prior lines of therapy subgroup (See clarification question A1).
2. Update to the EMIT prices for mycophenolate mofetil and imatinib (See clarification questions C4 & 5).
3. Updated adverse event profile to include diarrhoea.
4. Inclusion of concomitant medications in both arms of the model (See clarification question B23).
5. Reduction of dose intensity for belumosudil to observed study levels ([redacted]% for the QD arm, [redacted]% for the BID arm) (See clarification question B7).
6. Inclusion of recurrent malignancy after 36 months at a proportion of 5% of the failure events (See clarification question B9).

The results of the updated model with these adjustments are tabulated in Table 56.

Table 56. Deterministic alternative scenario with EAG plausible assumptions.

Iteration	ICER WITHOUT 1.2 severity modifier	ICER WITH 1.2 severity modifier	Incremental change to ICER
Update to the 2022 data cut including updated adverse events along with correction to prices for mycophenolate mofetil and imatinib (Points 1-3 above)	£3,571	£2,976	N/A
Addition of concomitant medications (assumes ECP received for 80% of the time on belumosudil)	£ 16,674	£13,895	366.96%
Reduction in belumosudil dose intensity from 100% to [redacted]% (QD), [redacted]% (BID)	£14,701	£12,251	-11.83%
Inclusion of recurrent malignancy after 36 months at a proportion of 5% of the failure events	£14,834	£12,362	0.91%

User instructions for alternative analysis:

- The 2022 data cut is already incorporated into the latest version of the model provided to the EAG.

- To add concomitant medications according to this scenario: Go to **[Settings]** sheet > Select "*Yes*" in cell **G91**.
- To adjust the dose intensity for belumosudil: Go to **[Costs]** sheet > Change cell **H58** from 100% to [redacted]% and cell **H59** from 100% to [redacted]%.
- To change the distribution of FFS events after 36 months: Go to [**Efficacy**] sheet > Change cells **G67**, **K67,** and **S67** to 95%. Change cells **H67**, **L67,** and **T67** to 5%.

Full cost-effectiveness results for the scenario analysis incorporating EAG plausible changes.

The cost-effectiveness estimates provided below include all of the changes made to the settings detailed in Table 56 above.

Table 57. Alternative scenario with EAG plausible assumptions: probabilistic results (PAS, WITHOUT and WITH severity modifier)

	WITHOUT severity modifier		WITH severity modifier	
Outcome	**Belumosudil**	**BAT**	**Belumosudil**	**BAT**
Total costs	[redacted]	£ 250,921	[redacted]	£ 250,921
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]		[redacted]	
Incremental LYs	[redacted]		[redacted]	
Incremental QALYs	[redacted]		[redacted]	
ICER (£/QALY)	£ 14,509		£ 12,091	
INHB (£20,000/QALY)	[redacted]		[redacted]	
INHB (£30,000/QALY)	[redacted]		[redacted]	
INMB (£20,000/QALY)	[redacted]		[redacted]	
INMB (£30,000/QALY)	[redacted]		[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

Table 58. Alternative scenario with EAG plausible assumptions: deterministic results (PAS, WITHOUT and WITH severity modifier)

	WITHOUT severity modifier		WITH severity modifier	
Outcome	**Belumosudil**	**BAT**	**Belumosudil**	**BAT**
Total costs	[redacted]	£ 249,277	[redacted]	£ 249,277
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]		[redacted]	
Incremental LYs	[redacted]		[redacted]	
Incremental QALYs	[redacted]		[redacted]	
ICER (£/QALY)	£ 14,834		£ 12,362	
INHB (£20,000/QALY)	[redacted]		[redacted]	
INHB (£30,000/QALY)	[redacted]		[redacted]	
INMB (£20,000/QALY)	[redacted]		[redacted]	
INMB (£30,000/QALY)	[redacted]		[redacted]	

BAT = best available therapy; ICER = incremental cost-effectiveness ratio; LY = life year; QALY = quality-adjusted life year; INHB = incremental net health benefit; INMB = incremental net monetary benefit

The results for the probabilistic analysis (without severity modifier) are plotted on the cost-effectiveness plane for belumosudil compared to BAT over 5,000 iterations in Figure 12 below (corresponding to Figure 35 in the CS). The majority of iterations ([redacted]%) fell in the north-east quadrant, indicating that treatment with belumosudil was consistently more effective and more costly than BAT. However, [redacted]% of the iterations fell in the south-east quadrant which indicates belumosudil being a dominant treatment option, providing better health outcomes at lower costs compared to BAT. 1 iteration fell in the SW quadrant.

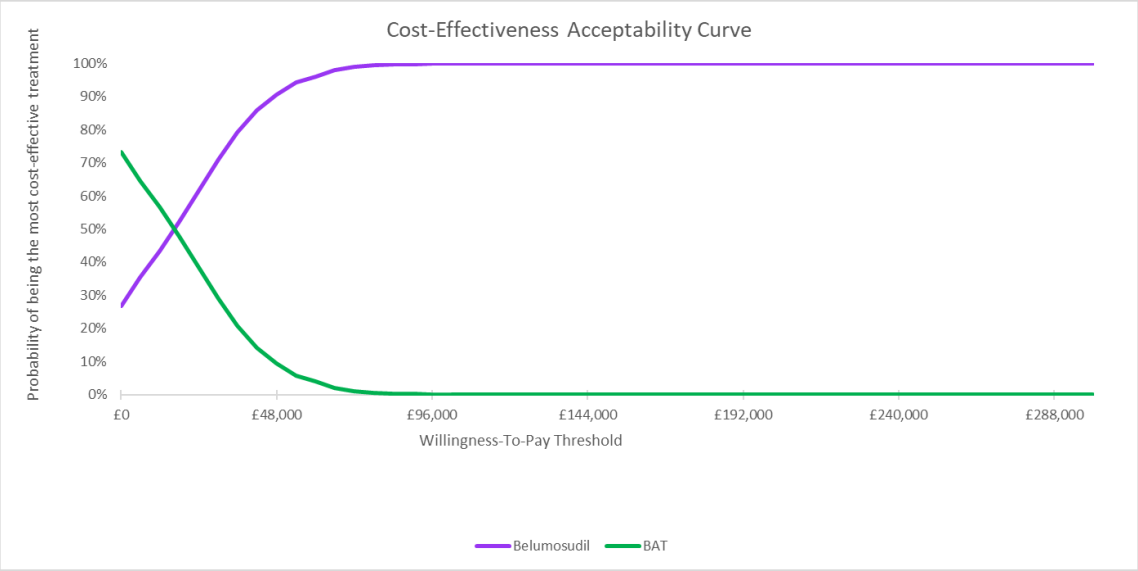
Figure 12. [redacted]

[redacted]

[redacted]

The cost-effectiveness acceptability curves are presented in Figure 13 below (Corresponding to Figure 36 in the CS). Belumosudil is more likely to be the optimal treatment choice vs. BAT at WTP thresholds above £16,500/QALY. The probability of belumosudil being the optimal treatment choice vs. BAT at a £30,000/QALY WTP threshold is [redacted]%.

Figure 13. Cost-effectiveness acceptability curves for belumosudil vs. BAT (PAS, without severity modifier)



BAT = best available therapy; PAS = patient access scheme

References

Akpek G, et al. A high-dose pulse steroid regimen for controlling active chronic graft-versus-host disease. Biology of Blood and Marrow Transplantation, 2001; 7:495-502.

Baird K, et al. Chronic graft-versus-host disease (GVHD) in children. Pediatr Clin North Am. 2010;57:297–322.

Bashey, A, et al. Improved Survival of Patients Diagnosed with Severe (Grade 3-4) Acute GVHD or Severe NIH Grade Chronic GVHD in the Current Era Compared to Historic Controls. *Blood* (2019) 134 (Supplement_1): 2006.

Crespo C, et al. Development of a population-based cost-effectiveness model of chronic graft-versus-host disease in Spain. Clin Ther. 2012;34(8):1774-87.

Dignan F, et al. Diagnosis and management of chronic graft-versus-host disease. British Journal of Haematology, 2012; 158(1):46-61.

Georgountzou A and Papadopoulos NG. Postnatal innate immune development: from birth to adulthood. Front Immunol. 2017;8:957.

Giaccone L, et al. 'Real-life' report on the management of chronic GvHD in the Gruppo Italiano Trapianto Midollo Osseo (GITMO). Bone Marrow Transplant. 2018;53:58–63.

Hartman J. e al (Hartman, J. D., Craig, B. M. (2018). Comparing and transforming PROMIS utility values to the EQ-5D. Qual Life Res. 27(3), 725-733.).

Jacobsohn, DA. Optimal management of chronic graft-versus-host disease in children. Br J Haematol. 2010;150:278–292.

Jacobsohn DA, et al. Risk factors associated with increased nonrelapse mortality and with poor overall survival in children with chronic graft-versus-host disease. Blood. 2011;118:4472-4479.

JAKAFI® (ruxolitinib) United States Prescribing Information, 2021. Available at: https://www.jakafi.com/pdf/prescribing-information.pdf.

Matza LS, Cong Z, Chung K, Stopeck A, Tonkin K, Brown J, et al. Utilities associated with subcutaneous injections and intravenous infusions for treatment of patients with bone metastases. Patient preference and adherence. 2013:855-65.

NHS England. Clinical Commissioning Policy: Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation. 16069/P. March 2017. Available at: https://www.england.nhs.uk/wp-content/uploads/2017/03/gvhd-heamatopoietic-stem-cell.pdf.

Penack O, et al. Prophylaxis and management of graft versus host disease after stem-cell transplantation for haematological malignancies: updated consensus recommendations of the European Society for Blood and Marrow Transplantation. Lancet Haematol. 2020;7:e157–e167.

Revicki DA, et al. Predicting EuroQol (EQ-5D) scores from the Patient-Reported Outcomes Measurement Information System (PROMIS) global items and domain item banks in a United States sample. Qual Life Res. 2009, 18(6), 783-91

REZUROCK (belumosudil) United States Prescribing Information, 2021. Available at: https://www.rezurock.com/full-prescribing-information.pdf.

RHOLISTIQ (belumosudil mesylate) Australian Product Information, November 2021.

Sanofi Data on File. Adelphi cGVHD Disease Specific Programmes (DSP) Analysis Tables, 2022.

Schain F, et al. Real-world study of direct medical and indirect costs and time spent in healthcare in patients with chronic graft versus host disease. The European journal of health economics : HEPAC : Health Economics in prevention and care. 2021;22(1):169-80.

Scheid C. Treatment Pathways and Health Outcomes of German Patients with Chronic Graft-Versus-Host Disease After Allogeneic Hematopoietic Cell Transplantation: A Retrospective Health Claims Data Analysis. Drugs Real World Outcomes. 2022 Dec;9(4):577-588.

TA539. National Institute for Health Care and Excellence. Single technology appraisal. Idelalisib for treating chronic lymphocytic leukaemia [ID764]. 2015. TA359.

Takahashi T, et al. Late-onset acute and chronic graft-versus-host disease in children: clinical features and response to therapy. Transplant Cell Ther. 2021;27:667.e1-667.e5.

Ward LM. Glucocorticoid-induced osteoporosis: why kids are different. Front Endocrinol (Lausanne). 2020;11:576.

Wolff D, Schleuning M, von Harsdorf S, et al. Consensus Conference on Clinical Practice in Chronic GVHD: Second-Line Treatment of Chronic Graft-versus-Host Disease. Biol Blood Marrow Transplant. 2011;17:1–17.

Yoshihiro I, et al. Failure-free survival after second-line systemic treatment of chronic graft-versus-host disease. Blood (2013) 121 (12): 2340–2346.

UK-WHO, Growth Charts 2-18 Years Old, 2012. Available at: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-2-18-years.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Clarification questions

May 2023

File name	Version	Contains confidential information	Date
ID4021_Belumosudil additional clarification questions_120523_ACIC	V1.0	Yes	12th May 2023

Section A: Clarification on effectiveness data

A1. Priority question: In the company submission (CS), the company stated a later data cut of ROCKstar was available in late 2022, but the economic model uses data from August 2021. Please update the pooled analysis for the ≥2 prior lines of therapy subgroup, and thus update the economic model, with the 2022 data cut.

The company asked for guidance from the EAG about what might be considered critical deliverables and an acceptable timeframe for their delivery.

We acknowledge these responses come a little later than the requested timeframe of 1 week but the answers required significant additional biostatistical analysis. We hope the EAG is able to accommodate this slight delay.

A) Please provide a table of baseline characteristics for the pooled ROCKstar and KD025-208 (≥ 2 prior lines of therapy subgroup) using the September data cut.

Table 1. Baseline characteristics for pooled ROCKstar and KD025-208 population (≥ 2 prior lines of therapy subgroup; 2022 data)

Baseline characteristic	200 mg once daily n=92	200 mg twice daily n=84	Combined 200 mg N=176
Median age (range), years	53.0 (20-77)	57.0 (18-77)	55.0 (18-77)
Males, n (%)	60 (65.2%)	42 (50.0%)	102 (58.0%)
GVHD prophylaxis after transplant, n (%)			
None	0	1 (1.2%)	1 (0.6%)
CNI only	5 (5.4%)	7 (8.3%)	12 (6.8%)
CNI + methotrexate	38 (41.3%)	35 (41.7%)	73 (41.5%)
CNI + methotrexate + other	10 (10.9%)	7 (8.3%)	17 (9.7%)
CNI + MMF	11 (12.0%)	14 (16.7%)	25 (14.2%)
CNI + MMF + other	5 (5.4%)	3 (3.6%)	8 (4.5%)
CNI + MMF + ATG	0	1 (1.2%)	1 (0.6%)
CNI + sirolimus	8 (8.7%)	8 (9.5%)	16 (9.1%)
CNI + corticosteroids	2 (2.2%)	1 (1.2%)	3 (1.7%)
Other regimen	12 (13.0%)	7 (8.3%)	19 (10.8%)
HLA matching of donor/recipient, n (%)			
Matched	80 (87.0%)	78 (92.9%)	158 (89.8%)
Partially matched	11 (12.0%)	5 (6.0%)	16 (9.1%)
Unknown	0	1 (1.2%)	1 (0.6%)
Missing	1 (1.1%)	0	1 (0.6%)
History of acute GVHD, n (%)	61 (66.3%)	63 (75.0%)	124 (70.5%)
Time from chronic GVHD diagnosis to enrolment, median (range), months	26.66 (1.6- 162.4)	29.91 (3.7- 144.1)	28.14 (1.6-162.4)
NIH chronic GVHD severity[a] n (%)			

Baseline characteristic	200 mg once daily n=92	200 mg twice daily n=84	Combined 200 mg N=176
Severe	66 (71.7%)	58 (69.0%)	124 (70.5%)
Moderate	24 (26.1%)	26 (31.0%)	50 (28.4%)
Mild	2 (2.2%)	0	2 (1.1%)
Organ involvement, n (%)			
≥4 organs involved	49 (53.3%)	44 (52.4%)	93 (52.8%)
≥6 organs involved	15 (16.3%)	10 (11.9%)	25 (14.2%)
Eyes	68 (73.9%)	59 (70.2%)	127 (72.2%)
Skin	75 (81.5%)	69 (82.1%)	144 (81.8%)
Mouth	52 (56.5%)	56 (66.7%)	108 (61.4%)
Joints/fascia	70 (76.1%)	63 (75.0%)	133 (75.6%)
Lungs	32 (34.8%)	27 (32.1%)	59 (33.5%)
Upper GI	16 (17.4%)	11 (13.1%)	27 (15.3%)
Oesophagus	25 (27.2%)	13 (15.5%)	38 (21.6%)
Lower GI	8 (8.7%)	8 (9.5%)	16 (9.1%)
Liver	10 (10.9%)	5 (6.0%)	15 (8.5%)
No. of organs involved, median (range)	4.0 (0-7)	4.0 (1-7)	4.0 (0-7)
Refractory to prior LOT, n (%)	60 (80.0%)	43 (65.2%)	103 (73.0%)
Number or prior lines of therapy, n (%)			
1	0	0	0
2	29 (31.5%)	18 (21.4%)	47 (26.7%)
3	27 (29.3%)	26 (31.0%)	53 (30.1%)
4	20 (21.7%)	18 (21.4%)	38 (21.6%)
5	14 (15.2%)	20 (23.8%)	34 (19.3%)
≥6	2 (2.2%)	2 (2.4%)	4 (2.3%)
Median	3.0	3.0	3.0
Prior systemic chronic GVHD therapies, n (%)[b]			
Prednisone	91 (98.9%)	83 (98.8%)	174 (98.9%)
Tacrolimus	56 (60.9%)	53 (63.1%)	109 (61.9%)
Sirolimus	43 (46.7%)	41 (48.8%)	84 (47.7%)
ECP	5 (5.4%)	4 (4.8%)	9 (5.1%)
Ibrutinib	27 (29.3%)	27 (32.1%)	54 (30.7%)
Mycophenolate mofetil	22 (23.9%)	22 (26.2%)	44 (25.0%)
Rituximab	23 (25.0%)	16 (19.0%)	39 (22.2%)
Ruxolitinib	29 (31.5%)	26 (31.0%)	55 (31.3%)
Concomitant systemic chronic GVHD therapies, n (%)			
Systemic hormonal preparations	91 (98.9%)	82 (97.6%)	173 (98.3%)
ECP	22 (23.9%)	28 (33.3%)	50 (28.4%)
Tacrolimus	30 (32.6%)	26 (31.0%)	56 (31.8%)
Sirolimus	19 (20.7%)	20 (23.8%)	39 (22.2%)
MMF	11 (12.0%)	3 (3.6%)	14 (8.0%)
Imatinib	1 (1.1%)	1 (1.2%)	2 (1.1%)
Rituximab	1 (1.1%)	0	1 (0.6%)

Baseline characteristic	200 mg once daily n=92	200 mg twice daily n=84	Combined 200 mg N=176
Ruxolitinib	1 (1.1%)	0	1 (0.6%)

[a] Severity was determined using the NIH Global Severity of chronic GVHD scoring
[b] This table includes the most common therapies for chronic GVHD (≥10%), as well as ECP
ATG = antithymocyte globulin, CNI = calcineurin inhibitor; ECP = extracorporeal photopheresis, GI = gastrointestinal; GVHD = graft-versus-host disease; HLA = human leukocyte antigen; MMF = mycophenolate mofetil; NIH = National Institutes of Health

B) Please provide KM plots of overall survival (OS), failure-free survival (FFS), time to response (TTR) and duration of response (DOR) in ROCKstar and KD025-208 (≥ 2 prior lines of therapy subgroup) studies across data cuts. Please provide plots for each study and also for the pooled analysis. These plots should include numbers at risk.

Overall survival (OS)

Table 2. KM plot of overall survival in ROCKstar (2022 data cut)

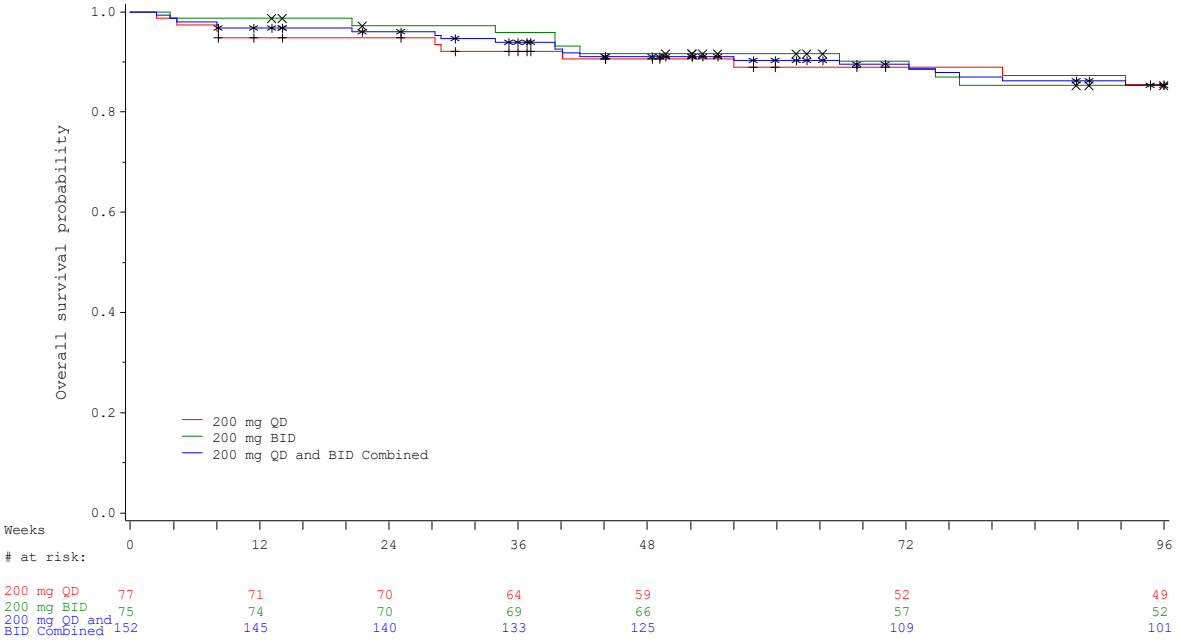


Table 3. KM plot of overall survival in KD025-208 (≥2 prior lines of therapy; 2022 data cut)

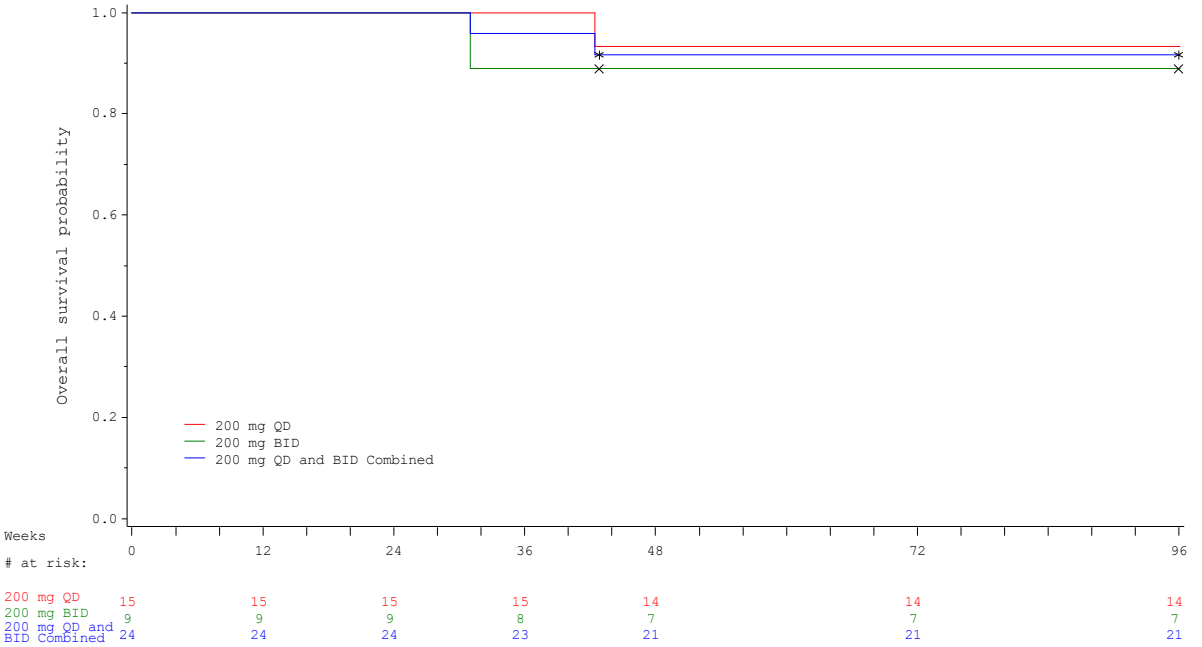
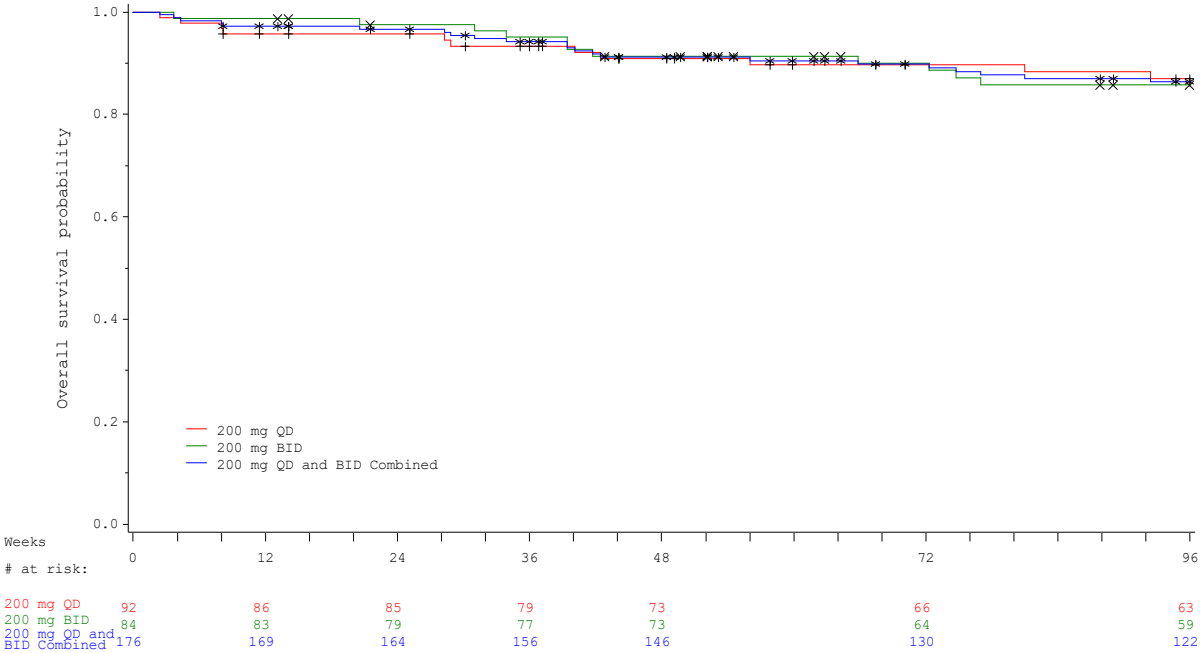
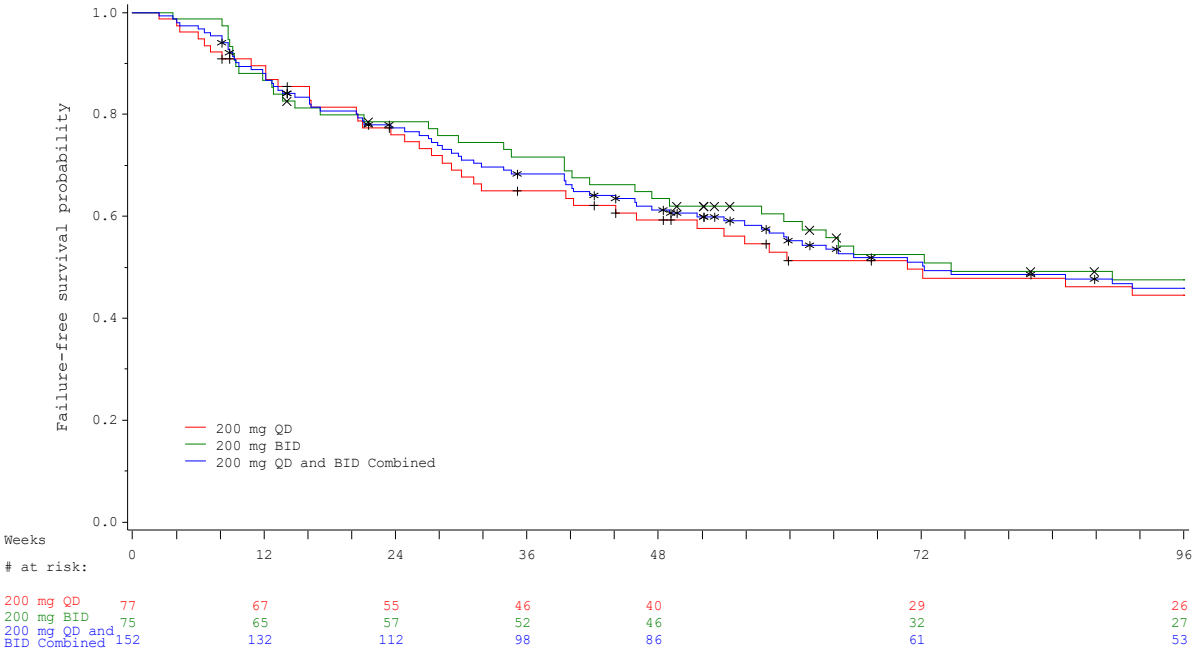


Table 4. KM plot of overall survival in pooled analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy; 2022 data cut)



Failure-free survival (FFS)

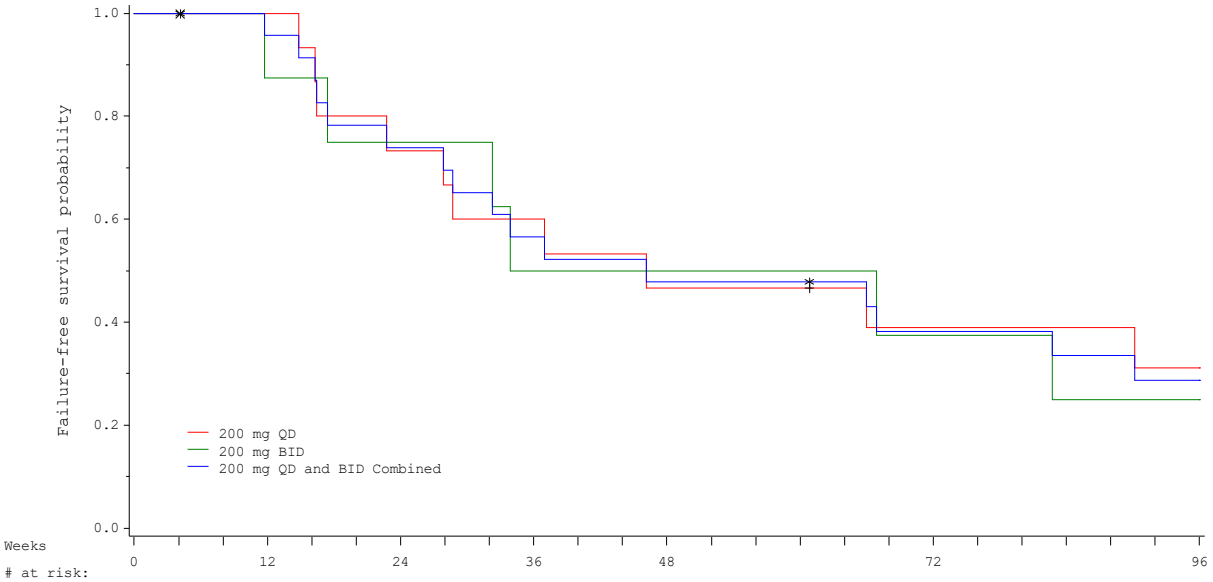
Table 5. KM plot of failure-free survival in ROCKstar (2022 data cut)



at risk:

200 mg QD	77	67	55	46	40	29	26
200 mg BID	75	65	57	52	46	32	27
200 mg QD and BID Combined	152	132	112	98	86	61	53

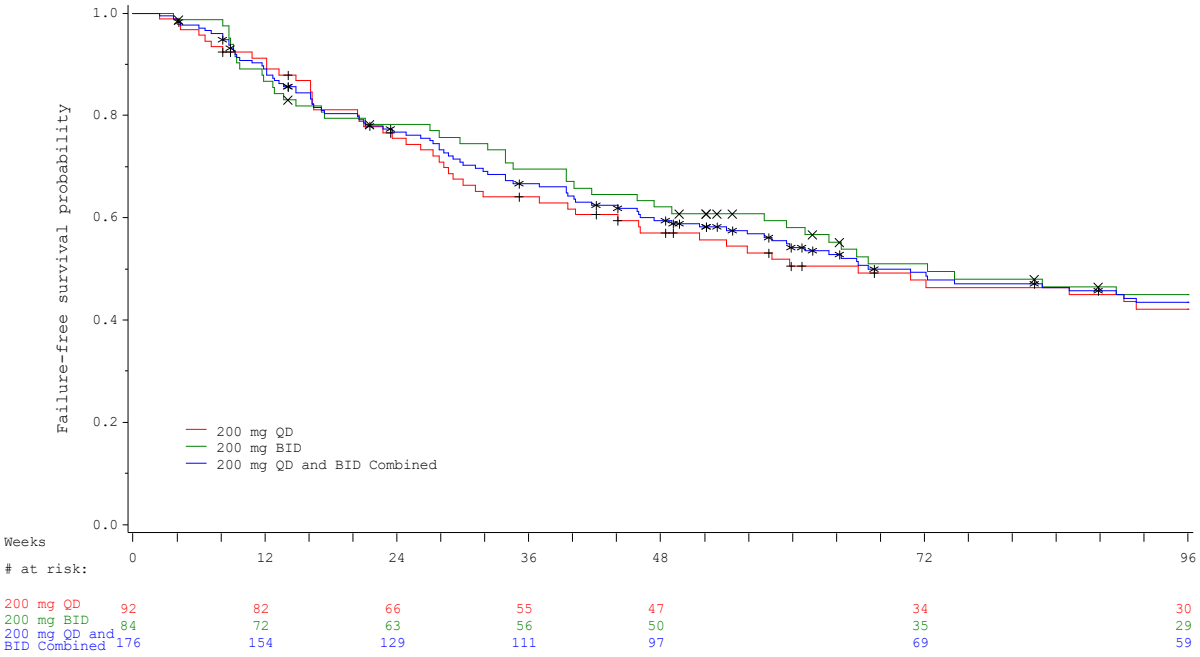
Table 6. KM plot of failure-free survival in KD025-208 (≥2 prior lines of therapy; 2022 data cut)



at risk:

200 mg QD	15	15	11	9	7	5	4
200 mg BID	9	7	6	4	4	3	2
200 mg QD and BID Combined	24	22	17	13	11	8	6

Table 7. KM plot of failure-free survival in pooled analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy; 2022 data cut)



Time to response (TTR)

Table 8. KM plot of time to response in ROCKstar (2022 data cut)

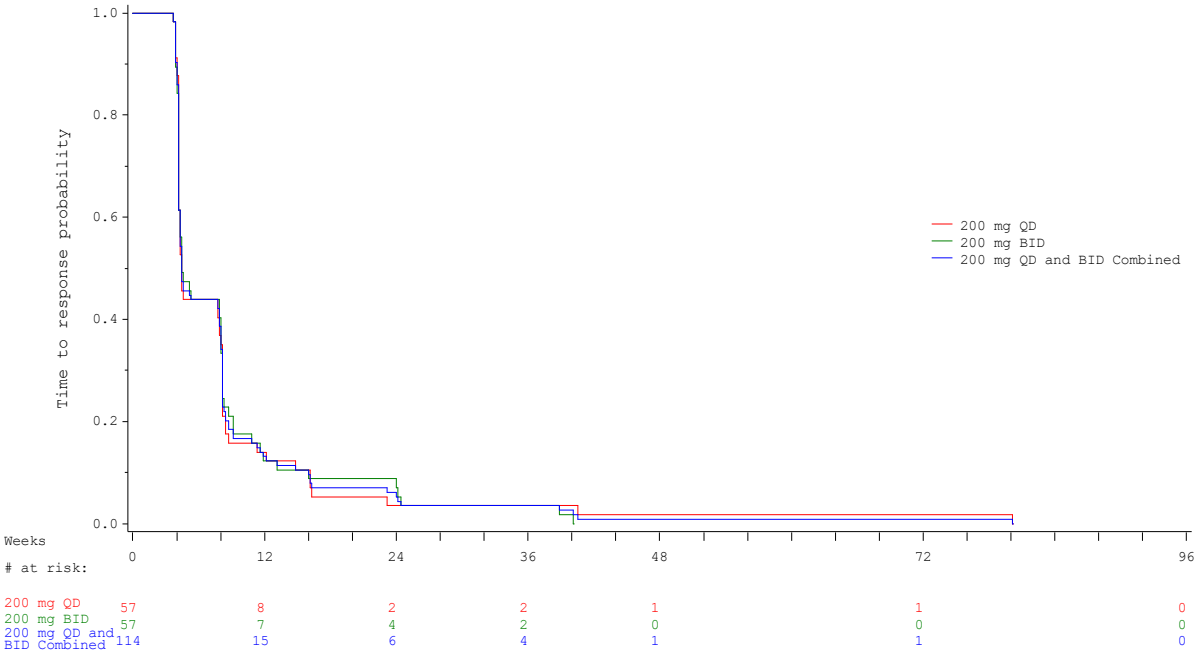


Table 9. KM plot of time to response in KD025-208 (≥2 prior lines of therapy; 2022 data cut)

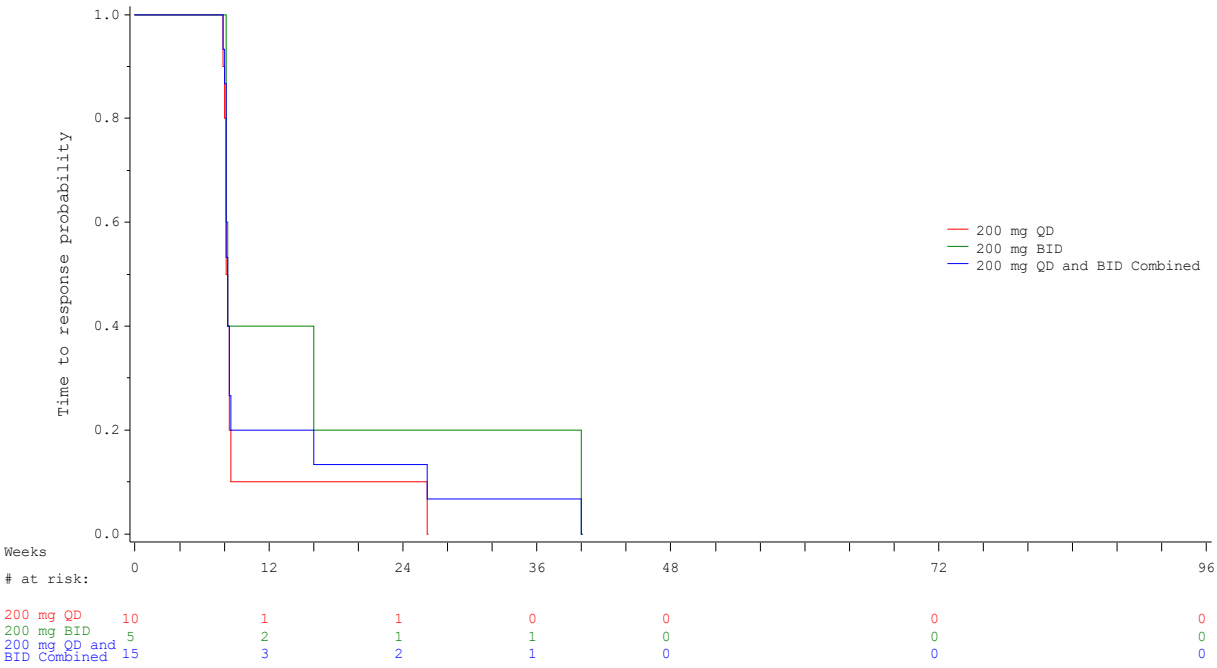
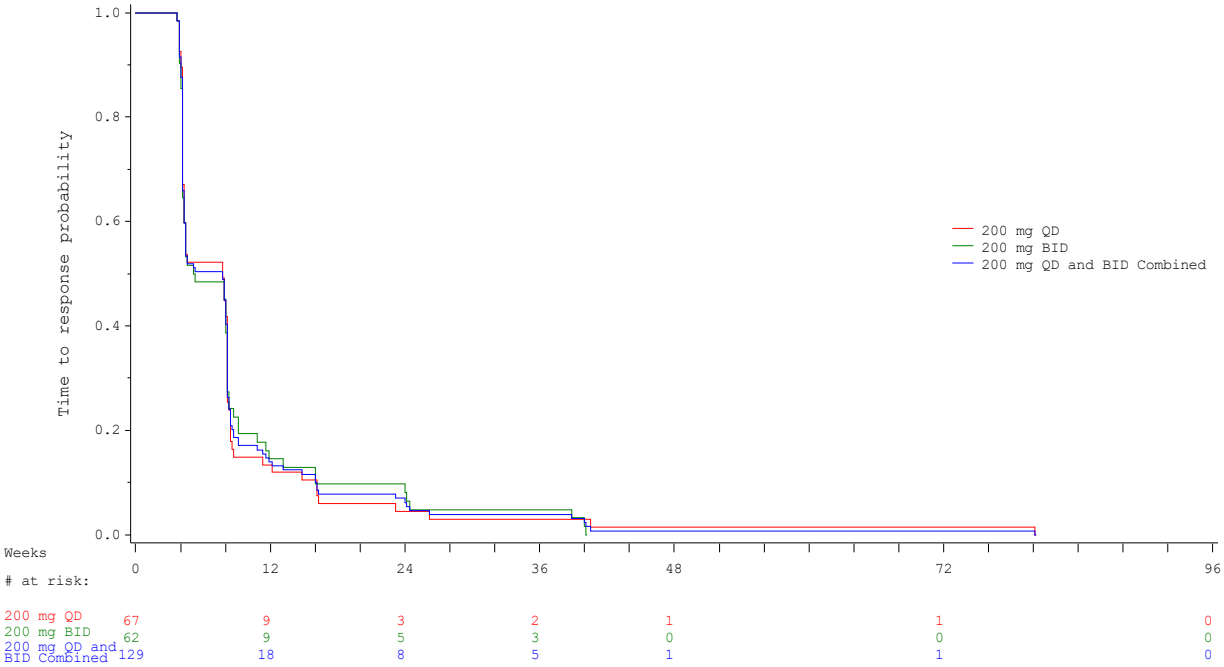


Table 10. KM plot of time to response in pooled analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy; 2022 data cut)



Duration of response (DOR)

Table 11. KM plot of duration of response in ROCKstar (2022 data cut)

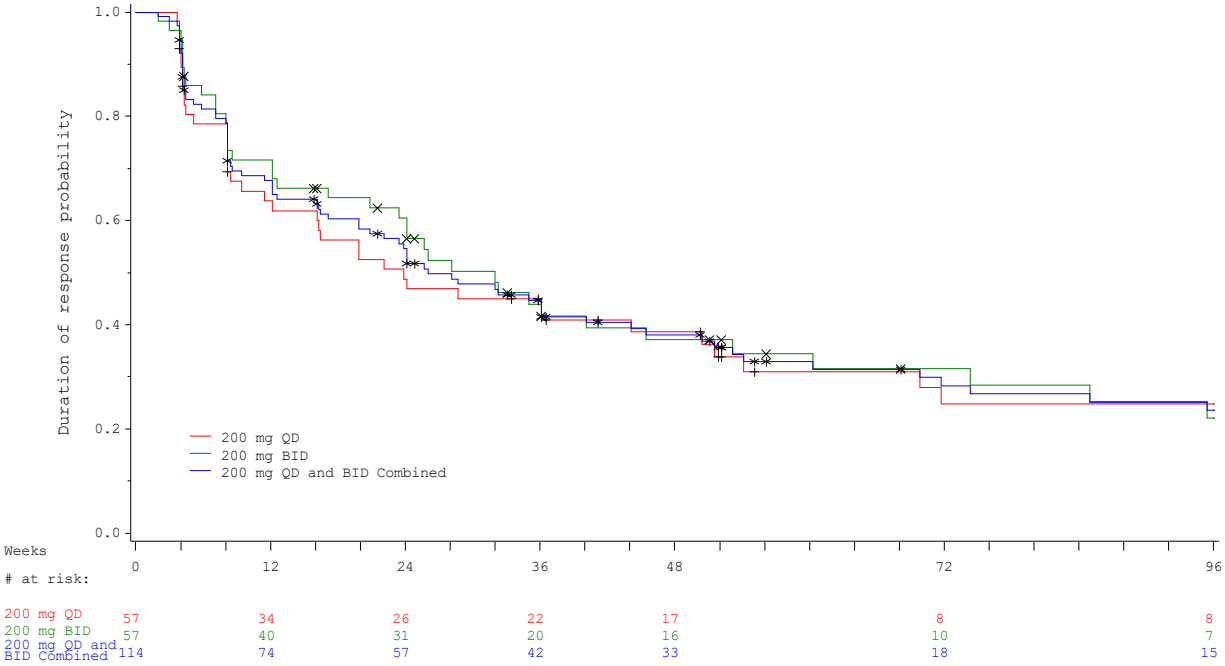


Table 12. KM plot of duration of response in KD025-208 (≥2 prior lines of therapy; 2022 data cut)

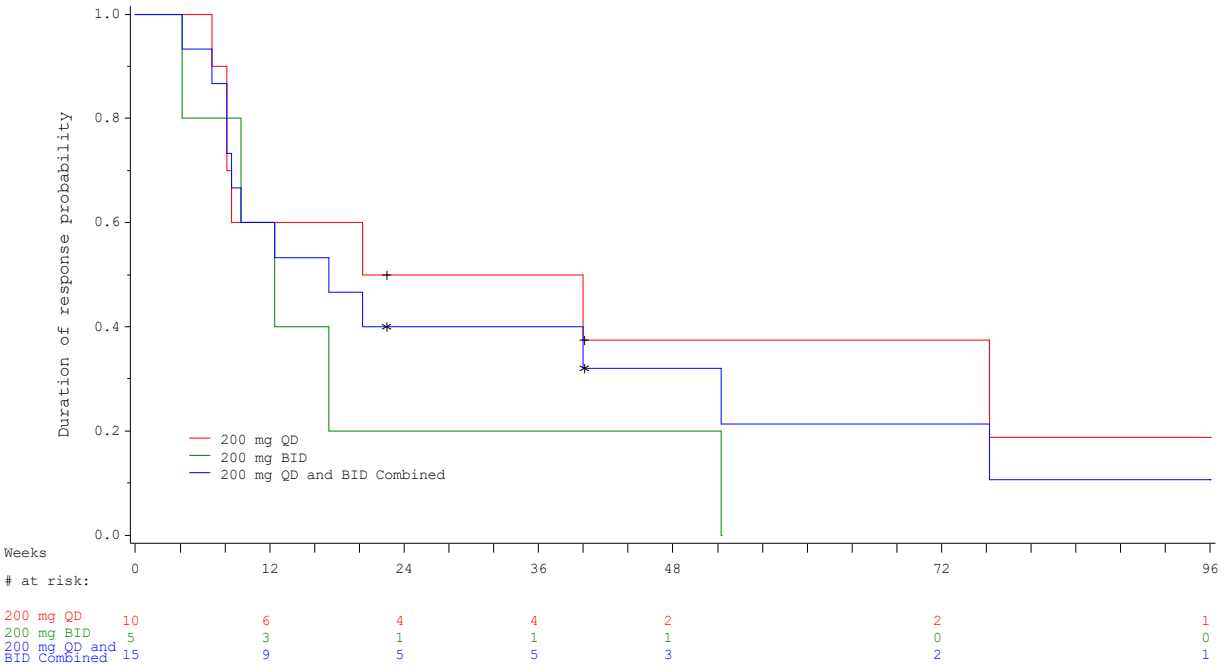
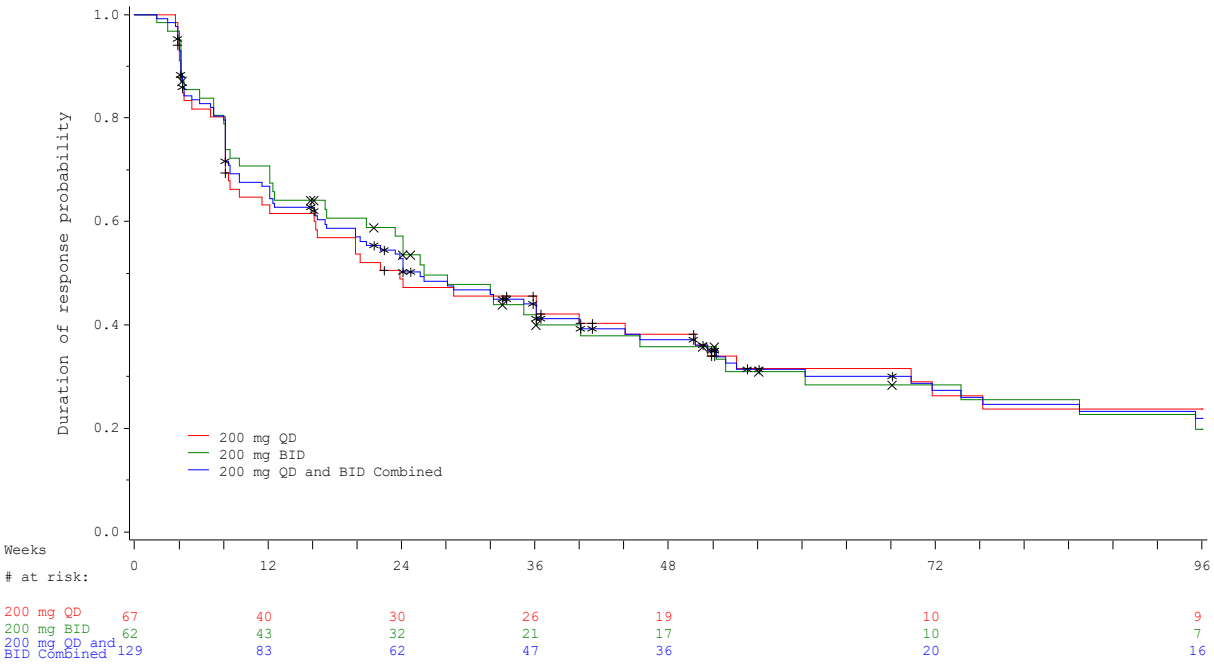


Table 13. KM plot of duration of response in pooled analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy; 2022 data cut)



C) Please fill these tables, previously provided in the CS for the August 2021 data cut, for the September 2022 data cut:

Table 14. Results of pooled efficacy analysis (≥2 prior lines of therapy; 2022 data)

	September 2022 data cut		
	200 mg once daily (n=92)	**200 mg twice daily (n=84)**	**Combined 200 mg (N=176)**
Median time to response, weeks (range)	7.71 (3.7-80.1)	5.21 (3.7-40.1)	7.71 (3.7-80.1)
Best ORR,[a] n (%)			
CR	4 (4.3%)	2 (2.4%)	6 (3.4%) 123 (69.9%)
PR	63 (68.5%)	60 (71.4%)	
Best response by organ system, n/N (%)			
Skin	23 / 75 (30.7%)	29 / 69 (42.0%)	52 / 144 (36.1%)
Eyes	28 / 68 (41.2%)	28 / 59 (47.5%)	56 / 127 (44.1%)
Mouth	28 / 52 (53.8%)	32 / 59 (57.1%)	60 / 108 (55.6%)
Oesophagus	14 / 25 (56.0%)	7 / 13 (53.8%)	21 / 38 (55.3%)
Upper GI	11 / 16 (68.8%)	5 / 11 (45.5%)	16 / 27 (59.3%)
Lower GI	5 / 8 (62.5%)	6 / 8 (75.0%)	11 / 16 (68.8%)
Liver	2 / 10 (20.0%)	1 / 5 (20.0%)	3/ 15 (20.0%)
Lungs	9 / 32 (28.1%)	6 / 27 (22.2%)	15 / 59 (25.4%)
Joints/fascia	48 / 92 (68.6%)	42 / 63 (66.7%)	90 / 133 (67.7%)
Median DOR in responders (primary/secondary)[b] weeks (95% CI)	23.9 (12.14, 50.43)	26.0 (17.14, 45.43)	25.7 (17.29, 36.14)
Median DOR in responders (quaternary), weeks (95% CI)	69.9 (28.29, 176.00)	74.3 (35.00, 114.57)	69.9 (40.43, 95.43)
Median FFS, months (95% CI)	15.2 (9.26, 24.02)	16.6 (11.27, 35.88)	15.4 (12.42, 22.74)
FFS, % (95% CI)			
FFS at 6 months	74% (0.64,0.82)	78% (0.68, 0.86)	76% (0.69, 0.82)
FFS within 12 months	56% (0.45,0.65)	61% (0.49, 0.70)	58% (0.50, 0.65)
FFS within 24 months	41% (0.30,0.51)	42% (0.31, 0.53)	41% (0.33, 0.49)
Median OS (months)	n/a	n/a	n/a
OS, % (95% CI)			
OS within 12 months	0.91 (0.83,0.95)	0.91 (0.83,0.96)	0.91 (0.86,0.95)
OS within 24 months	0.86 (0.76,0.92)	0.84 (0.74,0.91)	0.85 (0.78,0.90)
Median TTD , months (range)	9.18 (0.5, 64.2)	11.78 (0.4, 39.6)	10.38 (0.4 , 64.2)

Table 15. Safety profile from pooled analysis (safety population; ≥2 prior lines of therapy; 2022 data)

	September 2022 data cut			
	200 mg once daily (n=92)	**200 mg twice daily (n=84)**	**400 mg once daily (n=14)**	**Total (N=190)**
Any AE, n (%)	91 (98.9%)	84 (100.0%)	14 (100.0%)	189 (99.5%)
Any drug-related AE, n (%)	66 (71.7%)	55 (65.5%)	11 (78.6%)	132 (69.5%)
Grade ≥3 AEs, n (%)	59 (64.1%)	48 (57.1%)	10 (71.4%)	117 (61.6%)
Drug-related Grade ≥3 AEs, n (%)	18 (19.6%)	16 (19.0%)	2 (14.3%)	36 (18.9%)
SAE, n (%)	41 (44.6%)	34 (40.5%)	8 (57.1%)	83 (43.7%)
Drug-related SAE, n (%)	7 (7.6%)	5 (6.0%)	0	12 (6.3%)
Fatal AEs, n (%)	5 (5.4%)	5 (6.0%)	4 (28.6%)	14 (7.4%)
Infections and infestations (any grade), n (%)	57 (62.0%)	57 (67.9%)	9 (64.3%)	123(64.7%)
Grade ≥3, n (%)	18 (19.6%)	20 (23.8%)	3 (21.4%)	41 (21.6%)
Cytopenias[a]	16 (17.4%)	20 (23.8%)	1 (7.1%)	37 (19.5%)
Most common AEs (incidence ≥25%)				
Fatigue	39 (42.4%)	26 (31.0%)	8 (57.1%)	73 (38.4%)
Diarrhoea	41 (44.6%)	27 (32.1%)	4 (28.6%)	72 (37.9%)
Upper respiratory tract infection	26 (28.3%)	28 (33.3%)	4 (28.6%)	58 (30.5%)
Nausea	31 (33.7%)	25 (29.8%)	5 (35.7%)	61 (32.1%)
Dyspnoea	28 (30.4%)	20 (23.8%)	5 (35.7%)	53 (27.9%)
Cough	23 (25.0%)	23 (27.4%)	6 (42.9%)	52 (27.4%)
Oedema peripheral	25 (27.2%)	20 (23.8%)	4 (28.6%)	49 (25.8%)
Headache	26 (28.3%)	24 (28.6%)	3 (21.4%)	53 (27.9%)
Vomiting	25 (27.2%)	14 (16.7%)	2 (14.3%)	41 (21.6%)
Muscle spasms	13 (14.1%)	15 (17.9%)	3 (21.4%)	31 (16.3%)

D) Please provide the total number of responses to the PROMIS-GH questionnaire (total number of observations) based on the September 2022 data cut from ROCKstar that inform the utility regression analyses.

Table 16. PROMIS-GH Mental Health score from ROCKstar (2022 data)

PROMIS-GH Mental Health Score	**200 mg once daily (n=77)**	**200 mg twice daily (n=75)**	**Overall (n=152)**
Baseline, n	[redacted]	[redacted]	[redacted]
Cycle 2, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 3, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 4, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 5, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 7, day 1, n	[redacted]	[redacted]	[redacted]

Cycle 9, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 11, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 13, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 15, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 17, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 19, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 21, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 23, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 25, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 27, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 29, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 31, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 33, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 35, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 37, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 39, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 41, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 43, day 1, n	[redacted]	[redacted]	[redacted]
EOT, n	[redacted]	[redacted]	[redacted]

Table 17. PROMIS-GH Physical Health score from ROCKstar (2022 data)

PROMIS-GH Physical Health Score	200 mg once daily (n=77)	200 mg twice daily (n=75)	Overall (n=152)
Baseline, n	[redacted]	[redacted]	[redacted]
Cycle 2, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 3, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 4, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 5, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 7, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 9, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 11, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 13, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 15, day 1, n	[redacted]	[redacted]	[redacted]

Cycle 17, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 19, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 21, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 23, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 25, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 27, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 29, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 31, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 33, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 35, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 37, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 39, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 41, day 1, n	[redacted]	[redacted]	[redacted]
Cycle 43, day 1, n	[redacted]	[redacted]	[redacted]
EOT, n	[redacted]	[redacted]	[redacted]

A12. Priority question: Participants in ROCKstar and KD025-208, who were on-treatment, were assessed for response in 28-day treatment cycles. Based on the pooled and combined analysis of ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup), and using the 2022 data cut (if available), please provide details of the number of patients with FFS at defined timepoints during the study and the response or lack response to treatment within the FFS group. This is the response at the specified timepoint rather than best response at any prior timepoint.

A) This is a query on the response received for question A12. We requested the number of patients with FFS at 6 monthly time points through the study, and their response, or lack of response, at that time point. At the start of the study all patients have FFS and this decreases as patients start a new systemic cGvHD therapy, have non-relapse mortality or recurrent malignancy. However, the tables provided at clarification (10, 11, and 12) indicate an increase in the number of patients with FFS as the study progresses. Please provide updated tables with corrected proportions.

Table 18. Pooled and combined ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) (2022 data cut)

	Pooled and combined ROCKstar and KD025-208 belumosudil (N=176)						
	Patients in FFS, n (%)	**Responders, n (%)**	**CR, n (%)**	**PR, n (%)**	**LR-U, n (%)**	**LR-M, n (%)**	**LR-P, n (%)**
6 months	135 (76.7%)	88 (65.2%)	2 (1.5%)	86 (63.7%)	19 (14.1%)	18 (13.3%)	9 (6.7%)
12 months	105 (59.7%)	74 (70.5%)	3 (2.9%)	71 (67.6%)	17 (16.2%)	11 (10.5%)	3 (2.9%)
18 months	89 (50.6%)	61 (68.5%)	4 (4.5%)	57 (64.0%)	12 (13.5%)	13 (14.6%)	3 (3.4%)
24 months	81 (46.0%)	56 (69.1%)	2 (2.5%)	54 (66.7%)	9 (11.1%)	14 (17.3%)	2 (2.5%)
30 months	77 (43.8%)	47 (61.0%)	3 (3.9%)	44 (57.1%)	10 (13.0%)	18 (23.4%)	2 (2.6%)
36 months	74 (42.0%)	49 (66.2%)	3 (4.1%)	46 (62.2%)	9 (12.2%)	14 (18.9%)	2 (2.7%)
42 months	74 (42.0%)	50 (67.6%)	3 (4.1%)	47 (63.5%)	9 (12.2%)	13 (17.6%)	2 (2.7%)
48 months	74 (42.0%)	49 (66.2%)	3 (4.1%)	46 (62.2%)	9 (12.2%)	14 (18.9%)	2 (2.7%)

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 19. Pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) belumosudil 200 mg once daily (2022 data cut)

	Pooled ROCKstar and KD025-208 belumosudil 200 mg once daily (n=92)						
	Patients in FFS, n (%)	**Responders, n (%)**	**CR, n (%)**	**PR, n (%)**	**LR-U, n (%)**	**LR-M, n (%)**	**LR-P, n (%)**
6 months	69 (75.0%)	43 (62.3%)	1 (1.4%)	42 (60.9%)	10 (14.5%)	9 (13.0%)	7 (10.1%)
12 months	53 (57.6%)	36 (67.9%)	2 (3.8%)	34 (64.2%)	7 (13.2%)	8 (15.1%)	2 (3.8%)
18 months	46 (50.0%)	30 (65.2%)	2 (4.3%)	28 (60.9%)	5 (10.9%)	8 (17.4%)	3 (6.5%)
24 months	42 (45.7%)	28 (66.7%)	1 (2.4%)	27 (64.3%)	4 (9.5%)	8 (19.0%)	2 (4.8%)
30 months	39 (42.4%)	24 (61.5%)	1 (2.6%)	23 (59.0%)	4 (10.3%)	9 (23.1%)	2 (5.1%)
36 months	37 (40.2%)	25 (67.6%)	1 (2.7%)	24 (64.9%)	4 (10.8%)	6 (16.2%)	2 (5.4%)
42 months	37 (40.2%)	26 (70.3%)	1 (2.7%)	25 (67.6%)	4 (10.8%)	5 (13.5%)	2 (5.4%)
48 months	37 (40.2%)	25 (67.6%)	1 (2.7%)	24 (64.9%)	4 (10.8%)	6 (16.2%)	2 (5.4%)

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

Table 20. Pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) belumosudil 200 mg twice daily (2022 data cut)

	Pooled ROCKstar and KD025-208 belumosudil 200 mg twice daily (n=84)						
	Patients in FFS, n (%)	**Responders, n (%)**	**CR, n (%)**	**PR, n (%)**	**LR-U, n (%)**	**LR-M, n (%)**	**LR-P, n (%)**
6 months	66 (78.6%)	45 (68.2%)	1 (1.5%)	44 (66.7%)	22 (59.5%)	9 (13.6%)	2 (3.0%)
12 months	52 (61.9%)	38 (73.1%)	1 (1.9%)	37 (71.2%)	10 (19.2%)	3 (5.8%)	1 (1.9%)
18 months	43 (51.2%)	31 (72.1%)	2 (4.7%)	29 (67.4%)	7 (16.3%)	5 (11.6%)	0
24 months	39 (46.4%)	28 (71.8%)	1 (2.6%)	27 (69.2%)	5 (12.8%)	6 (15.4%	0
30 months	38 (45.2%)	23 (60.5%)	2 (5.3%)	21 (55.3%)	6 (15.8%)	9 (23.7%)	0
36 months	37 (44.0%)	24 (64.9%)	2 (5.4%)	22 (59.5%)	5 (13.5%)	8 (21.6%)	0
42 months	37 (44.0%)	24 (64.9%)	2 (5.4%)	22 (59.5%)	5 (13.5%)	8 (21.6%)	0
48 months	37 (44.0%)	24 (64.9%)	2 (5.4%)	22 (59.5%)	5 (13.5%)	8 (21.6%)	0

CR=complete response; PR=partial response; LR-M=lack of response-mixed; LR-P=lack of response-progression; LR-U=lack of response-unchanged

A17. Priority question: In Section B.3.12, the company state: "Based on the above pharmacokinetics and clinical trial outcomes evidence, we do not expect belumosudil to be more clinically or cost effective than BAT in the subgroups suggested in the scope and therefore no subgroups have been considered in this submission." Please can the company clarify the rationale for this conclusion regarding the efficacy of belumosudil in the subgroups suggested in the scope?

We think this may be an issue with how we originally worded the phrase cited. Our intention was not to suggest that belumosudil is not cost effective in these subgroups. However we acknowledge that the phrase: '*we do not expect belumosudil to be more clinically or cost effective than BAT in the subgroups*' could be interpreted as such.

To be clear, we do not expect belumosudil to be more or less cost effective than the base case when subgroups are considered. This is because the clinical and PK data suggest it is equally effective across the prespecified subgroups analysed. Hence we believe the full population should be considered in the analysis.

A18. Please provide data for following outcomes in the ROCKstar and KD025-208 (≥2L) trials. If data are not available at week 24 then provide the nearest available time point in each trial.

- Response at week 24 by organ;
- Best ORR by week 24 by organ.

Table 21. ROCKstar (2022 data cut) Response by organ at week 24

Organ	Rockstar (n=152) (2022 data cut)		
	200mg QD (N=77)	200mg BID (N=75)	Overall (N=152)
	Responders, n (%)	Responders, n (%)	Responders, n (%)
Skin	12 (19.0%)	15 (23.8%)	27 (21.4%)
Joints/fascia	29 (49.2%)	27 (47.4%)	56 (48.3%)
Eyes	9 (16.1%)	16 (30.8%)	25 (23.1%)
Mouth	13 (31.7%)	19 (37.3%)	32 (34.8%)
Lungs	3 (11.1%)	1 (4.0%)	4 (7.7%)
Oesophagus	7 (30.4%)	4 (30.8%)	11 (30.6%)
Upper GI	8 (57.1%)	4 (40.0%)	12 (50.0%)
Lower GI	4 (57.1%)	5 (62.5%)	9 (60.0%)
Liver	1 (10.0%)	0	1 (7.1%)

Table 22. ROCKstar (2022 data cut) Best ORR by organ by week 24

Organ	Rockstar (n=152) (2022 data cut)		
	200mg QD (N=77)	200mg BID (N=75)	Overall (N=152)
	Responders, n (%)	Responders, n (%)	Responders, n (%)
Skin	19 (30.2%)	22 (34.9%)	41 (32.5%)
Joints/fascia	42 (71.2%)	38 (66.7%)	80 (69.0%)
Eyes	16 (28.6%)	22 (42.3%)	38 (35.2%)
Mouth	22 (53.7%)	28 (54.9%)	50 (54.3%)
Lungs	5 (18.5%)	4 (16.0%)	9 (17.3%)
Oesophagus	11 (47.8%)	7 (53.8%)	18 (50.0%)
Upper GI	9 (64.3%)	4 (40.0%)	13 (54.2%)

Lower GI	5 (71.4%)	6 (75.0%)	11 (73.3%)
Liver	2 (20.0%)	0	2 (14.3%)

Table 23. KD25-208 (≥2 lines of prior therapy subgroup) (2022 data cut) Response by organ at week 24

Organ	KD025-208 (≥2 lines of prior therapy subgroup) (N=24) (2022 data cut)		
	200mg QD (N=15)	200mg BID (N=9)	Combined 200mg (N=24)
	Responders, n (%)	Responders, n (%)	Responders, n (%)
Skin	2 (16.7%)	0	2 (11.1%)
Joints/fascia	3 (27.3%)	1 (16.7%)	4 (23.5%)
Eyes	0	1 (14.3%)	1 (5.3%)
Mouth	3 (27.3%)	1 (20.0%)	4 (25.0%)
Lungs	0	0	0
Oesophagus	1 (50.0%)	0	1 (50.0%)
Upper GI	1 (50.0%)	0	1 (33.3%)
Lower GI	0	0	0
Liver	0	0	0

Table 24. KD025-208 (≥2 lines of prior therapy subgroup) (2022 data cut) Best ORR by organ by week 24

Organ	KD25-208 (≥2 lines of prior therapy subgroup) (N=24) (2022 data cut)		
	200mg QD (N=15)	200mg BID (N=9)	Combined 200mg (N=24)
	Responders, n (%)	Responders, n (%)	Responders, n (%)
Skin	2 (16.7%)	1 (16.7%)	3 (16.7%)
Joints/fascia	5 (45.5%)	3 (50.0%)	8 (47.1%)
Eyes	3 (25.0%)	2 (28.6%)	5 (26.3%)
Mouth	5 (45.5%)	1 (20.0%)	6 (37.5%)
Lungs	0	0	0
Oesophagus	1 (50.0%)	0	1 (50.0%)

Upper GI	2 (100.0%)	1 (100.0%)	3 (100.0%)
Lower GI	0	0	0
Liver	0	0	0

A19. Priority question: The final scope requests subgroup analyses by different organs or tissues affected by cGVHD should be conducted, if evidence allows. Response by organ at week 24 is presented in table S4 in the REACH-3 trial.[1] This is response at week 24 rather than best ORR at week 24.

Please provide subgroup analysis (tables and forest plots) comparing the response by organ at week 24 in the ROCKstar and KD025-208 (≥2L) trials (pooled data) to the REACH-3 trial. If data are not available at week 24 in the belumosudil trials, please use data from the nearest available time point. Please also offer an interpretation of the results.

We have not produced Forest plots as the data constitutes a naïve indirect comparison and this would therefore not be appropriate. REACH-3 data have been included in Table 25 below as well as belumosudil pooled analysis results only for the purposes of naïve side by side comparison and should not be overinterpreted for the reasons discussed below.

It is critical to consider these organ-specific responses in the context of the highly heterogenous nature of cGVHD pathogenesis and clinical presentation. Most patients receiving third line or later treatment for cGVHD will have multiple organs affected, with some organ manifestations affecting quality of life more than others (in the pooled ROCKstar and KD025-208 studies [≥2 lines of prior therapy subgroup] the median number of organs involved at baseline was 4.0). Therefore, evaluating response on an individual organ basis, as in Table 25, does not accurately reflect either the baseline manifestation or the individual patient benefit. Neither does it necessarily represent the primary organ impacted. For example, a patient showing lack of response of skin cGVHD may experience a complete response of eyes and mouth cGVHD, which could be a successful outcome for the individual resulting in significant quality of life benefit.

Inspection of Table 25 shows a directional benefit for belumosudil vs. BAT at week 24 in all organs affected except the liver; with responses for belumosudil, ranging from 6.7% (liver) to 56.3% (GI). However individual organ results should be interpreted with caution particularly in patients with liver involvement due to limited sample size and the exclusion criteria in ROCKstar of patients with liver transaminase (aspartate aminotransferase [AST] or alanine transaminase [ALT])>3 times the upper limit of normal for any reason.

Table 25. Naïve side by side comparison of response at week 24 from REACH-3 BAT and the pooled ROCKstar and KD025-208 (≥2 lines of prior therapy subgroup) (2022 data cut)

Organ	Pooled and combined ROCKstar and KD025-208 (n=176)		REACH-3 BAT (n=164)	
	Baseline involvement, n	Responders at week 24, n (%)	Baseline involvement, n	Responders at week 24, n (%)
Skin	144	29/144 (20.1%)	110 (67.1%)	17/110 (15.5%)
Joints/fascia	133	60/133 (45.1%)	44 (26.8%)	7/44 (15.9%)
Eyes	127	26/127 (20.5%)	93 (56.7%)	10/93 (10.8%)
Mouth	108	36/108 (33.3%)	99 (60.4%)	25/99 (25.3%)
Lungs	59	4/59 (6.8%)	49 (29.9%)	3/49 (6.1%)
Oesophagus	38	12/38 (31.6%)	17 (10.4%)	5/17 (29.4%)
Upper GI	27	13/27 (48.1%)	21 (12.8%)	8/21 (38.1%)
Lower GI	16	9/16 (56.3%)	10 (6.1%)	3/10 (30.0%)
Liver	15	1/15 (6.7%)	83 (50.6%)	18/83 (21.7%)
Overall response	-	89 (50.6%)	-	42 (25.6%)

Model approach

B1. Priority question: In Appendix N, the company assessed the proportional hazards (PH) assumptions for trial arms within ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup) and for REACH-3 separately. However, given the company has assumed a naive comparison of the pooled data from ROCKstar and KD025-208 and REACH-3, the EAG considers that the PH

assumption for belumosudil and BAT could have been assessed based on their Kaplan-Meier (KM) curves directly. As such, please provide an assessment of the PH assumption for belumosudil and BAT based on a comparison of the KM curves from the pooled ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup and REACH-3.

On 26th April we received the following clarification from the EAG: *In the model, FFS and OS are modelled using a joint-fit based on the company's indirect assessment of the PH assumption holding for treatment arms in pooled ROCKstar and KD025-208 and the treatment arms in REACH-3 and then making the assumption that the PH assumption would hold between trials. However, given the company's comments in this question, has the base case change to be independent fits? If so, the company's model does not reflect this. Given the company's implementation of data from REACH-3, the EAG considers that the KM curve for BAT from REACH-3 and the pooled belumosudil KM curve can be used directly to compare if the PH assumption holds based on this naive comparison, which is already used in the model.*

Company follow up response:

We would like to clarify that the parametric curves used in the model for BAT for FFS, OS, and DOR were always fitted independently from any belumosudil data (both in the original CS and in our response to the clarification questions). When referring to “joint fits” for BAT in the CS or the model, we were referring to the fact that these curves for BAT were obtained from models jointly fitted to the two arms in REACH-3 (i.e., BAT and ruxolitinib). This was not intended to refer to any kind of joint fitting between belumosudil and BAT (which was not attempted and this is why the PH assumption was not tested between trials). We acknowledge how this may have been confusing in parts of the CS. (Note: in the case of belumosudil, “joint fits” in the CS or the model refers to curves derived from models jointly fitted to the belumosudil 200 mg QD and belumosudil 200 mg BID data [from pooled ROCKstar and KD025-208]; again, this does not involve BAT in any way.)

Parametric fitting of survival curves and testing of the PH assumption were discussed in section B3.3 (page 89) of the CS. The tests of the PH assumption discussed in this section (and documented in Appendix N) referred to two separate series of tests: 1) Testing the PH assumption between the belumosudil 200 mg QD

and belumosudil 200 mg BID arms in the pooled ROCKstar and KD025-208 trials; 2) Testing the PH assumption between the BAT and ruxolitinib arms in the REACH-3 trial.

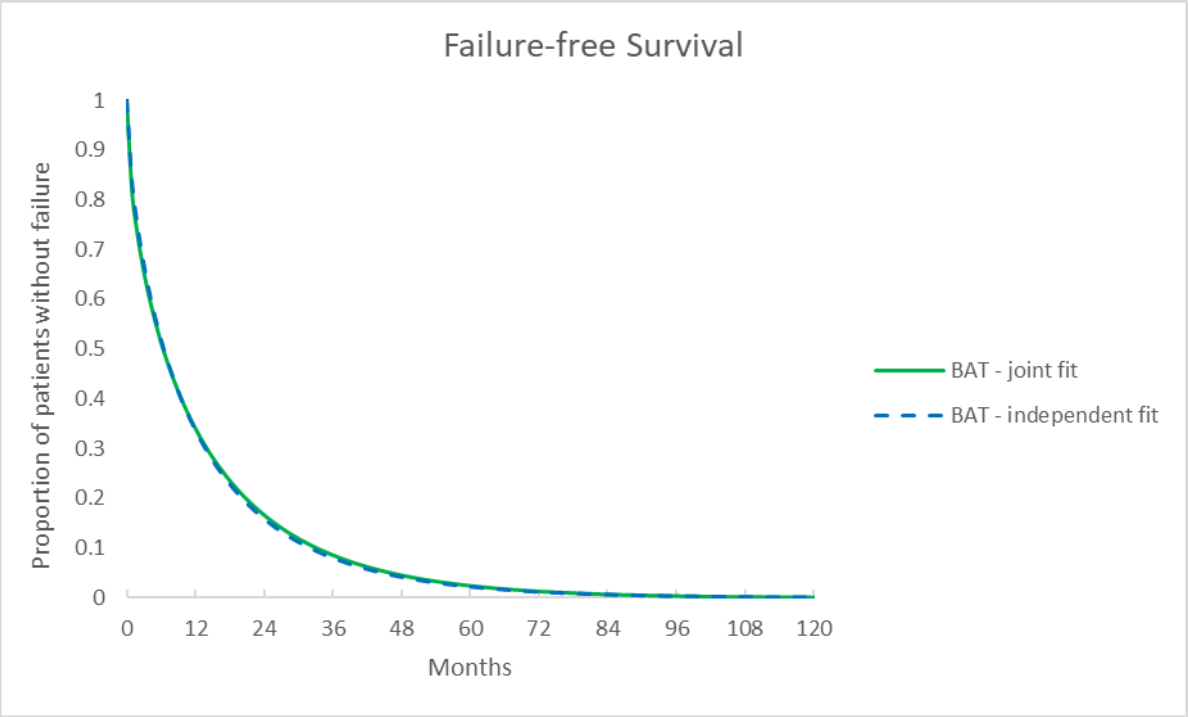
As summarised on page 89 of the CS: “For both the pooled analysis of ROCKstar and the Phase 2a study, and the REACH-3 trial, there were no concerns with respect to PH assumption and thus joint fits of the outcomes were considered. It can be noted that data for FFS and OS were immature for the treatments investigated in the ROCKstar, Phase 2a and REACH-3 trials and thus joint fits provide more reliable estimates of survival.”

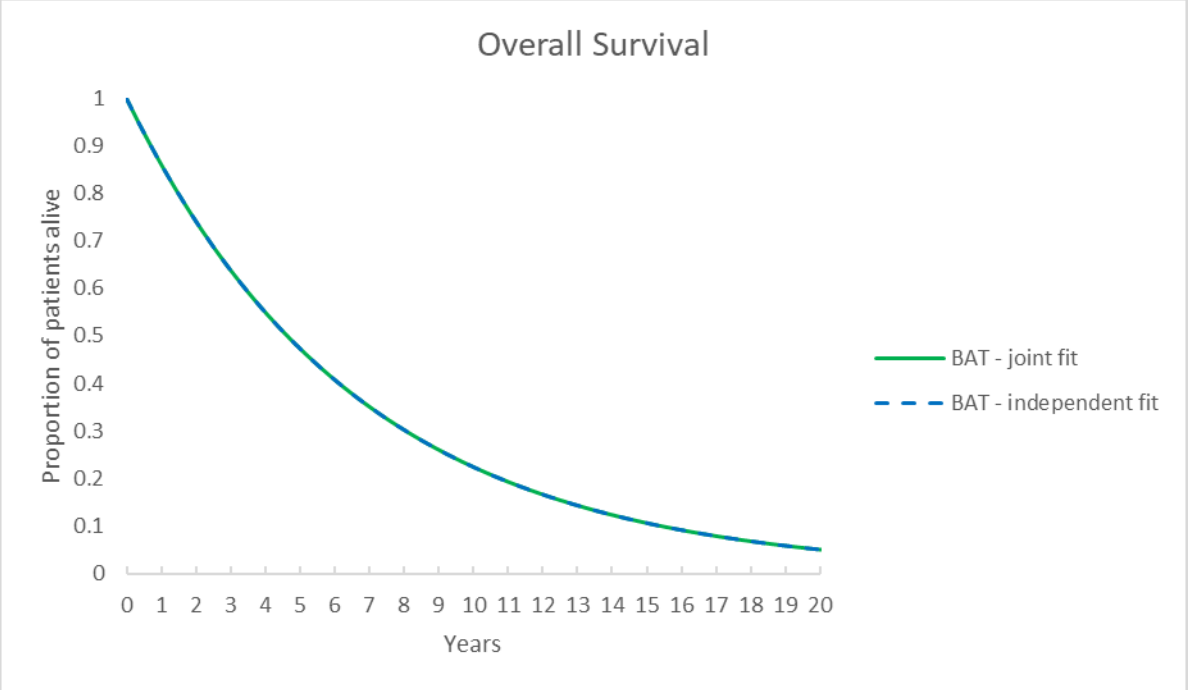
Therefore, in the base case, curves jointly fitted to belumosudil 200 mg QD and belumosudil 200 mg BID were used for these two arms in the model, and curves jointly fitted to BAT and ruxolitinib were used for the BAT arm in the model.

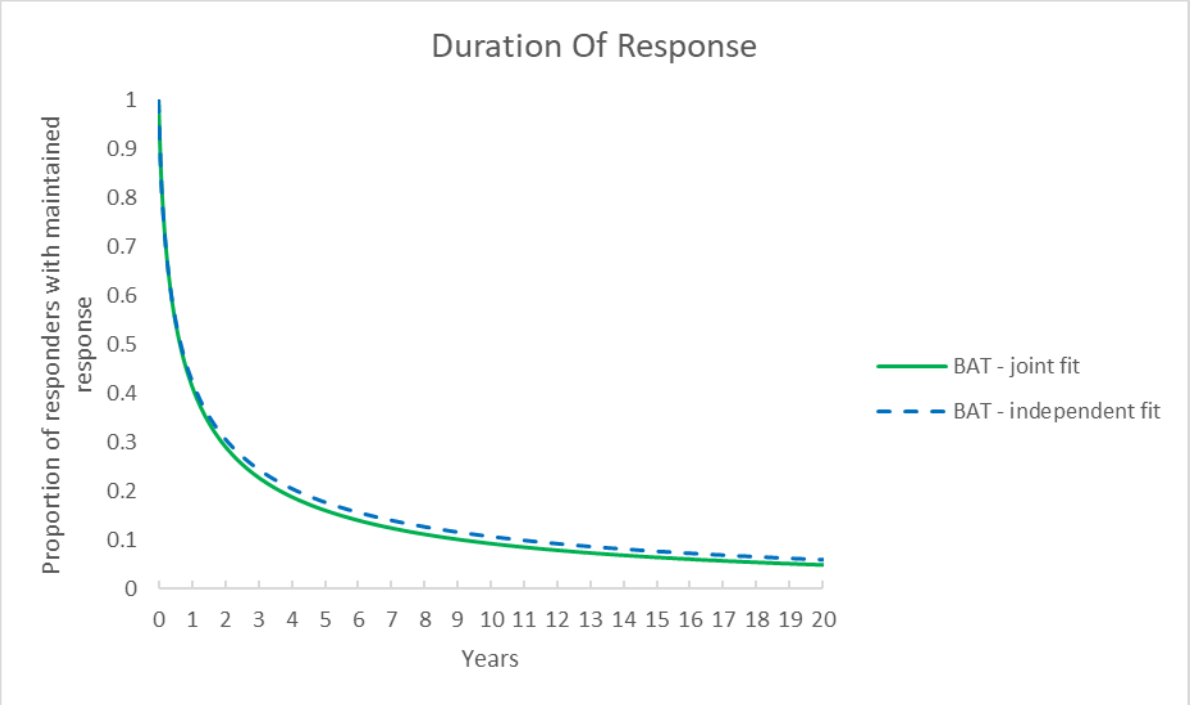
While joint fits were selected in the base case for BAT in order to leverage all available data from REACH-3, we have provided a graphical comparison below of the independently fitted curves vs. jointly fitted curves for the base-case distribution (Generalised Gamma for FFS; exponential for OS; log-normal for DOR) (Figure 1). This shows that the jointly and independently fitted curves are very similar in all cases (or even identical in the case of OS, since exponential was used in the base case).

Note that using independently fitted curves instead of jointly fitted curves for BAT for FFS and DOR would result in a lower ICER than in our base case presented in response to the clarification questions: deterministic ICERs of £3,077 without severity modifier and £2,565 with the modifier (compared to £3,571 and £2,976, respectively, when using joint fits).

Figure 1. Comparison of the independently fitted survival curves vs. jointly fitted curves for the base-case distribution







B17. Priority question: Please estimate the utility value for a partial responder by excluding data for complete responders from utility regression analysis and compare this to the utility values used in the base case for CR and PR. If the estimated utility value is lower than the base case value, please provide a scenario where the utility value for CR is maintained as per the base case ([redacted]) and the lower utility value is used for PR.

On 26th April we received the following clarification from the EAG: *The suggestion here was to drop the records completely (as stated in the question, exclude data for complete responders from the analysis).*

Company follow up response:

Regression coefficients estimates of the mixed model that excluded patients who achieved CR while using response and failure as covariates are presented in Table 26. Table 27 presents utilities in the associated health state. The model was estimated using September 2022 data cut of the ROCKstar study. Please note that in these analyses, all records of patients that achieved a CR and had utility recorded in CR were dropped, not just the records of utility while in CR and thus the number of patients included in these analyses differs from other analyses.

Table 26. Regression coefficient estimates from the mixed models including depth of response and treatment-failure as covariates.

Factor	Nr of pat.	Nr of obs.	Coef	Std Err	Low 95%CI	High 95%CI	P-value
Intercept	134	1,082	0.7215	0.0073	0.7072	0.7358	<0.001
Centered baseline utility score	134	1,082	0.7775	0.0432	0.6926	0.8633	<0.001
Response	103	670	0.0337	0.0063	0.0213	0.0462	<0.001
Treatment failure	25	74	0.0218	0.0128	-0.0033	0.0470	0.0881

Table 27. Predicted health state utilities from the mixed models including depth of response and treatment-failure as covariates.

Health states	Mean EQ-5D	Low 95%CI	High 95%CI
Failure-free, lack of response	[redacted]	[redacted]	[redacted]
Failure-free, partial response	[redacted]	[redacted]	[redacted]
Treatment failure	[redacted]	[redacted]	[redacted]

It is worth noting that the estimated utility value for the response health state ([redacted]) from this mixed model is slightly higher than the utility values used in the base case for CR and PR ([redacted]). The value of [redacted] was obtained by pooling CR and PR responses altogether and was derived from a model that included more patients. Therefore, this difference can be explained by the dropping of the records of patients who achieved a CR, who may have achieved a PR before achieving their CR.

As the estimated utility value for the response health state ([redacted]) from this mixed model is not lower than the utility value used in the base case for CR and PR ([redacted]), the deterministic ICER was calculated for a scenario using the value suggested in point B17 a) in the original set of clarification questions (i.e., utility value for CR is 10% higher than the value for PR, hence utility in CR was applied as 1.1*[redacted] = [redacted]) in the model. See Table 28 below.

Table 28. Deterministic results with updated utility values (PAS, WITHOUT and WITH severity modifier)

	ICER (£/QALY)		Change from Base case*
	WITHOUT severity modifier	WITH severity modifier	
Base case	£3,571	£2,976	N/A
Scenario	£3,564	£2,970	-0.18%

*Change from base case is the same in the analyses WITH and WITHOUT the severity modifier applied

User instructions: To run this analysis: Go to **Utilities** sheet > Change cell **G10** from [redacted] to [redacted].

References

1. Zeiser R, Polverelli N, Ram R, et al. Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. *The New England journal of medicine* 2021;385(3):228-38. doi: 10.1056/NEJMoa2033122 [published Online First: 2021/07/15]

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable.
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

1. Your name	[redacted]
2. Name of organisation	Anthony Nolan
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	**Anthony Nolan** saves the lives of people with blood cancer and other blood disorders. Founded in 1974 as the world's first stem cell register, we're motivated by a mother's determination to save her son, Anthony. Now saving three lives every day, our charity is a lifesaving legacy. By growing our register of potential stem cell donors, conducting ground-breaking research into improving transplant outcomes, and providing outstanding support and clinical care for patients and their families, Anthony Nolan cures people's blood cancer and blood disorders. In this submission, we are representing the views and experiences of stem cell transplant recipients, who have each experienced chronic Graft vs Host Disease (GvHD).
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.] If so, please state the name of the company, amount, and purpose of funding.	Pfizer paid a patient connected to Anthony Nolan, a member of our patient-led Policy Insights Panel, and an Anthony Nolan employee honorarium for attending a 3-hour workshop on developing a 'Blood Cancer Patient Charter' in November 2022: • Anthony Nolan employee - £300 • Patients' payments - £390 Anthony Nolan submitted a scoping document response on this appraisal but has not published any publicly available comment on Belumosudil or associated comparator technologies. Anthony Nolan's Patient Services team is having separate and firewalled discussions with Sanofi concerning consulting on a GvHD patient market research study. • Service contracts are TBC, with an estimated contract value of ~£4,200. • Neither Greg Judge, Niamh Buckingham nor Hugh Allen who oversee Anthony Nolan's appraisal submissions has direct contact with this study.

4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	None
5. How did you gather information about the experiences of patients and carers to include in your submission?	Information for this appraisal was gathered from a range of sources, including: • In-depth telephone and online video interviews with stem cell transplant recipients who have experienced chronic GVHD. • Insight from the Anthony Nolan Patient Services team. • Clinical nurse specialists were also consulted to build our understanding of the experiences of patients and carers.

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?	**Living with chronic GvHD**
	
	The onset of chronic GvHD is quantified as persistent inflammatory symptoms starting more than 100 days after a stem cell transplant, and these symptoms can continue for the remainder of a patient's life with a significant negative impact on their quality of life and lifestyle opportunities.
	
	Of the 6 patients who participated in our in-depth interviews, all reported living with chronic GvHD as difficult with a high level of uncertainty of continuing risks in progression. Key symptomatic factors and their common themes included:
	
	• **Eye inflammation** – donor's cells can attack the eye conjunctiva and glands, causing severe dryness and sensitivity to light:
	  ○ Dryness and sore eyes which one patient described as being "constant irritation, particularly in the early mornings when the light changes".
	  ○ Using tear substitutes was found to be problematic, from having tear ducts plugged to collect natural tears, steroid eye drops, to using plasma-derived eye drops and artificial tear products which can damage eyesight with prolonged use – all patients had difficulty in managing this.
	  ○ Limiting screen time was essential for two patients who work, splitting their part-time hours over four afternoons when their eyesight is at its best, but this requires supportive employers.
	  ○ Outside activities were severely limited for three patients, one patient no longer goes to theme parks because they can't cope with the transition of constantly going inside buildings and then outside. Another can no longer drive at night or in dusk light conditions, making them more reliant on family and friends for transport.
	  ○ Inside activities such as reading books were also reported as being curtailed, with patients reading fewer books or relying on more expensive audiobooks.
	
	• **Skin inflammation** – donor's cells can attack a patient's skin causing colour changes, thinning or thickening, hardening and rashes etc.:
	  ○ Several patients noticed changes in their skin within the first-year post-transplant, and not instantly after Day 0. Reports of skin thickening and mottling on the torso and arms were one patient's experience – this quickly led to difficulty in joint mobility as skin became tighter in these areas.
	  ○ Redness of the skin spreading across the body was a familiar symptom for patients. One used Dermovate steroid cream with limited relief in itchiness and swelling.
	  ○ An increase in sensitivity to heat and light was significant, with one patient saying, "I can't go out without covering up, it feels like I'm in an oven, in the sunlight on cold days has an effect on my skin".

	• **Mouth and dental impact** - donor's cells attack the saliva glands and soft tissues inside the mouth:
	  ○ Two patients reported symptoms in the mouth with one person particularly affected by a very dry and sore mouth with no saliva being naturally produced.
	  ○ They said, “it makes it a challenge to exercise, and means I need to keep regularly drinking water and then stopping for toilet breaks”.
	  ○ Eating can become a chore and generally less enjoyable with the need for hydration always in the back of your mind.
	  ○ Flare-ups can occur with painful spots appearing on the tongue and cheeks, relieved slightly with teething gels.
	  ○ One treatment has been beta-methadone mouthwash, but this can impact dental care – one patient has already needed 15 fillings in just two years, with good dental health before his transplantation.
	• **Lung inflammation** - donor's cells attack the lungs and cause shortness of breath, wheezing and a persistent cough:
	  ○ Patients reported that changes in their lung function seemed to “creep upon me” as it wasn’t initially noticeable without direct screening.
	  ○ Two patients now have permanent lung damage which cannot be improved, despite the GvHD now being resolved. If they catch a cold or respiratory infection, it can take much longer to begin the recovery process.
	  ○ One patient said it now means “I can’t do any real exercise, no cardio but I can walk normally for short distances” which they said has contributed to weight gain.
	  ○ Chronic bronchitis is an ongoing consequence for many with steroid inhalers being used – one patient said they had been told that the condition could be life-limiting but is currently under control.
	• **Other symptoms:**
	  ○ Stomach and colon – patients can experience severe inflammation with constipation and diarrhoea being symptoms to manage.
	  ○ Sinuses can be inflamed and/or acquire regular infections with some patients having to be admitted to A&E and screened for more serious infections.
	Mental health and wellbeing impact
	The onset of chronic GvHD can firstly be very demoralising for patients who have been showing positive signs of post-transplant recovery. Worried that this may in fact be a relapse of their disease, these new symptoms can develop a range of strong emotions. Patients expressed feelings of their ‘lives being limited’ by the GvHD and having to adapt to what they were still able to do.

	Exploring their treatment histories for GvHD, many cited the uncertainty of needing to have different therapies, only for their effectiveness to be limited. Couple this with new inflammation occurring elsewhere in their bodies and it feels like a “never-ending rollercoaster”. An issue which was raised by two women was the impact on intimacy with their partners and their sense of sexuality. GvHD can occur within the vagina, as well as areas of skin across the lower body. For some, this can make sexual intercourse painful and create a physical and emotional barrier between them and their loved ones. A common theme was that many patients described taking a “massive step backwards” in their recoveries. They saw the stem cell transplant as a potentially lifesaving treatment, which was hindered by the chronic GvHD. Patients talked of the symptoms limiting “how much I can work, exercise and socialise” with this creating a feeling of loneliness and exclusion from the activities they were used to participating in. **Effect on daily life** Patients told us that living with chronic GvHD had a significant effect on their day-to-day life, including their ability to work, have a social life, travel, and live life with spontaneity. • Managing their eyesight and transitioning between dark and light places made visiting some locations difficult, compounded for some by skin sensitivity. • Changes in lung function makes exercising difficult, with weight gain creating its own negative impacts on people’s lives. • The inability to drive in all conditions limits people’s access to both work and socialising, making them reliant on others for transport or to take turns driving. • Many of these GvHD symptoms require medication with one patient taking ‘nearly 30 pills a day’, creating its burden of organising doses, arranging repeat prescriptions and the general admin of timing medication around meals. **Carers** Patients did speak of the pressure that chronic GvHD has placed on their families. The uncertainty of changing and progressive symptoms was very apparent, along with a realisation that whilst many of the symptoms can and do subside, their effects are often life-long and have consequences for what people can do together.

Current treatment of the condition in the NHS

7. What do patients or carers think of current treatments and care available on the NHS?	None of the patients interviewed had experience with Belumosudil, and only two had direct experience with third-line comparators such as Methotrexate and Extracorporeal Photopheresis (ECP). It is hoped that Belumosudil will be able to provide an alternative treatment to persistent chronic GvHD symptoms and help control their side effects before any permanent damage to patients' eyesight or internal organs is sustained. All the patients we interviewed remarked on the side effects of high-dose steroids, whether that was indigestion, dizziness and changes in blood pressure, swings in body temperature and the long tapering off from the high doses. They initially made them feel a lot worse and none said they would favour having to go through the same experience. Patients also reflected on the need to take multiple drugs for managing their chronic GvHD over a prolonged period as well as the need to go into and remain in hospital for emergencies, usually within the first two years post-transplant. **Quality of life** Patients did note that treatments such as ECP require intravenous access, an issue for one patient for suffered from trypanophobia - an extreme fear of needles. This had a significant effect on patients' ability to have a normal life, including working and having a social life. Patients expressed a preference for oral 'at home' treatments that allowed them to leave the hospital. Belumosudil is administered orally, and this could be of benefit to patients' well-being.
8. Is there an unmet need for patients with this condition?	All patients reported the need to take control of their chronic GvHD sooner, being bounced from one treatment to the next was described as "exhausting" by one patient. More treatments are required that can be administered at home and require fewer pills overall. Patients also highlighted the extremely unpleasant side effects of many of the currently available medications and noted the importance of any new treatment that is better tolerated or has fewer serious and unpleasant side effects.

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	Belumosudil, as an oral therapy, is likely to improve patients' experience of treatment and quality of life, due to its convenience and the option to possibly take it at home. Given the challenges with comparator treatments, patients favour another option for the treatment of chronic GvHD. Another treatment option is a particularly acute need for those who may have tried all the comparator treatments already or may be unwilling to try them based on their previous experiences with existing treatments. Continued use of steroids cannot be tolerated and ECP is not always deemed suitable for patients. Other immunosuppressant drugs also carry their own risks for a patient's overall health.

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	Belumosudil is only recommended for those over 12, meaning some will not be able to benefit from the availability of this treatment. The ROCKstar study stated that "Adverse events (AEs) were consistent with those expected in patients with cGVHD receiving corticosteroids and other immunosuppressants. Sixteen subjects (12%) discontinued belumosudil because of possible drug-related AEs."[1] Patients would hope that new treatments carried fewer AEs than their comparators and were proven to have greater tolerance levels than the standard of care. Of course, overall efficacy in managing chronic GvHD remains a leading priority. *1 - Cutler C, Lee SJ, Arai S, Rotta M, Zoghi B, Lazaryan A, Ramakrishnan A, DeFilipp Z, Salhotra A, Chai-Ho W, Mehta R, Wang T, Arora M, Pusic I, Saad A, Shah NN, Abhyankar S, Bachier C, Galvin J, Im A, Langston A, Liesveld J, Juckett M, Logan A, Schachter L, Alavi A, Howard D, Waksal HW, Ryan J, Eiznhamer D, Aggarwal SK, Ieyoub J, Schueller O, Green L, Yang Z, Krenz H, Jagasia M, Blazar BR, Pavletic S. Belumosudil for chronic graft-versus-host disease after 2 or more prior lines of therapy: the ROCKstar Study. Blood. 2021 Dec 2;138(22):2278-2289. doi: 10.1182/blood.2021012021. Erratum in: Blood. 2022 Mar 17;139(11):1772. PMID: 34265047; PMCID: PMC8641099.*

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients who are eligible for a third-line therapy are first and foremost to benefit from newly available treatments at this level. However, it has been demonstrated that the treatment pathways for chronic GvHD can range between international standards, that are used within NHS commissioning policy versus the pathway used for NIH clinical trials. Belumosudil could be given as a therapy beyond third-line, and in time could be a candidate for severe cases that require a second-line therapy. The treatment pathway has not reached complete consensus in the UK; further mapping and harmonisation are required.

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	We have not identified any equality issues.

Other issues

13. Are there any other issues that you would like the committee to consider?	Patients have commented on lung damage creeping up on them and finding it to be irreversible. Similarly, resolving eye management therapies has left some patients with a worsening prognosis. It is clear that improved screening of potential chronic GvHD symptoms is needed, and this must be monitored during the introduction of new treatments in this area.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission.	• Chronic GvHD onset varies significantly from one patient to another, both in terms of timing and severity, with multiple treatments having to be administered. • The impact of quality of life can be significant, affecting people's eyesight, lung capacity, dietary needs, personal relationships, and capacity to work and have a social life. • Managing the inflammatory symptoms can take months or several years, with long-term side effects potentially leading to life-long disabilities. • It is not uncommon for some patients to be referred for 4th, 5th, or 6th line therapies; finding an effective 3rd line therapy would be beneficial to the patient and cost-effective in the long term. • Patients favour a treatment that can be administered orally; there is the potential for both quality of life and cost-saving benefits of Belumosudil over other treatments.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - **YES** or ~~NO~~

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

NHS organisation submission (ICBs and NHS England)

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 10 pages.

About you

1. Your name	[redacted]
2. Name of organisation	NHS England, Specialised Commissioning

3. Job title or position	[redacted]
4. Are you (please select Yes or No):	Commissioning services for an ICB or NHS England in general? Yes Commissioning services for an ICB or NHS England for the condition for which NICE is considering this technology? Yes Responsible for quality of service delivery in an ICB (for example, medical director, public health director, director of nursing)? No An expert in treating the condition for which NICE is considering this technology? No An expert in the clinical evidence base supporting the technology (for example, an investigator in clinical trials for the technology)? No Other (please specify):
5a. Brief description of the organisation (including who funds it).	NHS England is a DHSC arm's length body. Specialised Commissioning
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No

Current treatment of the condition in the NHS

6. Are any clinical guidelines used in the treatment of the condition, and if so, which?	NHS England Specialised Commissioning have a published policy for the management of GvHD: https://www.england.nhs.uk/wp-content/uploads/2017/03/gvhd-heamatopoietic-stem-cell.pdf
	
	For chronic GvHD (cGvHD), the policy states that:
	- Corticosteroids are first line treatment, with initial starting dose 1mg/kg
	- Where patients are at risk of developing adverse effects or becoming steroid dependent, calcineurin inhibitors can be used
	- For steroid-refractory cGvHD (failed response), sirolimus is indicated, with second line options as below:
	  ○ Pentostatin
	  ○ Skin, liver, pulmonary and oral: Extracorporeal photophoresis
	  ○ Refractory cutaneous or musculoskeletal cGvHD: rituximab
	  ○ Refractory pulmonary or sclerodermatous cGvHD: Imatinib
	- Third line treatments are: mycophenolate mofetil, methotrexate and pulsed corticosteroids.
	
	British Committee for Standards in Haematology (BCSH) and the British Society for Bone Marrow Transplantation (BSBMT) Guidelines for the diagnosis and management of acute graft-versus-host disease https://academy.myeloma.org.uk/library/guideline/guidelines-for-the-diagnosis-and-management-of-acute-graft-versus-host-disease/

7. 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.)	The pathway of care is well defined, however, although there are identified treatments, there is an unmet need for individuals who don't respond to first and second line treatments. NHS England is the commissioner responsible for 30 days before transplant, until 100 days post transplant.

8. What impact would the technology have on the current pathway of care?	This technology would have a significant impact on the current pathway of care. Approximately 30-40% of patients who have undergone an allo-HSCT will develop chronic GvHD, of which 5-6% will require second or subsequent lines of therapy. Belumosudil may be used in place of extracorporeal photophoresis (ECP) thereby reducing the need for patients to attend hospital for two days every two weeks, for a period of between 6 – 18 months.

The use of the technology

9. To what extent and in which population(s) is the technology being used in your local health economy?	This technology is currently not being used. There are currently no compassionate use schemes or trials available.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	This technology will be used in the same way as current care in the NHS.
10a. How does healthcare resource use differ between the technology and current care?	It is anticipated that no additional healthcare resource will be required with the implementation of this treatment. Resource may in fact be reduced because it may lead to fewer supportive care admissions for patients with chronic GvHD that are treated successfully. It may also reduce the requirement for ECP, which would in turn have an impact on the current provision in terms of
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	This technology will be used in specialist clinics in secondary care.

10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Belumosudil is an oral formulation (tablet) which is administered once daily. There is no anticipated investment required to introduce the technology. No additional clinic visits are expected for patients who take this treatment.
10d. If there are any rules (informal or formal) for starting and stopping treatment with the technology, does this include any additional testing?	No additional testing would be required for this intervention as based on clinical response parameters reported in trials.
11. What is the outcome of any evaluations or audits of the use of the technology?	This technology has not been used in the UK and we are not aware of any no real world data.

Equality

12a. Are there any potential equality issues that should be taken into account when considering this treatment?	There are no anticipated potential equality issues in access to belumosudil for chronic GvHD.
12b. Consider whether these issues are different from issues with current care and why.	There are currently reported issues in accessing ECP, which is not available at all transplant centres, and patients may be required to travel to another site to receive ECP. Should belumosudil become available, this would potentially improve equitable access to chronic GvHD across the country.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy [ID4021]

STA Report

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135825.

Title:	Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy [ID4021]
Produced by:	BMJ Technology Assessment Group (BMJ TAG)
Authors:	Steve Edwards, Director of Health Technology Assessment, BMJ-TAG, London Alexander Allen, Senior Clinical Evidence Analyst, BMJ-TAG, London Tracey Jhita, Health Economist Manager, BMJ-TAG, London Victoria Wakefield, Principal Clinical Evidence Analyst.
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	26/05/2023
Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135825.
Declared competing interests of the authors	No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The EAG would like to thank Professor Andrew R Gennery, Sir James Spence Professor of Child Health at Great North Children's Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust, and Dr Varun Mehra, Consultant Haematologist at King's College Hospital NHS Foundation Trust, for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report. The EAG would also like to thank Dr Jeff Davies, Reader in Haemato-oncology and Consultant Haemato-oncologist at Barts Cancer Institute and Barts Health NHS Trust for providing feedback on the clinical sections of 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.
Report reference:	Edwards SJ, Allen A, Jhita T and Wakefield V. Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy: A Single Technology Appraisal. BMJ Technology Assessment Group, 2023.

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; validated the statistical analyses; provided feedback on all versions of the report. Guarantor of the report
Alexander Allen	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, background and clinical results sections
Victoria Wakefield	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the clinical results sections
Tracey Jhita	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections

All authors read and commented on draft versions of the EAG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Adverse event
AIC	Akaike Information criterion
ALL	Acute lymphocytic leukaemia
alloHSCT	Allogenic haematopoietic stem cell transplant
AML	Acute myeloid leukaemia
BAT	Best available therapy
BIC	Bayesian information criterion
BID	Twice daily
BNF	British National Formulary
BSA	Body Surface Area
CEAC	Cost-effectiveness acceptability curve
cGvHD	Chronic graft-versus-host disease
CI	Confidence interval
CLL	Chronic lymphocytic leukaemia
CML	Chronic myeloid leukaemia
CMU	Commercial Medicines Unit
CNI	Calcineurin inhibitors
CR	Complete response
CS	Company submission
CSR	Clinical study report
CYP3A	Human cytochrome P450 3A
DOR	Duration of response
DSA	Deterministic sensitivity analysis
DSP	Disease specific programme
DSU	Decision support unit
EAG	External assessment group
ECA	External control arm
ECP	Extracorporeal photopheresis
eMIT	Drugs and pharmaceutical electronic market information tool
EPAR	European public assessment report
FFS	Failure-free survival
GI	Gastrointestinal
HES	Hospital episode statistics
HR	Hazard ratio
HRQoL	Health-related quality of life
HSCT	Haematopoietic stem cell transplant
ICER	Incremental cost-effectiveness ratio
INHB	Incremental net health benefit

INMB	Incremental net monetary benefit
IPD	Individual patient-level data
IV	Intravenous
K-M	Kaplan-Meier
KOL	Key opinion leader
LOT	Line of therapy
LR	Lack of response
LSS	Lee Symptom Scale
LY	Life year
MAIC	Matching-adjusted indirect comparison
mITT	Modified intent-to-treat population
MMF	Mycophenolate mofetil
MS	Multiple sclerosis
NHB	Net health benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMB	Net monetary benefit
ONS	Office of National Statistics
ORR	Objective response rate
OS	Overall survival
PAIC	population adjusted indirect comparison
PAS	Patient Access Scheme
PDDS	Patient determined disease steps
PH	Proportional hazards
PPI	Proton pump inhibitor
PR	Partial response
PROMIS	Patient-Reported Outcomes Measurement Information System
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
RCT	Randomised controlled trial
RDI	Relative dose intensity
QALY	Quality-adjusted life year
QD	Once per day
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
TEAE	Treatment emergent adverse event

TSD	Technical support document
TTD	Time to treatment discontinuation
TTR	Time to response
UK	United Kingdom
WTP	Willingness to pay

1 Executive summary

This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) as being potentially important for decision making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.5 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of NICE.

1.1 Overview of the EAG's key issues

Table 1 presents a summary of the EAG's key issues on the evidence submitted on the clinical and cost effectiveness belumosudil for treating chronic graft versus host disease (cGvHD) after two or more lines of systemic therapy.

Table 1. Summary of key issues

ID	Summary of issue	Report sections
1	Evidence for adolescents not available from ROCKstar and KD025-208	2.3.1, 3.2.4
2	Inclusion of concomitant medication costs for belumosudil, such that the intervention for the cost-effectiveness analysis is belumosudil in addition to BAT (belumosudil+BAT)	2.3.2, 4.2.7
3	Naïve comparison of belumosudil versus BAT	3.4.3
4	Removal of response outcomes from the economic model	4.2.4
5	Removal of OS benefit for belumosudil+BAT	4.2.4

Abbreviations: BAT, best available therapy; OS, overall survival

The key differences between the company's preferred assumptions and the EAG's preferred assumptions include:

- Inclusion of concomitant medications for belumosudil, resulting in the intervention being belumosudil in addition to best available therapy (BAT), hereafter referred to as belumosudil+BAT;
- Removal of response outcomes and overall survival (OS) benefit for belumosudil+BAT; and

Secondary differences between the company and EAG preferred assumptions include:

- Treatment discontinuation (TTD) approach for treatment arms;

- Removal of accommodation costs for patients on extracorporeal photopheresis (ECP);
- Maximum subsequent treatment duration of five years (except for rituximab);
- Alternative utility value for patients whose failure event was initiation of a new cGvHD systemic therapy;
- Caregiver disutility for patients whose failure event was initiation of a new cGvHD systemic therapy equal to caregiver disutility for patients who are failure-free and have a partial response (PR) or lack of response (LR);
- Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689; and
- Removal of intravenous (IV) disutility for BAT.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increasing failure-free survival (FFS) and OS;
- Increasing the time patients spend in response to treatment;
- Reducing the impact of adverse event (AEs), including impact of IV infusions, on patients;
- Reducing impact on caregiver health-related quality of life (HRQoL).

Overall, the technology is modelled to affect costs by:

- Its higher unit price than current treatments;
- Being given as a tablet, rather than intravenously at hospital as with ECP and rituximab;
- Increasing FFS, thus reducing the proportion of patients and the length of time spent occupying the failure health state;
- Reducing the impact of AEs on patients.

The modelling assumptions that have the greatest effect on the ICER are:

- Inclusion of concomitant medications for belumosudil, such that the intervention in the model is belumosudil+BAT;
- Inclusion of an OS benefit for belumosudil+BAT.

1.3 The decision problem: summary of the EAG's key issues

Table 2. Issue 1: Evidence for adolescents not available from ROCKstar and KD025-208

Report section	2.3.1 and 3.2.4
Description of issue and why the EAG has identified it as important	The marketing authorisation for belumosudil and the population included in the NICE final scope are patients aged 12 years and over with cGvHD after 2 or more lines of systemic therapy. However, no adolescents, aged 12 to 18 years old, were recruited to the ROCKstar and KD025-208 trials at the time of the latest data cut (September 2022). The company stated that the manageable safety profile, in relation to adverse reactions of concern to adolescent patients, can reliably be expected to be the same in adolescents as in adults. This is due to the similarity of the disease pathophysiology, general response to treatment, pharmacokinetics modelling and flat exposure-safety relationship. The company asserts that the efficacy of belumosudil is expected to be similar in adolescents as in adults. The EAG's clinical experts agreed, that from a biological perspective, there is no reason why belumosudil would not work as effectively as in adults. Many of the drugs used for cGvHD in adolescents do not have marketing authorisation due to a lack of research.
What alternative approach has the EAG suggested?	Given there are no efficacy and safety data for belumosudil in adolescents, the EAG cannot confirm if the clinical outcomes for adults would be seen in adolescents.
What is the expected effect on the cost-effectiveness estimates?	Without any evidence on the direction of the treatment effect for adolescents, the EAG is unable to comment on the expected impact on the ICER.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers this to be an unresolvable issue. However, the EAG recommends the committee obtain advice from its clinical experts on the generalisability of the evidence in adults to adolescents.

Abbreviations: BAT, best available therapy; CS, company submission; EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; N/A, not applicable; NICE, National Institute of Health and Care Excellence

1.4 The clinical and cost-effectiveness evidence: summary of the EAG's key issues

Table 3. Issue 2: Concomitant medications for belumosudil only

Report section	2.3.2, 4.2.7.9
Description of issue and why the EAG has identified it as important	The NICE final scope specifies that the intervention for this appraisal is belumosudil with established clinical management (hereafter referred to as belumosudil+BAT). However, the company included belumosudil in the model as a monotherapy and costs used reflect this assumption. In ROCKstar and KD025-208, concomitant medications were permitted. Additionally, clinical outcomes in the model include the efficacy of concomitant medications (the composition of which is similar to the basket of treatments included in BAT, but the proportions of use of each treatment may be different). The EAG considers the exclusion of concomitant medications, which would be considered akin to established clinical management, a significant omission in the cost-effectiveness analysis as these costs would be incurred

	in UK clinical practice. Furthermore, the EAG considers that the proportions of use for each treatment that makes up BAT for the belumosudil arm is likely to differ to what is assumed for BAT as a comparator. Thus, the intervention in the model should reflect belumosudil+BAT, rather than belumosudil monotherapy.
What alternative approach has the EAG suggested?	The company provided the breakdown of concomitant medications for the treatment of cGvHD used for at least five trial subjects in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup based on the 2021 data cut and provided a scenario exploring concomitant medications for belumosudil but also included concomitant medications for the BAT arm of the model for "fairness". The EAG notes concomitant medications were not permitted in REACH-3. The EAG considers that inclusion of concomitant medications for the BAT arm of the model for the scenario analysis is not clinically valid. By definition, BAT is a composition of treatments that reflects established clinical management and thus concomitant medications are implicitly part of the basket. Additionally, the usage of concomitant medications for BAT is not based on evidence as it was not permitted in REACH-3. Additionally, the company's concomitant medication scenario included tacrolimus, corticosteroids, and antibiotics, which are considered background therapies and would be equally given in all treatment arms. However, as the company estimates a survival benefit with belumosudil, the total cost of background treatments will likely exceed that of BAT. As discussed in Issue 5, the EAG considers there is substantial uncertainty with the estimated survival benefit for belumosudil and to limit the decision risk, prefers to remove this assumption. As such, the EAG prefers the following assumptions to model concomitant medications: • Removal of concomitant medication costs for BAT; • Inclusion of concomitant medication costs for belumosudil, such that the intervention in the model is belumosudil+BAT; • Exclusion of concomitant tacrolimus, corticosteroid and antibiotics.
What is the expected effect on the cost-effectiveness estimates?	The company's ICER post clarification changed from £3,571 to £16,716.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence required as the scenario resolves the issue. However, the company indicated that due to a paucity of time, they were unable to obtain concomitant medication data for the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (September 2022 data cut). As such the EAG recommends that the model is updated with the pooled analysis and employs the EAG's preferred assumptions.

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; LOT, lines of therapy; OS, overall survival;

Table 4. Issue 3: Naïve comparison of belumosudil versus BAT

Report section	3.4.3
Description of issue and why the EAG has identified it as important	The key trials for belumosudil (ROCKstar and KD025-208) were Phase II trials and did not have suitable comparator arms that met the decision problem under review. The comparator in the NICE final scope is established clinical management (BAT) without belumosudil. The company investigated using a PAIC, [redacted] and using a MAIC. None of these approaches were found to be feasible to produce a robust comparison with belumosudil+BAT. Instead, the company used the BAT arm from the Phase III, REACH-3 trial of ruxolitinib vs investigator's choice (BAT) after one prior line of therapy in a naïve direct comparison with belumosudil+BAT. The EAG's clinical experts assessed the eligibility criteria and baseline characteristics of the patients in the belumosudil+BAT trials and the REACH-3 trial to assess the direction of bias of the naïve comparison. They concluded that while many factors indicate REACH-3 BAT as a more treatable arm, a key factor, specific organ involvement, indicates the REACH 3 arm is more complex to treat. Thus, the EAG and its clinical experts consider it impossible to predict the overall likely direction of any bias resulting from differences in the patients recruited to each treatment arm.
What alternative approach has the EAG suggested?	At the clarification stage, the EAG requested the company to perform an MAIC using the IPD of the combined belumosudil+BAT arms (200 mg QD and 200 mg BID) from the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) to match to the REACH-3 BAT arm. The company stated that the composition of the organs affected within each patient in REACH-3 was not reported, and this would likely be correlated with outcomes. The EAG agrees with the company that it is not be possible to perform a robust MAIC as published IPD from REACH-3 are not available to allow matching by the composition of organs affected within each patient.
What is the expected effect on the cost-effectiveness estimates?	Given that the only option to compare clinical outcomes for belumosudil+BAT with BAT is via a naïve comparison, there is no expected impact on the cost-effectiveness results. However, the EAG emphasises the uncertainty associated with naïve comparisons of clinical outcomes from different trials.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers that with the current clinical evidence available, there is no alternative approach that can resolve the issue. Ideally, the EAG considers that the company should undertake a Phase III RCT of belumosudil+BAT versus BAT alone in the UK.

Abbreviations: BAT, best available therapy; BID, twice daily; DOR, duration of response; EAG, external assessment group; IPD, individual patient-level data; LOT, line of therapy; MAIC, matching-adjusted indirect comparison; N/A, not applicable; c, population adjusted indirect comparison; QD, once daily; RCT, randomised controlled trial; SLR, systematic literature review.

Table 5. Issue 4: Removal of response outcomes from the economic model

Report section	4.2.4
Description of issue and why the EAG has identified it as important	The EAG considers that inclusion of response in the model to add granularity to QALYs and costs in the model is potentially adding unnecessary complexity to the analysis. Changes to the 'in response' curves and response data have limited impact on the ICER and as such is not a primary driver of cost-effectiveness. Additionally, given the naïve comparison of belumosudil and BAT, inclusion of response in the model is another source of unresolvable uncertainty in the model. The EAG's clinical experts advised that in clinical practice, while response to treatment is monitored as it affects how treatments will be delivered or adjusted (thus affecting costs), failure-free survival is a more clinically relevant outcome for patients.
What alternative approach has the EAG suggested?	The EAG considers that the company's scenario where response is excluded from the model is a more appropriate approach to the cost-effectiveness analysis and removes a source of unresolvable uncertainty in the analysis, thus limiting decision risk.
What is the expected effect on the cost-effectiveness estimates?	The company's scenario removing response has limited impact on the ICER, reducing it from £3,571 to £3,434.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence required as the scenario resolves the issue.

Abbreviations: BAT, best available therapy; EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

Table 6. Issue 5: Removal of OS benefit for belumosudil+BAT.

Report section	4.2.4
Description of issue and why the EAG has identified it as important	Observed OS for both belumosudil from the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and BAT from REACH-3 is immature, with neither dataset reaching median. As such, combined with the issue of a naïve treatment comparison, the EAG considers there is substantial uncertainty in the estimated OS benefit associated with belumosudil. The uncertainty in OS due to immature data is further exacerbated as 37% of BAT patients in REACH-3 crossed over to ruxolitinib after the Week 24 assessment point. The EAG's clinical experts advised that treatments which are effective at treating cGvHD will improve survival, but that if a patient survives beyond five years, then OS between treatments are likely to be similar. Additionally, the company explained that there is no strong rationale for belumosudil to be life extending beyond the benefit of treating cGvHD and keeping patients' failure-free. In their base case, the company assumed that the risk of death for belumosudil patients would be equal to BAT after five years resulting in an estimated difference of [redacted] life years between belumosudil and BAT.
What alternative approach has the EAG suggested?	The EAG considers that removal of the OS benefit for belumosudil+BAT excludes another source of unresolvable uncertainty in the model.
What is the expected effect on the cost-effectiveness estimates?	The EAG's scenario removing the OS benefit for belumosudil+BAT changes the company's ICER post clarification from £3,571 to dominant. The change in the ICER is driven by a reduction in the time belumosudil+BAT patients spend in the failure health state ([redacted] years versus [redacted] years), which is associated with significant cost savings, while maintaining the FFS benefit.
What additional evidence or analyses might help to resolve this key issue?	Only mature OS data from both ROCKstar, KD025-208 and REACH-3 can resolve the issue. Nonetheless, for the purposes of this appraisal, the EAG's scenarios can allow committee to consider the uncertainty with assuming an OS benefit with belumosudil+BAT. However, the EAG notes that the impact of including an OS benefit in the EAG base case changes the ICER from dominant to £28,943. Therefore, the EAG recommends the committee obtain advice from its clinical experts on the clinical plausibility of an OS benefit for belumosudil+BAT.

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; FFS, failure-free survival; ICER, incremental cost-effectiveness ratio; LOT, lines of therapy; OS, overall survival; QALY, quality-adjusted life-year.

1.5 Other key issues: summary of the EAG's view

Secondary issues identified for committee consideration include the following:

- The three trials primarily included in the company's submission, ROCKstar, KD025-208 and REACH-3, have established clinical management that is reflective of care in the USA, which differs to the UK. The EAG are concerned this may limit the generalisability of the clinical evidence to cGvHD patients in the UK – Section 3.4.3.1.3.
- Approach to estimating TTD for belumosudil and BAT – Section 4.2.7.

- Company's assumption of accommodation costs for patients on ECP reimbursed by the NHS - Section 4.2.7.
- Duration of subsequent treatments (except rituximab) for patients who fail third-line treatment and initiate treatment (60%) assumed to be lifetime – Section 4.2.7.
- Utility value for patients whose failure event was initiation of a new cGvHD systemic therapy assumed to be the same as patients with a recurrent malignancy – Section 4.2.6.
- Impact on health-related quality of life (HRQoL) for caregivers of patients whose failure event was initiation of a new cGvHD systemic therapy assumed to be the same as patients with a recurrent malignancy – Section 4.2.6.
- Cost of treatment for central line-related infections was relatively high compared with other adverse events but impact on HRQoL was assumed to be zero – Section 4.2.6.
- Company included a disutility for intravenous (IV) infusions for the BAT arm of the model. However, in their scenario exploring concomitant medications for belumosudil, the company considered the utility value for the failure-free health state included the impact of IV infusions and thus removed the disutility for belumosudil only – Section 4.2.6.

The EAG notes that issues around preferred alternative assumptions around costs and utility values are important for the committee to consider, but the impact of these assumptions become secondary when the assumption of no OS benefit is employed in the model (ICER becomes dominant). However, the choice of preferred utility values informs the application of the severity modifier, discussed in Section 7. The company's base case is associated with a severity modifier of 1.2, but the EAG's preferred base case estimated a severity modifier of 1. However, the EAG notes that the both the company's and EAG's base case ICERs are below the lower bound of the cost-effectiveness range typically used by NICE, £20,000 per QALY.

1.6 Summary of EAG's preferred assumptions and resulting ICER

Table 7 presents the EAG's preferred assumptions as well as the EAG deterministic and probabilistic base case ICER. Table 8 presents deterministic scenarios around the EAG base case.

Table 7. EAG preferred assumptions and deterministic base case ICER – belumosudil+BAT versus BAT

Scenario	Incremental costs (£)	Incremental QALYs	Cumulative ICER (£/QALY)
Company base case – post clarification	[redacted]	[redacted]	3,571
Removal of response outcomes – company scenario	[redacted]	[redacted]	3,434
Removal of OS benefit	[redacted]	[redacted]	Dominant

Concomitant medication costs for belumosudil only	[redacted]	[redacted]	Dominant
Removal of cost of background therapies	[redacted]	[redacted]	Dominant
KM TTD data for belumosudil	[redacted]	[redacted]	Dominant
Exponential distribution for BAT TTD	[redacted]	[redacted]	Dominant
Removal of accommodation costs for patients on ECP	[redacted]	[redacted]	Dominant
Maximum subsequent treatment duration of five years (except for rituximab)	[redacted]	[redacted]	Dominant
Midpoint utility value of 0.608 for failure new cGvHD systemic therapy utility value	[redacted]	[redacted]	Dominant
Caregiver disutility for failure – new cGvHD systemic therapy equal to failure-free (PR/LR)	[redacted]	[redacted]	Dominant
Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689	[redacted]	[redacted]	Dominant
Removal of IV disutility for BAT	[redacted]	[redacted]	Dominant
EAG's preferred deterministic base case - combination of all scenarios	[redacted]	[redacted]	**Dominant**
EAG's preferred probabilistic base case - combination of all scenarios	[redacted]	[redacted]	**Dominant**

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; KM, Kaplan-Meier; OS, overall survival; QALY, quality adjusted life year; TTD, time to treatment discontinuation.

Table 8. Deterministic scenario analyses around the EAG base case

	Results per patient	Belumosudil+BAT	BAT	Incremental value
0	**EAG base case**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
1	**Inclusion of OS benefit**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	28,449
2	**Adelphi DSP treatment failure utility value (0.52) for the failure – new cGvHD systemic therapy**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
3	**Utility value of 0.696 from Crespo *et al* for the failure – new cGvHD systemic therapy**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant

4	**Removal of caregiver disutility**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
5	**Scenario 1 + 4**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	26,749

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; KM, Kaplan-Meier; OS, overall survival; QALY, quality adjusted life year; TTD, time to treatment discontinuation

For further details of the exploratory and sensitivity analyses done by the EAG, see Section 6.2.

2 Introduction and background

2.1 Introduction

This report contains an assessment of the company submission (CS) submitted for the Single Technology Appraisal (STA) of belumosudil (Rezurock®, Sanofi) for treating chronic graft versus host disease (cGvHD) after 2 or more lines of systemic therapy.

Belumosudil has marketing authorisation for the treatment of patients aged 12 years and older with cGvHD who have received at least two prior lines of systemic therapy.[1]

2.2 Background

GvHD is a serious complication of allogeneic haematopoietic stem cell transplant (alloHSCT). An alloHSCT is most often used to treat blood cancers, such as leukaemia and lymphoma, and certain types of blood or immune system disorders. GvHD happens when particular types of white blood cell (T cells) in the donated stem cells or bone marrow attack the hosts own healthy tissues. This is because the donated cells (the graft) see the host's body cells as foreign and attack them.

People may be diagnosed with acute or chronic GvHD. Acute GvHD (aGvHD) generally starts within 100 days of the transplant and is not an indication relevant to this decision problem. Chronic GvHD, occurs in 20% to 50% of people who undergo alloHSCT. It starts more than 100 days after the transplant and may involve a single organ or several organs: skin, joints/fascia, eyes, mouth, lungs, oesophagus, gut, liver. A patient's involved organs undergo tissue inflammation and fibrosis that often results in permanent organ dysfunction. The severity of cGvHD is graded as mild, moderate or severe, based on the number of organs involved and the severity within each organ.[2]

Section B.1 of the company submission (CS) provides an overview of cGvHD. Based on advice from the EAG's clinical experts, the CS presents an accurate overview of diagnosis and classification, clinical presentation, development, epidemiology and disease burden. The company also expand on the humanistic burden on the patients with cGvHD and the impact of the disease on caregivers and family members.

2.2.1 Positioning of belumosudil in the UK treatment pathway

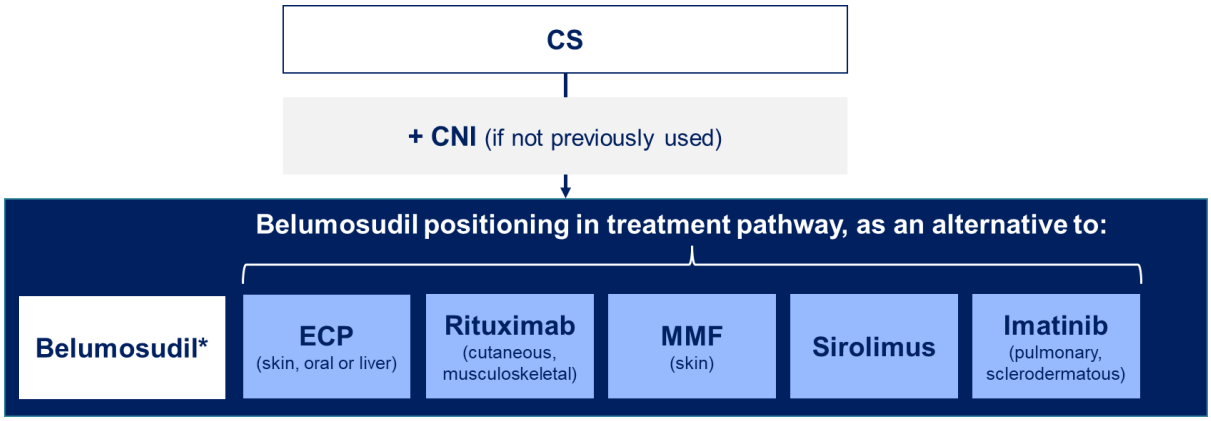
The company detail two published treatment pathways for cGvHD care in the UK. In 2012, a joint working group established by the Haemato-oncology subgroup of the British Committee for Standards in Haematology (BCSH) and the British Society for Bone Marrow Transplantation (BSBMT) made recommendations for the diagnosis and management of cGvHD (Table 5, CS).[3] This was

followed by the 2017 Clinical Commissioning Policy issued by NHS England to define a treatment pathway for the clinical management of cGvHD in England (Table 6, CS).[4]

The BCSH/BSBMT 2012 report and the NHS England 2017 Clinical Commissioning Policy present similar treatment pathways. First line (1L) treatment is corticosteroids with or without calcineurin inhibitor (CNI). Second line (2L) treatment is extracorporeal photopheresis (ECP), pentostatin, rituximab, and/or imatinib. Third line (3L) treatment is mycophenolate mofetil (MMF), methotrexate (MTX), or pulsed corticosteroids.

The company presented a treatment pathway for cGvHD they developed at an advisory board in January 2023 with clinical and health economic experts. This places corticosteroids at 1L, CNIs at 2L, and other relevant treatments at 3L (Figure 4, CS). The company positioned treatment with belumosudil at 3L in their treatment pathway "as an alternative" to ECP, rituximab, MMF, sirolimus, and imatinib (Figure 1, below). The company state that they intend for belumosudil to be used as a monotherapy, in the cGVHD treatment pathway, after oral corticosteroids (with or without the addition of CNIs) and at least one other systemic therapy, such as sirolimus or the later addition of a CNIs.

Figure 1. Treatment pathway for patients with chronic GVHD in England (reproduced from Figure 5 of the CS)



Belumosudil is positioned after oral corticosteroids (with or without the addition of CNIs) and at least one other systemic therapy, such as sirolimus or the later addition of a CNI.

Abbreviations: CS: corticosteroids; CNI: calcineurin inhibitors; ECP: extracorporeal photopheresis; MMF: mycophenolate mofetil.

The EAG's clinical experts provided the treatment pathway used in their clinical care which is different to the national guidance and the company's view. They also noted where belumosudil would be positioned, given the marketing authorisation restricts the use of belumosudil to patients who have received at least two prior systemic therapies. The EAG's clinical experts consider 1L

treatment to be corticosteroids with or without CNIs, 2L treatment is ECP in eligible patients, and 3L treatment are other therapies, including belumosudil. The EAG's clinical experts highlighted that ECP is limited to five NHS Blood and Transplant (NHSBT) Therapeutic Apheresis Units in England and this can limit the accessibility of treatment. The treatment pathway recommended by the EAG's clinical experts is noted below. This would be adjusted based on the specific organs involved and a patient's access to ECP.

First line: corticosteroids +/- CNIs;

Second line: ECP;

Third line: belumosudil, imatinib, MMF, pentostatin, pulsed corticosteroids, rituximab, sirolimus.

The EAG's clinical experts noted potential advantages to using belumosudil alongside ECP. Belumosudil is an oral medication and can be easily dispensed and may have a fast therapeutic response. Access to ECP can be slower but the experts considered the effect may be more sustained.

The EAG's clinical experts also remarked on ECP in adolescents who are 12 to 18 years old (12-18yo). ECP requires large veins or a central line is used and, due to this, adolescents (12-18yo) more frequently require a central line than adults. This can delay starting ECP for as long as a week as this requires an apheresis team and it is invasive and intensive. In these cases, patients may prefer belumosudil as it is non-invasive, oral treatment and can be started immediately.

2.2.1.1 cGvHD treatment weaning

Once a patient's organ systems are responding to treatment then they are gradually weaned off cGvHD medications to reduce the deleterious effects of treatment, such as excessive off target immunosuppression. This manifests itself as infection that can be life threatening. Therefore, patients who respond to treatment for 1-2 years would be gradually weaned off treatment over the next 1-2 years. This would occur in people who have partial response (PR) as well as those with complete response (CR). Five years after diagnosis with cGvHD, the EAG's clinical experts stated that patients would likely have either discontinued treatment or have died. However, cGvHD symptoms can worsen when treatment is stepped down and treatment can be stepped up again.

2.3 Critique of the company's definition of the decision problem

A summary of the final scope issued by NICE, together with the company's rationale for any deviation from this is provided in Table 9.[5] Key differences between the decision problem addressed in the CS and the scope are discussed in greater detail in the sections that follow below.

Table 9. Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Company's rationale if different from the scope	EAG comment
Population	People aged 12 years and over with chronic GVHD after 2 or more lines of systemic therapy.	As per final scope	N/A	The eligibility criteria for ROCKstar met the final scope issued by NICE. The eligibility criteria for the KD025-208 trial included patients at a minimum of 1 LOT and the age criterion did not include adolescents (12-18yo).Therefore, no adolescents (12-18yo) were recruited to the trial. Patients recruited to the BAT arm in the REACH-3 trial were appropriate, except they were at an earlier stage of the treatment pathway. Forty-nine percent of the patients had only received corticosteroids as prior cGvHD therapy and 42% had only received corticosteroids with CNI as prior cGvHD therapy. Patients with 3 previous lines of therapy (LOTs) were excluded. See Section 2.3.1.
Intervention	Belumosudil with established clinical management.	As per final scope	N/A	The interventions in the ROCKstar and KD025-208 trials matched the NICE final scope. Both trials included an arm where belumosudil 200 mg once daily (QD) was used and an arm where belumosudil 200 mg twice daily (BID) was used. Belumosudil 200 mg QD is the recommended dose but people on PPIs or strong CYP3A inducers require an increased dosage of 200 mg BID. However, patients in the belumosudil trials who were on PPIs did not exclusively receive belumosudil 200 mg BID. The established clinical management in both trials was appropriate for the USA, which would have been different in the UK. In UK care, a higher proportion

				would have received ECP and a lower proportion received sirolimus. See Section 2.3.2.
Comparator(s)	Established clinical management without belumosudil, including: • ECP • Imatinib • Rituximab • Sirolimus • MMF • Tacrolimus • Cyclosporine	As per final scope	N/A	The comparator arm in REACH-3 received best available therapy (BAT) chosen by the investigators from a list of 10 commonly used options. This list included: ECP, imatinib, everolimus, low-dose methotrexate, rituximab, pentostatin, sirolimus, and MMF. CNIs could be given with corticosteroids alongside BAT. The EAG's clinical experts noted that the number receiving ECP would be substantially higher in the UK and that ibrutinib, low-dose methotrexate, everolimus, and infliximab are not commonly used for this indication. Also, 61 (38%) patients received ruxolitinib after week 24 in the BAT arm and there is currently no NICE guidance on using ruxolitinib for cGvHD in the UK. However, for the economic analysis, the company adjusted the proportions of treatments received for BAT in REACH-3 to be reflective of UK clinical practice. See Section 2.3.3.
Outcomes	The outcome measures to be considered include: • Response to treatment (including complete response and overall response) • Immunosuppressant sparing • Mortality • Treatment AEs	As per final scope	N/A	The outcomes in the submission largely match the outcomes in the NICE final scope. The EAG notes that ROCKstar, KD025-208 and REACH-3 trials report on the steroid sparing within the trial. HRQoL data collected in ROCKstar using PROMIS Global Health 10 (PROMIS-GH) was mapped to EQ-5D-5L to provide utility data for the economic model. The EAG also notes that there were inconsistencies in the outcome definitions and time points between the belumosudil trials and REACH-3. See Section 2.3.4.

	• FFS • HRQoL			
Economic analysis	The reference case stipulates that: • The cost-effectiveness of treatments should be expressed in terms of incremental cost per QALY. • 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 availability of any managed access arrangement for the	As per final scope	N/A	N/A, as per final scope.

	intervention will be taken into account. • The availability and cost of biosimilar and generic products should be taken into account.			
Subgroups to be considered	If the evidence allows the following subgroups will be considered. These include: • Different organs or tissues affected by chronic GVHD • Number and type of previous treatments	As per final scope	N/A	The EAG is aware that the earlier LOT in REACH-3 does not allow for subgroup analysis linked to the number and type of previous treatments. However, data are reported for both treatment arms addressing different organs or tissues affected by cGvHD. Response by organ at week 24 is presented in table S4 in the REACH-3 trial appendix. At the clarification stage, the company provided a table of response by organ at week 24 using the naïve comparison to REACH-3 (Section 3.5.3). See Section 2.3.5 for further details on subgroups to be considered.

Abbreviations: 12-18yo, 12-18 years old; cGvHD, chronic graft versus host disease; BAT, best available therapy; CNI, calcineurin inhibitors; ECP, extracorporeal photopheresis; LOTs, lines of therapy; ORR, overall response rate; PPIs, proton pump inhibitors; PROMIS, Patient-Reported Outcomes Measurement Information System.

2.3.1 Population

The population in the NICE final scope are patients aged 12 and over with cGvHD after 2 or more lines of systemic therapy (LOTs). The populations recruited for ROCKstar, KD025-208, and REACH-3, the trials primarily used to support this submission, are reported below.

ROCKstar is an ongoing, Phase II, randomised, open-label multicentre trial comparing belumosudil 200 mg once daily (QD) and belumosudil 200 mg twice daily (BID) in alloHSCT recipients aged ≥12 years with persistent cGvHD after receiving 2 to 5 prior systemic LOTs. ROCKstar recruited 152 patients in 28 centres across the USA. The current eligibility criteria for the ROCKstar trial matches the population in the NICE final scope. However, the original age criterion were adults (≥18 years old) until a protocol amendment was made on 01 June 2020. Despite the change to the eligibility criteria, no adolescents (12-18yo) were recruited to the trial for the September 2022 data cut used in the CS. However, the trial is ongoing and in response to clarification question A2, the company stated at clarification that there were two adolescents (12-18yo) currently enrolled in ROCKstar. At the clarification stage, the company justify using the same dose and assuming the same efficacy and safety in adolescents (12-18yo) as was demonstrated in adults. See the EAG's critique in Section 3.2.4.

KD025-208 was a Phase IIa, open-label, dose-escalation, multicentre study comparing belumosudil 200 mg QD, belumosudil 200 mg BID, and belumosudil 400 mg QD in patients with cGvHD. KD025-208 completed in May 2022. KD025-208 enrolled 54 patients in seven centres across the USA. The patients recruited were alloHSCT recipients aged ≥18 years with persistent cGvHD who had received 1-3 LOTs. The eligibility criteria in the KD025-208 trial differed to the NICE final scope in two ways: adolescents (12-18yo) were not recruited and patients with a minimum of one prior LOT were recruited. The company address the LOT discrepancy by using data from the subgroup of patients who had received ≥2 prior LOTs in the economic model.

REACH-3 was a Phase III open-label, randomised, multicentre trial which evaluated the efficacy and safety of ruxolitinib at a dose of 10 mg twice daily, as compared with best available therapy (BAT). REACH-3 was conducted across 149 centres in 28 countries, including the UK. The BAT control arm from REACH-3 was used by the company to provide the data for the comparators in the CS. The patients recruited to the REACH-3 trial were alloHSCT recipients aged ≥12 years with moderate or severe glucocorticoid-refractory cGvHD. REACH-3 did include adolescents (12-18yo) with the age of

participants ranging from 12 to 76 years old. However, patients treated previously with 2 or more systemic therapies for cGvHD in addition to corticosteroids with or without CNIs were ineligible. Forty-nine percent of the patients in the control arm of the REACH-3 trial had only received corticosteroid treatment and 42% had only received corticosteroid with CNIs, at entry to the trial. Therefore, there is proportion of people in the REACH-3 control arm who had not previously received 2 or more LOTs for cGvHD and did not match the population in the NICE final scope. Lines of therapy is discussed further in Section 2.2.1.

The population considered in the economic model are patients aged 12 years and older with cGvHD who have had ≥2 prior LOTs, which is reflective of the NICE final scope.[5] Clinical data from ROCKstar and KD025-208 for belumosudil QD and BID were pooled and analysed for the subgroup of patients who have had ≥2 prior LOTs and used to inform the economic model. Baseline characteristics included in the model were obtained from the pooled ROCKstar and KD025-208 (combined dose data) for the ≥2 LOTs subgroup (presented in Table 10). Please see Section 3.2.3 for further details on baseline characteristics from the belumosudil trials.

Table 10. Modelled population baseline characteristics (taken from the company's post-clarification model)

Baseline characteristic	Value used in the economic model
Mean age (years)	53.9
Proportion males (%)	58.0
Body surface area (m^2)	1.90

2.3.2 *Intervention*

Belumosudil (Rezurock®) is a selective Rho-associated coiled-coil kinase 2 (ROCK2) inhibitor, which targets both immune response dysfunction and downregulates fibrotic processes associated with cGvHD.[1, 6] Belumosudil received marketing authorisation in Great Britain on 7 July 2022 for the treatment of patients aged 12 years and older with cGvHD who have received at least two prior LOTs.

The recommended dose, for people aged 12 years and older, is belumosudil 200 mg, administered orally, once daily.[1] The Summary of Product Characteristics (SmPC) also states that strong CYP3A4 inducers and proton pump inhibitors (PPIs) may decrease belumosudil exposure.[1] Therefore, a dosage increase to 200 mg twice daily is recommended in people who are co-administering with strong CYP3A4 inducers or PPIs.

In ROCKstar, patients were randomised to two belumosudil intervention arms; 200 mg QD and 200 mg BID, in addition to established clinical management. Randomisation was not stratified or dependent on CYP3A4 inducer or PPI usage. This is discussed in Sections 3.2.1 and 3.3.5. KD025-208 was a dose-finding trial where patients were assigned to belumosudil in addition to established clinical management; 200 mg QD, 200 mg BID, or 400 mg QD. The EAG notes that the marketing authorisation does not permit belumosudil 400 mg once daily and this dosing regimen is not considered further.

In Section 3.2.3 of the CS, the company note that due to the heterogenous nature of cGvHD and the prescription of different medicines according to manifestation and disease stage, it is appropriate to consider BAT as a 'basket of therapies'. People may receive a single therapy or combination of therapies depending on their symptoms and their response to previous LOTs. The patients recruited to ROCKstar and KD025-208 were treated with belumosudil, with what were referred to in the CS, as concomitant therapies. The concomitant medications permitted in the ROCKstar trial were corticosteroids, CNIs, sirolimus, MMF, methotrexate, rituximab and ECP (Table 9 in Section B.2.3.1, CS). The EAG understand these therapies to represent the basket of therapies (BAT) that a patient with cGvHD might receive at the third LOT. The company support this in their reply to clarification question B23, where they state that they recognise that concomitant medications are used in clinical practice and many of these treatments will be common between the belumosudil with established care, and established care without belumosudil, arms in the model. Therefore, the treatment regimens used in the ROCKstar and KD025-208 trials will henceforth be referred to as belumosudil+BAT.

[redacted]

The EAG notes, that in the ROCKstar and KD025-208 trials, [redacted]

[redacted] [7, 8]

In Table 10 (February 2020 data cut) in the CS, 99.2% in the patients in ROCKstar, and 100% of the patients in KD025-208, were using concomitant corticosteroids. Corticosteroids are linked to irritation of the stomach lining and PPIs are used to address irritation of the stomach lining. Belumosudil+BAT has been shown to have a steroid sparing effect (see Section 3.3.3) and the company assumes that a reduction in the use of corticosteroids would lead to a reduction in the use of PPIs.

In the economic model, both the belumosudil 200 mg QD and BID regimen are included. However, given the SmPC guidance for patients on CYP3A inducers and PPIs, the company assumed that only 5% of patients would be on these treatments and would require the belumosudil 200 mg BID regimen. Therefore, in the economic model, 95% of patients are assumed to receive belumosudil 200 mg QD and the remaining 5% are assumed to receive belumosudil 200 mg BID. The EAG's clinical experts considered this to be an optimistic assessment of the steroid sparing effect of belumosudil+BAT and estimated that up to 10% of patients would still be on PPIs. Pooled clinical efficacy data from ROCKstar and KD025-208 for both belumosudil+BAT regimens are provided in the CS and the economic model for the ≥2 prior LOTs subgroup and these are discussed further in Section 4.2.4.

As mentioned previously, ROCKstar and KD025-208 were USA-based studies and thus established clinical management is representative of USA care. The EAG's clinical experts advised that established clinical management in the UK is different to the USA with higher use of ECP and a lower use of sirolimus. In the combined belumosudil+BAT arm (Table 11), 27.9% of patients were concomitant with ECP at baseline and the EAG's clinical experts considered that in the UK 60–65% of patients would be expected to be on ECP. Sirolimus was used by 21.2% in the belumosudil+BAT trials, whereas this would be closer to 5% of patients in UK care. Additionally, there is currently no NICE guidance available for the use of ruxolitinib for patients with cGvHD.

Table 11. Concomitant systemic chronic GvHD therapies in ROCKstar and KD025-208[a]

Baseline characteristic	200 mg once daily n=92[a]	200 mg twice daily n=84 [a]	Combined 200 mg N=176 [a]
Systemic hormonal			

preparations	91 (98.9%)	82 (97.6%)	173 (98.3%)
ECP	22 (23.9%)	28 (33.3%)	50 (28.4%)
Tacrolimus	30 (32.6%)	26 (31.0%)	56 (31.8%)
Sirolimus	19 (20.7%)	20 (23.8%)	39 (22.2%)
MMF	11 (12.0%)	3 (3.6%)	14 (8.0%)
Imatinib	1 (1.1%)	1 (1.2%)	2 (1.1%)
Rituximab	1 (1.1%)	0	1 (0.6%)
Ruxolitinib	1 (1.1%)	0	1 (0.6%)

Abbreviations: CNI, calcineurin inhibitor; ECP, extracorporeal photopheresis; GVHD, graft-versus-host disease; MMF, mycophenolate mofetil.

[a] This uses the September 2022 data cut and includes patients in KD025-208 who had ≥2 prior lines of treatment

In the economic model, the company did not include concomitant medications for the belumosudil+BAT arm and did not provide any justification for their exclusion, which the EAG considers to be an important omission. During the clarification stage, the EAG requested the company provide an explanation for why concomitant medications were excluded from their base case (Question B23). The company explained that concomitant medications were excluded based on guidance from their clinical trials team, who considered that efficacy of a patient's existing treatment package would not "boost" the efficacy of belumosudil.

The EAG considers the company's rationale is not satisfactory as it ignores the fact that belumosudil was given in addition to BAT in ROCKstar and KD025-208 (which aligns with the NICE final scope) and that this is how it will likely be provided to patients in the NHS. Thus, the total costs for belumosudil should include the costs of concomitant medications (BAT) as these will still be incurred by the NHS, but the usage of these treatments is likely to differ to BAT provided without belumosudil. Inclusion of the cost of concomitant medication for belumosudil in the economic model is discussed further in Section 4.2.7.9.

2.3.3 Comparator

The comparator listed in the NICE final scope is established clinical management without belumosudil, including:

- ECP;
- Imatinib;
- Rituximab;
- Sirolimus;
- MMF;

- Tacrolimus;
- Cyclosporine .

The ROCKstar and KD025-208 trials did not provide suitable comparator data for established clinical management and so the company explored alternative sources of comparator data. One method used by the company was the [redacted] using [redacted] This is discussed further in Section 3.4.2, but the company concluded that the [redacted]

The company also explored using a population adjusted indirect comparison (PAIC) to enable a comparison with established clinical management without belumosudil. The company undertook a systematic literature review (SLR) and identified no studies that could be reliably compared with the belumosudil+BAT trials using a PAIC. This is further discussed in Section 3.4.1.

The EAG notes that the company used the trial-based BAT control arm from the REACH-3 study, identified in the SLR, to allow comparison of belumosudil+BAT to currently available treatments in the economic model through a naïve direct comparison.

In REACH-3, patients in the BAT arm received a therapy chosen by the investigators from a list of 10 commonly used options (Table 12). Patients in the BAT arm of REACH-3 who were on corticosteroids with or without CNIs at baseline could continue with these treatments through the trial. The trial protocol allowed patients in the BAT arm to crossover to ruxolitinib, on or after, week 24. Ruxolitinib was used by 61 (38%) of the patients in the REACH-3 BAT arm after week 24.

The EAG's clinical experts indicated that BAT received in REACH-3 generally reflected established clinical management that would be given in the USA and were appropriate given that patients were receiving 2L or 3L treatment. As mentioned previously, the EAG's clinical experts advised that established clinical management in the UK is different to the USA. Specifically, usage of ECP would be around 60–65% and ibrutinib, low-dose methotrexate, everolimus, and infliximab are not commonly used for this indication in the UK. Up until April 2022, ruxolitinib was reimbursed by NHS England through an interim rapid commissioning policy in the context of the COVID-19 pandemic. However, use in new patients is no longer permitted due to the withdrawal of the interim commissioning policy.[9]

The EAG's clinical experts noted that that all three trials, ROCKstar, KD025-208 and REACH-3, have established clinical management that is reflective of USA care. It is likely that established clinical management was consistent between the three trials despite it being different to UK care.

For the economic model, the company reweighted the proportions of each treatment included in BAT from REACH-3 to be reflective of UK clinical practice, presented in Table 12. The EAG's clinical experts considered the proportions of each treatment in BAT included in the economic model was appropriate and reflective of UK clinical practice. Please refer to Section 4.2.7 for more details on the dosing regimen and administration for each of the treatments considered as part of BAT.

Table 12. BAT treatments in REACH-3 (adapted from Table S1, Zeiser 2021)[10] and included in the economic model

Treatment	Best available therapy (n=158) – REACH-3	Adjusted proportions used for the company base case
ECP	55 (34.8)	64.6%
MMF	35 (22.2)	22.2%
Ibrutinib	27 (17.1)	-
Low-dose methotrexate	10 (6.3)	-
Imatinib	8 (5.1)	5.1%
Sirolimus	7 (4.4)	4.4%
Rituximab	6 (3.8)	3.8%
Everolimus	5 (3.2)	-
Infliximab	5 (3.2)	-
Pentostatin	0	-

Abbreviations: ECP, extracorporeal photopheresis; MMF, mycophenolate mofetil.

The EAG notes that the company did not justify the exclusion of tacrolimus and cyclosporine from the comparator arm of the model, given that these treatments were listed in the NICE final scope.[5] Additionally, in the company's NICE advisory board report, participants advised that [redacted]

CNIs can be added to the treatment regimen for cGvHD to decrease steroid dosage and duration. When consulted, the EAG's clinical experts advised that tacrolimus and cyclosporine are used in UK clinical practice but earlier in the pathway. However, the EAG's clinical experts considered once patients required tacrolimus and/or cyclosporine, they would continue with treatment regardless of what line of therapy they were on and as such could be considered as "background" therapies.

During the clarification stage, the EAG requested the company to justify the exclusion of tacrolimus and cyclosporine from the model (Question B24) and the company agreed that tacrolimus and cyclosporine are used as background therapies and that usage of these treatments is likely to be the same for both belumosudil+BAT and BAT. Additionally, the company explained that in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup only one patient received cyclosporine.

Given the EAG's clinical experts' view that tacrolimus and cyclosporine are background therapies for patients with cGvHD, the EAG agrees with the company that usage of these treatments is unlikely to differ between belumosudil+BAT and BAT patients. Thus, exclusion of tacrolimus and cyclosporine as comparators to belumosudil+BAT is not unreasonable.

2.3.4 Outcomes

Outcomes relevant to the decision problem reported in the belumosudil+BAT trials and the REACH-3 BAT arm are detailed in this section. ROCKstar, KD025-208 and REACH-3 use the 2014 National Institutes of Health (NIH) consensus criteria to define a patient's response to cGvHD treatment (Table 13).[12]

Table 13. 2014 National Institutes of Health (NIH) presented consensus criteria[12]

Complete Response (CR)	Resolution of all manifestations in each organ or site
Partial Response (PR)	Improvement in at least one organ or site without progression in any other organ or site
Lack of response (LR)	Progression: progression in at least one organ or site without a response in any other organ or site
	Mixed response: complete or partial response in at least one organ accompanied by progression in another organ
	Unchanged: outcomes that do not meet the criteria for complete response, partial response, progression or mixed response are considered unchanged

The ROCKstar (September 2022 data cut) and KD025-208 trials reported response to treatment (including complete response and overall response), mortality, treatment adverse events (AEs), and failure-free survival (FFS). The following outcomes, reported in ROCKstar and KD025-208, are relevant to the decision:

- Response to treatment:

    - Best overall response rate (ORR) at any time: the percentage of patients that had either complete response (CR) or partial response (PR), using the 2014 National Institutes of Health (NIH) Consensus Criteria as assessed by investigators;[12]
    - Best ORR by organ system: best response at any time (CR or PR) for the 9 individual organs (skin, eyes, mouth, oesophagus, upper GI, lower GI, liver, lungs, and joints and fascia);
    - Time to response (TTR) reported in ROCKstar;
    - Duration of response (DOR):
        - Primary/secondary: time of first documentation of response to deterioration from best response, initiation of new systemic therapy for chronic GVHD, or death
        - Quaternary: time of documentation of response to the time of documented lack of response, initiation of new systemic therapy for chronic GVHD, or death. Durations were summed for multiple response/lack of response episodes;
    - Time to treatment discontinuation (TTD);
- FFS: the time from the first dose of belumosudil+BAT to the earliest of:
    - new systemic chronic GVHD therapy;
    - non-relapse mortality;
    - recurrent malignancy;
- Overall survival (OS): time from date of randomisation to date of death due to any cause;
- Immunosuppressant sparing: Corticosteroid reduction and corticosteroid discontinuation was addressed (Table 29 of the CS).The proportion of patients who discontinued CNIs during ROCKstar (August 2021 data cut) is reported in Table 18 of the CS;
- Treatment AEs;
- Quality of life:
    - Lee Symptom Scale;
    - PROMIS-GH: collected in the ROCKstar trial and mapped to EQ-5D-5L.[13]

REACH-3 reported the following outcomes that are relevant to the decision problem:

- Response to treatment (including complete response and overall response):
    - Proportion of patients with CR or PR at 24 weeks, according to the NIH Consensus Criteria.[12]

    - Best ORR at any time within 24 weeks: the percentage of patients that had either CR or PR, using the 2014 National Institutes of Health (NIH) Consensus Criteria;[12]
    - Response rate by organ system (CR or PR) for the 9 individual organs (skin, eyes, mouth, oesophagus, upper GI, lower GI, liver, lungs, and joints and fascia) at 24 weeks;
    - Duration of response: time from first documented CR or PR to cGvHD progression, death, or the additional of new systemic therapies for cGvHD;
- FFS: time from date of randomisation to the earliest of:
    - Relapse or recurrence of underlying disease or death due to underlying disease;
    - Non-relapse mortality;
    - Addition or initiation of another systemic therapy for cGvHD;
- OS: time from date of randomization to date of death due to any cause;
- Immunosuppressant sparing: corticosteroid dose up to Week 24 is presented as a graph but no data on reduction in the use of CNIs;
- Treatment AEs;
- Quality of life:
    - Lee Symptom Scale.

The EAG notes a number of inconsistencies between the outcomes reported in the belumosudil+BAT trials (ROCKstar and KD025-208) and the REACH-3 trial.

The company uses best ORR at any time, from the belumosudil+BAT trials, and best ORR at 24 weeks, from the REACH-3 trial, in the economic model. This approach favours belumosudil+BAT as a patient could first respond to treatment after more than 24 weeks of treatment. The company provided best ORR at 24 weeks in the belumosudil+BAT trials in response to clarification question (CQ) A11 and this is reported in Section 3.5.2.

In the model, the company used the quaternary DOR reported in the belumosudil+BAT trials. In Section B.3.3.5 of the CS, they state that this was selected for comparability reasons as it best matches the outcome reported in REACH-3. It is unclear to the EAG whether primary/secondary DOR or quaternary DOR best matches the DOR used in REACH-3. Quaternary DOR allows for multiple response/lack of response episodes and this appears to have a substantial impact on the effect estimate. See Section 3.5.5 for a critique of this analysis.

The company did not state that ROCKstar and KD025-208 used mortality linked to relapse as a failure event within FFS as it was in REACH-3. However, the company confirmed at the clarification stage that mortality linked to relapse would be a failure event and this would be categorised as "recurrent malignancy". Therefore, the EAG consider the definitions of FFS is consistent across the three studies.

2.3.5 Subgroups to be considered

The final scope suggests the following subgroups be considered, if the evidence allows:

- Different organs or tissues affected by chronic GVHD;
- Number and type of previous treatments.

In Section B.3.12 of the CS, the company explain that the subgroups in the final scope have not been considered in this submission. The company also state that, "… *based on the above pharmacokinetics and clinical trial outcomes evidence, we do not expect belumosudil to be more clinically or cost effective than BAT in the subgroups suggested in the scope and therefore no subgroups have been considered in this submission*."

The EAG understand that there is little data on the number and type of previous treatments of patients in the REACH-3 trial and, as such, it is not possible to undertake comparative analysis for this subgroup. However, there are data in the belumosudil trials and the REACH-3 trial addressing different organs or tissues affected by chronic GVHD. At the clarification stage, the company provided a table of response by organ at week 24 using the naïve comparison to REACH-3, this is presented in Section 3.5.3.

3 Clinical effectiveness

3.1 Critique of the methods review

The company conducted a systematic literature review (SLR) to identify studies reporting on the clinical efficacy and safety of treatment options for adult patients with chronic graft vs host disease (cGvHD) after allogenic haematopoietic stem-cell transplantation (alloHSCT), who have failed at least one prior line of therapy (LOT). The company detailed the methods of the SLR in Appendix D of the company submission (CS), and the EAG's critique is presented in Table 14. The company carried out the SLR in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[14] and methods published in the Cochrane Handbook for Systematic Reviews of Interventions.[15]

There are a number of inconsistencies between the objective of the SLR and the decision problem in the NICE final scope. The stated objective of the SLR was to identify studies in adult patients with cGvHD after alloHSCT, who have failed at least one prior LOT.

Firstly, the population relevant to this appraisal also includes adolescents, 12–18 years old (12-18yo), but this population is not included in the SLR. The search strategies in MEDLINE, Embase, and CENTRAL are consistent with the stated objective of the SLR and contain lines that remove studies in adolescents. In addition, the inclusion criteria are limited to studies in adults.

Secondly, the company were aware the included population for the SLR, patients failing at least one prior LOT, was broader than the population of interest for the submission. They explain that this is to ensure that no studies reporting data on the population of interest were missed as the definition of prior therapies and how it is reported can be variable in the scientific literature.

The EAG also has some concerns linked to MEDLINE (Line 13), Embase (Line 16), and CENTRAL (line 13) search strategies. This line in each strategy uses the 'NOT' operator to remove papers that refer to acute GvHD, GvHD prophylaxis, and GvHD prevention in the title or abstract. Relevant cGvHD studies may mention these aspects of GvHD care in the paper's title or abstract.

The EAG considers it unlikely that the company would have missed any relevant studies involving treatment with belumosudil. However, an aim of the review was to find studies that could be used as a comparator to belumosudil+BAT and the SLR deliberately excludes relevant studies linked to adolescents and potentially removal of relevant studies using the 'NOT' operator. However, he EAG's

clinical experts did not identify any relevant studies that were missed in the SLR and could have been used in the analysis.

The searches were conducted in January 2023. A total of 26 unique studies that were present in 38 publications were eligible for inclusion in the SLR. Of the 26 primary publications, three were RCTs and were appraised using The Cochrane Risk of Bias Assessment Tool 2.0.[16] Twenty studies were single-arm, non-randomised trials that were appraised using the Downs and Black Checklist.[17] The company noted that three single-arm studies were reported in conference abstracts and did not appraise these studies.

In the submission, the company focuses on evidence from the belumosudil+BAT trials (ROCKstar and KD025-208). However, the ROCKstar and KD025-208 trials did not provide suitable comparator data for established clinical management and so the company explored alternative sources of comparator data. As discussed in Sections 2.3.3 and 3.4.1, the company use the SLR to assess the feasibility of a population adjusted indirect comparison (PAIC) and a matching adjusted indirect comparison (MAIC). The company also used the studies identified in the SLR to explore the most appropriate study to use as a trial-based best available therapy (BAT) control arm through a naïve direct comparison.

Table 14. A summary of the EAG's critique of the systematic literature review

Systematic review step	Section of CS in which methods are reported	EAG's assessment of robustness of methods
Data sources	Appendix D.1	**The EAG considers the sources and dates searched to be comprehensive.** Databases searched: • MEDLINE and MEDLINE-In-Process • Embase • Cochrane Database of Systematic Reviews (CDSR; reviews only, not including protocols) • Cochrane Central Register of Controlled Trials (CENTRAL) Grey literature searches were also conducted to identify recent, relevant research that may not have been published in peer-reviewed journals. Specifically, clinicaltrials.gov and the conference proceedings from four key conferences listed below were searched (from 2019 to 18 January 2023). • European Hematology Association • American Society of Hematology • International Society for Laboratory Hematology • European Society for Medical Oncology

Search strategies	Appendix D.1.1	**The EAG considers the search strategy could fail to identify evidence relevant to the decision problem.** The search strategies for the literature review used free-text keywords, MeSH and EMTREE terms for the population and interventions of interest. • MEDLINE (Line 13). Embase (Line 16), CENTRAL (line 13): this line removes papers that refer to acute GvHD, GvHD prophylaxis, and GvHD prevention in the title or abstract. It is possible relevant cGvHD studies may mention these aspects of GvHD care in the paper's title or abstract. • MEDLINE (Line 14) , Embase (Line 17), CENTRAL (line 14): this line removes studies linked to adolescents. Studies in adolescents are relevant to the decision problem and may be missed.
Inclusion criteria	Appendix D.1.2 (Table 5)	**The inclusion criteria were limited to adults with cGvHD but adolescents (12–18 year olds) are relevant to the decision problem and were potentially missed.** The eligibility criteria for the SLR were adult patients with chronic GVHD after allogeneic haematopoietic cell transplant who have failed at least one prior LOT. Failure of a single prior LOT is broader than the NICE final scope. The interventions/comparators were appropriate on the advice of the EAG's clinical experts. Outcomes were in line with those defined by NICE in the final scope.[5] Records were limited to English language studies.
Screening	Appendix D.1.2	**The EAG considers the reporting of methods for screening to be adequate.** The screening process involved two stages, with dual screening conducted by two independent investigators and any discrepancies resolved by a third investigator at both stages. Full-text articles of abstracts that were deemed relevant during the first level of review were retrieved and reviewed.
Data extraction	Appendix D.1.2	**The EAG is satisfied with the data extraction process** Data from the included studies were extracted into a single data extraction template designed in Microsoft Excel. Data were independently extracted from each included study by a single investigator, with validation performed by a second, senior investigator.
Tool for quality assessment of included study or studies	Appendix D.3	**The EAG considers the company's choice of quality assessment tool for RCTs and non-randomised studies to be reasonable** The company use the Cochrane Risk of Bias Assessment Tool 2.0 (RoB 2) for the three included RCTs and the Downs and Black (D&B) checklist for 20 single-arm studies. The company state that quality assessment was not conducted on three studies as they were published as conference abstracts because there were insufficient details available for quality assessment. RoB 2 and the D&B checklist are designed to be used for comparative studies but do have limited applicability to single arm studies.

Abbreviations: EAG, External Assessment Group; LOT, line of therapy; SLR, systematic literature review.

3.2 Critique of trials of the technology of interest

In this section, the EAG critiques the ROCKstar and KD025-208 trials as the primary source of data in patients with cGvHD who were treated with belumosudil+BAT, for the economic model. The

economic model uses data from the September 2022 data cut in the ROCKstar trial and the subgroup of patients in KD025-208 who had received ≥2 prior LOTs.

Table 15. EAG's summary of the design, conduct and analysis of ROCKstar and KD025-208

Aspect of trial design or conduct	Section of CS in which information is reported	EAG's critique
Randomisation	B.2.3.1.1	**ROCKstar: appropriate** Treatment consisted of belumosudil 200 mg QD+BAT (n=77) or 200 mg BID+BAT (n=75) administered orally in subjects with cGvHD. Randomization was stratified (1:1) by cGvHD severity and prior exposure to ibrutinib. **KD025-208: non-randomised clinical trial** Patients were enrolled into three sequential cohorts: cohort one received belumosudil 200 mg QD+BAT (n=17), cohort two received belumosudil 200 mg BID+BAT (n=16), and cohort three received belumosudil 400 mg QD+BAT (n=21).
Concealment of treatment allocation	NR	**ROCKstar: unclear** No allocation concealment method was reported. **KD025-208: no allocation concealment.** Patients were enrolled into sequential cohorts and there was no allocation concealment.
Eligibility criteria	B.2.3.1.2	**ROCKstar: current eligibility criteria are representative of the population eligible for belumosudil in the UK population. However, adolescents (12–18 year olds) were only eligible after a protocol amendment.** Summary of inclusion criteria: • Patients ≥12 years who had received alloHSCT and were experiencing persistent cGvHD manifestations after receiving 2 to 5 prior LOTs. The inclusion of adolescents was only allowed after a protocol amendment in June 2020. **KD025-208: eligibility criteria differed from the NICE final scope. Adolescents were not recruited, and patients who had a minimum of one prior LOT were recruited.** Summary of inclusion criteria: • Patients ≥18 years who had received allogeneic bone marrow transplant or alloHSCT and were experiencing persistent chronic GVHD manifestations after receiving 1 to 3 prior LOTs. Full details of the eligibility criteria for ROCKstar and KD025-208 are available in the CS, Table 9.
Blinding	B.2.3.1	**ROCKstar and KD025-208 were not blinded studies** • ROCKstar: open-label study; • KD025-208: open-label study.

Baseline characteristics	B.2.3.2.1 and B.2.3.2.2	**Baseline characteristics of ROCKstar and KD025-208 were appropriate given the eligibility criteria** The EAG's clinical experts considered the baseline characteristics reported in ROCKstar and KD025-208 were appropriate for the population sampled. They were reflective of a USA cGvHD population. **ROCKstar** The EAG considers there to be an imbalance in the proportions of males and females between the study arms with a higher proportion of males enrolled in the 200 mg QD arm (64%) compared to the 200 mg BID arm (50%). The EAG also notes that randomisation was not stratified by sex. With the exception of sex, the EAG considers the baseline characteristics to be reasonably well balanced between the treatment arms. The applicability of the baseline characteristics in ROCKstar and KD025-208 to the decision problem and UK practice is discussed in Section 2.3.1.
Dropouts	Appendix D.2.1	**ROCKstar** The proportion of patients discontinuing belumosudil+BAT treatment were balanced between study arms. In the 200 mg QD arm, 50.0% of patients discontinued, and in the 200 mg BID, 56.3% of patients discontinued. The reasons for discontinuation were consistent between arms. Discontinuation was mainly due to cGvHD progression, voluntary withdrawal, or adverse events. **KD025-208** A high proportion of patients discontinued belumosudil+BAT treatment, but this was balanced between study arms. • 200 mg QD arm: 88% of patients discontinued; • 200 mg BID, 88% of patients discontinued; • 400 mg QD arm: 86% of patients discontinued. Discontinuation was primarily due to disease progression.
Statistical analysis		
Sample size and power	B.2.3.1.1 and B.2.4	**The sample size used in the pooled analysis is appropriate** **ROCKstar** The sample size was based on the primary efficacy endpoint, with one planned interim analysis and a target ORR of 55%. Approximately 63 participants per treatment arm were required to provide 90% power to yield a 95% CI of ORR that excluded 30% as the lower bound. **KD025-208** A sample size of 16 subjects per dose group was planned to provide >90% chance of ≥1 subject experiencing an AE with an underlying rate of ≥14%. Assuming a best ORR of 25%, the study was expected to have approximately 90% probability to show a response in ≥2 patients per dose group. The sample size of the in the pooled analysis (September 2022 data cut and ≥2 prior LOTs) was:

		• 92 patients in the belumosudil 200 mg QD+BAT; • 84 patients in the belumosudil 200 mg BID+BAT.
Handling of missing data	Appendix B.2.4.1.3, B.2.4.2.3, D.2.1, D.2.2	**Unclear how missing data were handled in ROCKstar but there was a small proportion of patients missing after the first data cut.** **ROCKstar** It is unclear if any adjustments were made to handle missing data in the ROCKstar study. However, the participant flow presented in D.2.1 indicates that a total of 119 (90.2%) subjects were participating in the study, either being actively treated with belumosudil+BAT or being followed for survival, in the February 2020 data cut. **KD025-208** In Section B.2.4.2.3 of the CS, it is stated that no missing values were imputed. At the February 2020 data cut, a total of 37 (68.5%) subjects were participating in the study, either being actively treated with belumosudil+BAT or being followed for survival.
Outcome assessment	B.2.4 and B.2.10	**Appropriate** The ROCKstar and KD025-208 (≥2 LOT) trials outcomes were reported in pooled and combined analysis. The outcome data used in the company's economic model (base case) came from the pooled analysis. In this analysis, the belumosudil 200 mg QD+BAT arms from each study were pooled, and the belumosudil 200 mg BID+BAT arms were pooled. The data from the KD025-208 trial used in the efficacy analysis came from the trial subgroup who had ≥2 prior LOTs. The company also presented combined analyses where the pooled belumosudil 200 mg QD=BAT and pooled belumosudil 200 mg BID+BAT arms were combined together to form a single belumosudil 200 mg+BAT treatment arm. The EAG consider this a reasonable approach as the pooled belumosudil 200 mg QD+BAT and pooled belumosudil 200 mg BID+BAT management arms had similar efficacy estimates. All safety analyses were reported in the safety population which consisted of all patients who were randomised and who received ≥ 1 dose of study medication with assignment by actual treatment received. It was reported in a pooled analysis. The company note that in order to inform the economic model, an additional analysis was performed to identify Grade ≥3 AEs occurring in patients who received ≥2 prior LOTs from the pooled analysis.

Abbreviations: AE, adverse event; alloHSCT, Allogeneic hematopoietic stem cell transplant; BID, twice daily; CS, company submission; EAG, External Assessment Group; LOTs, lines of therapy; ORR, overall response rate; QD, once daily.

3.2.1 Co-administering PPIs

The ROCKstar and KD025-208 trials each had a belumosudil 200 mg once daily (QD)+BAT treatment arm and a belumosudil 200 mg twice daily (BID)+BAT treatment arm. Belumosudil 200 mg QD is the indicated dose appropriate for patients 12 years and over. Belumosudil 200 mg BID dose is indicated in people who co-administer with proton pump inhibitors (PPIs) or strong CYP3A inducers. However,

in the ROCKstar and KD025-208 trials, patients were assigned belumosudil dose irrespective of their use of PPIs or strong CYP3A inducers. In response to clarification question A4, the company note that use of strong CYP3A4 inducers was prohibited in the ROCKstar protocol and other CYP3A4 inhibitors/inducers were to be used with caution. Therefore, co-administration of strong CYP3A inducers for patients in the belumosudil+BAT trials is expected to be very low. However, PPIs were widely used in the belumosudil+BAT trials, 50.6% of patients in the pooled belumosudil 200 mg QD+BAT arm and 50.0% of patients in the pooled belumosudil 200 mg BID+BAT arm, were co-administering with PPIs. Given the usage of PPIs, the belumosudil dose received by 50.3% of patients is not reflective of how they would currently be treated in UK clinical practice. The company provided overall response rate (ORR) in the subgroup of people in ROCKstar in each dosing arm who were, or were not, receiving concomitant PPI. This is presented in Section 3.3.5.

3.2.2 Population analysed

The ROCKstar and KD025-208 trials outcomes (for ≥2 LOTs) were reported in pooled and combined analysis. The outcome data used in the company's economic model (base case) came from the pooled analysis. In the pooled analysis, the belumosudil 200 mg QD+BAT treatment arms from each study were pooled, and the belumosudil 200 mg BID+BAT arms were pooled. The data from the KD025-208 trial used in the efficacy analysis came from the trial subgroup who had ≥2 prior LOTs.

The company also presented combined analyses where the pooled belumosudil 200 mg QD+BAT and pooled belumosudil 200 mg BID+BAT arms were combined together to form a single belumosudil 200 mg with BAT arm. The EAG considers this a reasonable approach as the pooled belumosudil 200 mg QD+BAT and pooled belumosudil 200 mg BID+BAT arms had similar efficacy estimates. This analysis was not used in the economic model presented by the company or in the EAG's base case.

3.2.3 Baseline characteristics

The baseline characteristics of the patients (≥2 LOT) in the combined ROCKstar (September 2022 data cut) and KD025-208 trials are presented in Table 16. The EAG's clinical experts considered the patients recruited across the belumosudil trials to be broadly appropriate, although they were more representative of a USA cGvHD population than a UK cGvHD population.

The EAG's clinical experts also commented on what can be inferred from the baseline characteristics of a person with cGvHD. A patient's sex and age are not thought to have a significant impact on

cGvHD severity or response to GvHD treatment. However, older patients may be less tolerant to treatment while adolescents (12-18yo), tend to be more responsive to treatment.

Patients may have a history of acute GvHD (aGvHD), prior to developing cGvHD. A history of aGvHD can indicate a patient is in worse physical condition, and they may have already tried a number of the treatments used for cGvHD. It is unclear if this would affect the efficacy of other treatments but could lead to disease that is more complex to treat.

After initial treatment with corticosteroids with or without CNIs, patients with cGvHD may cycle through treatments before finding an effective treatment, or combinations of treatments. This can take a period of months or years and living with cGvHD without effective treatment leads to more developed disease with more organs involved, and more fibrosis that is difficult to reverse. Therefore, time from cGvHD diagnosis to effective treatment, and the number of prior LOTs, can be signals of disease that is more developed and more complex to treat.

The EAG's clinical experts noted that different organ systems have different responses to treatment. Lung, liver and gut are known to be more difficult to treat with lower anticipated response rates. These are often the key organs that are associated with the highest disease morbidity and are the targets for treatment.

Table 16.Baseline characteristics of pooled analysis of ROCKstar and KD025-208 (reproduced from clarification response A1)

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)			
	200 mg once daily (n=92)	200 mg twice daily (n=84)	Combined 200 mg (N=176)
Median age (range), years	53.0 (20 to 77)	57.0 (18- to 7)	55.0 (18 to 77)
Males, n (%)	60 (65.2%)	42 (50.0%)	102 (58.0%)
GVHD prophylaxis after transplant, n (%)			
None	0	1 (1.2%)	1 (0.6%)
CNI only	5 (5.4%)	7 (8.3%)	12 (6.8%)
CNI + methotrexate	38 (41.3%)	35 (41.7%)	73 (41.5%)
CNI + methotrexate + other	10 (10.9%)	7 (8.3%)	17 (9.7%)
CNI + MMF	11 (12.0%)	14 (16.7%)	25 (14.2%)
CNI + MMF + other	5 (5.4%)	3 (3.6%)	8 (4.5%)
CNI + MMF + ATG	0	1 (1.2%)	1 (0.6%)
CNI + sirolimus	8 (8.7%)	8 (9.5%)	16 (9.1%)
CNI + corticosteroids	2 (2.2%)	1 (1.2%)	3 (1.7%)
Other regimen	12 (13.0%)	7 (8.3%)	19 (10.8%)
HLA matching of donor/recipient, n (%)			

Matched	80 (87.0%)	78 (92.9%)	158 (89.8%)
Partially matched	11 (12.0%)	5 (6.0%)	16 (9.1%)
Unknown	0	1 (1.2%)	1 (0.6%)
Missing	1 (1.1%)	0	1 (0.6%)
History of acute GVHD, n (%)	61 (66.3%)	63 (75.0%)	124 (70.5%)
Time from chronic GVHD diagnosis to enrolment, median (range), months	26.66 (1.6 to 162.4)	29.91 (3.7 to 144.1)	28.14 (1.6 to 162.4)
NIH chronic GVHD severity[a] n (%)			
Severe	66 (71.7%)	58 (69.0%)	124 (70.5%)
Moderate	24 (26.1%)	26 (31.0%)	50 (28.4%)
Mild	2 (2.2%)	0	2 (1.1%)
Organ involvement, n (%)			
≥4 organs involved	49 (53.3%)	44 (52.4%)	93 (52.8%)
≥6 organs involved	15 (16.3%)	10 (11.9%)	25 (14.2%)
Skin	68 (73.9%)	59 (70.2%)	127 (72.2%)
Joints/fascia	75 (81.5%)	69 (82.1%)	144 (81.8%)
Eyes	52 (56.5%)	56 (66.7%)	108 (61.4%)
Mouth	70 (76.1%)	63 (75.0%)	133 (75.6%)
Lungs	32 (34.8%)	27 (32.1%)	59 (33.5%)
Oesophagus	16 (17.4%)	11 (13.1%)	27 (15.3%)
Upper GI	25 (27.2%)	13 (15.5%)	38 (21.6%)
Lower GI	8 (8.7%)	8 (9.5%)	16 (9.1%)
Liver	10 (10.9%)	5 (6.0%)	15 (8.5%)
No. of organs involved, median (range)	4.0 (0 to 7)	4.0 (1 to 7)	4.0 (0 to 7)
Refractory to prior LOT, n (%)	60 (80.0%)	43 (65.2%)	103 (73.0%)
Number or prior lines of therapy, n (%)			
1	0	0	0
2	29 (31.5%)	18 (21.4%)	47 (26.7%)
3	27 (29.3%)	26 (31.0%)	53 (30.1%)
4	20 (21.7%)	18 (21.4%)	38 (21.6%)
5	14 (15.2%)	20 (23.8%)	34 (19.3%)
≥6	2 (2.2%)	2 (2.4%)	4 (2.3%)
Median	3.0	3.0	3.0
Prior systemic chronic GVHD therapies, n (%)[b]			
Prednisone	91 (98.9%)	83 (98.8%)	174 (98.9%)
Tacrolimus	56 (60.9%)	53 (63.1%)	109 (61.9%)
Sirolimus	43 (46.7%)	41 (48.8%)	84 (47.7%)
ECP	5 (5.4%)	4 (4.8%)	9 (5.1%)
Ibrutinib	27 (29.3%)	27 (32.1%)	54 (30.7%)
Mycophenolate mofetil	22 (23.9%)	22 (26.2%)	44 (25.0%)
Rituximab	23 (25.0%)	16 (19.0%)	39 (22.2%)
Ruxolitinib	29 (31.5%)	26 (31.0%)	55 (31.3%)

Concomitant systemic chronic GVHD therapies, n (%)			
Systemic hormonal preparations	91 (98.9%)	82 (97.6%)	173 (98.3%)
ECP	22 (23.9%)	28 (33.3%)	50 (28.4%)
Tacrolimus	30 (32.6%)	26 (31.0%)	56 (31.8%)
Sirolimus	19 (20.7%)	20 (23.8%)	39 (22.2%)
MMF	11 (12.0%)	3 (3.6%)	14 (8.0%)
Imatinib	1 (1.1%)	1 (1.2%)	2 (1.1%)
Rituximab	1 (1.1%)	0	1 (0.6%)
Ruxolitinib	1 (1.1%)	0	1 (0.6%)

[a] Severity was determined using the NIH Global Severity of chronic GVHD scoring.

[b] This table includes the most common therapies for chronic GVHD (≥10%), as well as ECP.

Abbreviations: ATG, antithymocyte globulin; CNI, calcineurin inhibitor; ECP, extracorporeal photopheresis; GI, gastrointestinal; GvHD, graft-versus-host disease; HLA, human leukocyte antigen; MMF, mycophenolate mofetil; NIH, National Institutes of Health.

3.2.4 Adolescents: 12 to 18 years old

As noted in Section 2.3.1, adolescents (12-18yo) are relevant to the decision problem but KD025-208 did not recruit adolescents (12-18yo) and no data in adolescents (12-18yo) is reported in the ROCKstar September 2022 data cut. In clarification question A2, the EAG requested the basis on which belumosudil has been granted marketing authorisation in adolescents (12-18yo) and a rationale for the use of the same dose in adults and adolescents (12-18yo).

In response, the company stated that the manageable safety profile, in particular in relation to adverse reactions of concern to adolescent patients, can reliably be expected to be the same in adolescents as in adults. This is due to the similarity of the disease pathophysiology, general response to treatment, pharmacokinetics (PK) modelling and flat exposure-safety relationship. In addition to this, the efficacy demonstrated in adults, not just the ORR but the reduction of steroid dose, indicates that adolescent patients would benefit from treatment with belumosudil+BAT. The choice of dose is supported by the population PK modelling showing a large overlap between adult and adolescent area under the curve (AUC), as well as the flat exposure–response and exposure–safety relationships. Therefore, the benefit/risk balance is positive for belumosudil+BAT in the treatment of cGvHD in adolescent patients at a dose of 200mg QD.

Given there are no efficacy data in adolescents (12-18yo), the EAG is unclear on the effectiveness of belumosudil+BAT in this population. All of the efficacy and safety data in the CS relates to adults and the EAG cannot confirm the same results would be seen in adolescents (12-18yo).

3.2.5 Statistical analysis

The company provided a summary of statistical analyses used for the belumosudil+BAT trials (Table 17). In the ROCKstar trial, the Hochberg procedure was used for multiplicity adjustment for the primary endpoint; best overall response rate (ORR) at any time. The method of multiplicity adjustment is presented in Figure 7 in the CS. The EAG agrees with the company that the multiplicity adjustment is appropriate for this outcome. The sample size and power calculations are appropriate for trials without control groups.

Table 17. Summary of statistical analyses (reproduced from Table 12, CS)

Trial number (acronym)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
ROCKstar	The primary efficacy endpoint was best ORR at any time. The null hypothesis was that ORR was ≤30%.	The Hochberg procedure was used for multiplicity adjustment for the primary endpoint. The primary analysis was conducted using the mITT population. Descriptive statistics, without multiplicity adjustment, were provided for all secondary and exploratory endpoints.	The sample size was based on the primary efficacy endpoint, with one planned interim analysis and a target ORR of 55%. Approximately 63 participants per treatment arm were required to provide 90% power to yield a 95% CI of ORR that excluded 30% as the lower bound.	Reason for discontinuation or withdrawal was documented in the eCRF according to treatment group.
KD025-208	The primary efficacy endpoint was best ORR at any time. A null hypothesis was not stipulated.	The study was not powered to show significant differences between dose groups with respect to efficacy, AEs or PD analyses. The primary analysis was conducted using the safety population.	A sample size of 16 subjects per dose group was planned to provide >90% chance of ≥1 subject experiencing an AE with an underlying rate of ≥14%. Assuming a best ORR of 25%, the study was expected to have approximately 90% probability to show a response in ≥2 patients per dose group.	Reason for discontinuation or withdrawal was documented in the eCRF according to dose group.

mITT population: all randomised participants who received at least one dose of study drug

Abbreviations: AE, adverse events; eCRF, electronic case report form; ORR, overall response rate; PD, pharmacodynamics.

3.2.6 Participant flow through ROCKstar and KD025-208

The participant flow was presented for the ROCKstar trial (February 2020 data cut) and KD025-208 trial (February 2020 data cut) in Appendix D.2 of the CS. The company state in D.2.1 that as of February 2020, a total of 119 (90.2%) subjects were participating in the study, either being actively treated with belumosudil+BAT or being followed for survival. The most common reasons for treatment discontinuation were progression of cGvHD in 16 (12.1%) patients and 14 (10.6%) patients decision to discontinue. In total, 13 (9.8%) patients discontinued from the study for the following reasons: 8 (6.1%) patients died, patient decision in 3 (2.3%), 1 (0.8%) patient was lost to follow-up, and 1 (0.8%) patient had missing data.

Participant flow for KD025-208 trial, presented in D.2.2, is in the full trial population and not only the ≥2 LOT subgroup used in the analysis. At the February 2020 data cut, 37 (68.5%) patients were participating in the study, either being actively treated with belumosudil+BAT or being followed for survival. The most common reasons for treatment discontinuation were progression of cGvHD in 22 (40.7%) patients and patient decision in 11 (20.4%) patients.

The EAG note that the company could provide participant flow using the September 2022 data cut from ROCKstar, and the final participant flow from KD025-208 which completed in May 2022. This would provide a better measure of missing data from the studies.

3.3 Critique of the clinical effectiveness analysis and interpretation

The results of the ROCKstar and KD025-208 trials are presented below. Where possible the results reported are from the September 2022 data cut from the ROCKstar trial and the ≥2 LOT subgroup from KD025-208 trial.

3.3.1 Response to treatment, failure-free survival , and mortality

Results for response to treatment outcomes, failure-free survival (FFS) and mortality outcomes reported in the CS for the pooled and combined belumosudil+BAT analysis are presented in Table 18 below. Any patient lost to follow-up without any response assessment was counted as a non-responder in the pooled analysis. The key efficacy inputs in the economic model are FFS, overall survival (OS), time to response (TTR), ORR, duration of response (DOR), and time to treatment discontinuation (TTD). The pooled analyses for each dosing regimen show a consistent effect for ORR, FFS, and OS. There appears to be a small additional benefit for belumosudil 200 mg BID+BAT in DOR and TTD compared to belumosudil 200 mg QD+BAT, but this was not statistically significant.

Table 18. Results of pooled efficacy analysis (3-year analysis)

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)			
Outcome	**200 mg once daily (n=92)**	**200 mg twice daily (n=84)**	**Combined 200 mg (N=176)**
Median time to response, weeks (range)	7.71 (3.7 to 80.1)	5.21 (3.7 to 40.1)	7.71 (3.7 to 80.1)
Best ORR,[a] n (%) CR PR	67 (72.8%) 4 (4,3%) 63 (68.5%)	62 (73.8%) 2 (2.4%) 60 (71.4%)	114 (73.1%) 6 (3.4%) 123 (69.9%)
Median DOR in responders (primary/secondary)[b] weeks (95% CI)	23.9 (12.14 to 50.43)	26.0 (17.14 to 45.43)	25.7 (17.29 to 36.14)

Median DOR in responders (quaternary), weeks (95% CI)	69.9 (28.29 to 176.00)	74.3 (35.00 to 114.57)	69.9 (40.43 to 95.43)
Median FFS, months (95% CI)	15.2 (9.26 to 24.02)	16.6 (11.27 to 35.88)	15.4 (12.42 to 22.74)
FFS, % (95% CI) FFS at 6 months FFS within 12 months FFS within 24 months	 74% (0.64 to 0.82) 56% (0.45 to 0.65) 41% (0.30 to 0.51)	 78% (0.68 to 0.86) 61% (0.49 to 0.70) 42% (0.31 to 0.53)	 76% (0.69 to 0.82) 58% (0.50 to 0.65) 41% (0.33 to 0.49)
Median OS (months)	n/a	n/a	n/a
OS, % (95% CI) OS within 12 months OS within 24 months	 91% (83 to 95) 86% (76 to 92)	 91% (83 to 96) 84% (74 to 91)	 91% (86 to 95) 85% (78 to 90)
Median TTD, months (range)	9.18 (0.5 to 64.2)	11.78 (0.4 to 39.6)	10.38 (0.4 to 64.2)
Median TTR in responders, weeks (range)	n=59 7.86 (3.7 to 80.1)[c]	n=55 5.29 (3.7 to 40.1)[c]	n=114 7.86 (3.7 to 80.1)[c]

[a] Best ORR at any time was defined as the percentage of patients that had either complete response or partial response, using the 2014 NIH Consensus Criteria

[b] DOR was measured from the time of first documentation of response to the time of first documented deterioration from best response (primary), to the time of first documented lack of response (secondary), or to the time of first documented lack of response with durations summed for multiple response/lack of response episodes (quaternary)

[c] August 2021 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)

Abbreviations: CI, confidence interval; CR, complete response; DOR, duration of response; FFS, failure-free survival; GI, gastrointestinal; NA, not available; NIH, National Institutes of Health; ORR, overall response rate; OS, overall survival; PR, partial response; TTD, time to treatment discontinuation.

3.3.1.1 Duration of response

Quaternary DOR is used in the company's economic model. Median (quaternary) DOR was 62.3 weeks in the pooled belumosudil 200 mg QD+BAT arm and 74.3 weeks in the pooled belumosudil 200 mg BID+BAT arm. Median (primary/secondary) DOR substantially lower, 22.1 weeks in the pooled belumosudil 200 mg QD+BAT arm and 24.1 weeks in the pooled belumosudil 200 mg BID+BAT arm. This discrepancy is further discussed in Section 3.5.5.

3.3.2 Response by organ system

At the clarification stage, the company provided best response from baseline by organ system in the combined analysis of the belumosudil+BAT trials. Of the organ systems assessed, belumosudil+BAT was most effective in joints/fascia, mouth, oesophagus, and upper/lower GI. The proportion responding was lowest for liver, lungs, and skin compared to the other organ systems assessed. The company also provided a table comparing response by organ system at week 24 as was reported in the REACH-3 trial. This is presented in Section 3.5.3.

Table 19. Response from baseline of organ involvement in the combined ROCKstar and KD025-208

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)					
Organ	**Number affected, n**	**Responders, n (%)**	**CR, n (%)**	**PR, n (%)**	**LR, n (%)**
Skin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Joints/fascia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eyes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lungs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oesophagus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upper GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lower GI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Liver	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CR, complete response; GI, gastro-intestinal; PR, partial response; LR, lack of response.

3.3.3 *Discontinuations and reductions of concomitant medications*

The belumosudil+BAT trials reported reduction in corticosteroid use during the trials in the pooled and the combined analysis (Table 20). The mean reduction in corticosteroid dose from baseline in the combined analysis was 51.0%, with 26.3% of patients discontinuing treatment.

Table 20. Discontinuations and reductions of concomitant medications in pooled analysis (2 year analysis)

August 2021 data cut data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)			
Outcome	**200 mg once daily (n=81)**	**200 mg twice daily (n=87)**	**Combined 200 mg (N=156)**
CS reduction, n (%)	55 (67.9%)	53 (70.7%)	108 (69.2%)
Mean change in CS dose from baseline, %	-49.72%	-52.24%	-50.95%
CS discontinuation, n (%)	23 (28.4%)	18 (24.0%)	41 (26.3%)
CNI discontinuation in ROCKstar, %	21%	33%	27%

Abbreviations: CNI, calcineurin inhibitors; CS, corticosteroid; NR, not reported.

3.3.4 *Quality of life*

In ROCKstar and KD025-208, quality of life (QoL) and symptoms were measured using the 7-day Lee Symptom Scale (LSS) summary score. A clinically meaningful improvement from baseline was defined as a reduction of ≥7 points and reported for each trial in Table 21 and Table 22 below. Fifty-seven percent of patients in the pooled belumosudil 200 mg QD arm+BAT, and 60% of patients in the belumosudil 200 mg BID arm+BAT, achieved a reduction of ≥7 points on the LSS summary score.

ROCKstar also assessed QoL with an exploratory endpoint using the Patient-Reported Outcomes Measurement Information System Global Health (PROMIS-GH) questionnaire. The PROMIS-GH questionnaire consists of 10 items that measure physical health, physical functioning, general mental health, emotional distress, satisfaction with social activities and relationships, ability to carry out usual social activities and roles, pain, fatigue, and overall quality of life. Two 4-item summary scores: a Global Physical Health (GPH) score and a Global Mental Health (GMH) score can be used. GPH and GMH scores were mapped to the EQ-5D-3L measure to provide utility scores for use in economic model (Section B.3.4.2 of the CS).[13]

The company reported the number of patients who had an improvement of ≥4.7 points, previously identified as a clinically meaningful difference in cGvHD.[18] In the combined analysis of ROCKstar, 45.4% of patients achieved a clinically meaningful improvement in GMH and 50.7% in GPH.

Table 21. ROCKstar: Improvement in HRQoL scores (mITT)

September 2022 data cut			
Outcome	[redacted]	[redacted]	[redacted]
Improvement in LSS score ≥7 points from baseline Overall, n (%) Responders, n/N (%) Non-responders, n/N (%)	[redacted]	[redacted]	[redacted]
Improvement in PROMIS raw mental health score ≥4.7 points from baseline, n (%)	[redacted]	[redacted]	[redacted]
Improvement in PROMIS raw physical health score ≥4.7 points from baseline, n (%)	[redacted]	[redacted]	[redacted]

mITT: randomized subjects who received ≥1 dose of belumosudil.

Abbreviations: HRQoL, health-related quality of life; LSS, Lee Symptom Scale; mITT, modified intent-to-treat population; NR, not reported; PROMIS, Patient-Reported Outcomes Measurement Information System.

Table 22. KD025-208: Improvement in LSS score (mITT)

Outcome	200 mg once daily (n=17)	200 mg twice daily (n=16)	400 mg once daily (n=21)	Total (N=54)
Improvement in LSS score ≥7 points from baseline Overall, n (%) Responders, n/N (%) Non-responders, n/N (%)	 9 (53%) 8/11 (73%) 1/6 (17%)	 7 (44%) 3/11 (27%) 4/5 (80%)	 11 (52%) 9/13 (69%) 2/8 (25%)	 27 (50%) 20/35 (57%) 7/19 (37%)

mITT: randomized subjects who received ≥1 dose of belumosudil.16 subject received <2 LOT.

Abbreviations: HRQoL, health-related quality of life; LSS, Lee Symptom Scale; LOT, line of therapy; mITT, modified intent-to-treat population; NR, not reported; PROMIS, Patient-Reported Outcomes Measurement Information System.

3.3.5 Subgroup analysis: concomitant PPIs

The company provided Table 23 of ORR in the subgroup of patients who were, or were not, receiving concomitant PPIs during the trial. In the subgroup of patients co-administering with PPIs, belumosudil 200 mg QD+BAT was numerically more effective than belumosudil 200 mg BID+BAT for ORR (84.6% vs 71.4%, respectively). However, belumosudil 200 mg BID is in the recommended dose in patients co-administering with PPIs.

In the subgroup of patients who are not co-administering with PPIs, the reverse is found and belumosudil 200 mg QD+BAT is numerically less effective than belumosudil 200 mg BID+BAT in ORR (63.2% vs 80.0%, respectively). However, belumosudil 200 mg QD is the recommended dose in patients who are not co-administering with PPIs.

The EAG notes that a comparison of ORR for the two belumosudil+BAT treatment regimens was not subject to statistical significance testing by the company and cautions against drawing strong conclusions from small subgroup analyses. However, based on these results and the despite the information the company provided during the clarification stage, the EAG remains unclear about the evidence base supporting the use of different daily doses of belumosudil based on concomitant use of PPIs or strong CYP3A4 inducers.

Table 23. Subgroup analysis of ORR for patients receiving concomitant PPI in ROCKstar (mITT, 2022 data cut, reproduced from clarification response A3)

Treatment arm	200 mg QD (n=77)		200 mg BID (n=75)	
Concomitant PPI?	Yes (n=39)	No (n=38)	Yes (n=35)	No (n=40)
ORR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
CR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
PR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BID: twice per day; ORR: objective response rate; PPI: proton pump inhibitor; QD: once per day.

3.3.6 Safety

The company reported the safety profile of the patients in the safety population in the belumosudil+BAT trials (Table 24). The safety population which was defined as enrolled patients who received at least one dose of study treatment. The company also reported the Grade ≥3 AEs occurring in over 5% of patients in either treatment arm in the pooled analysis, presented in Table 25, below. The company state that for missing start or end dates of AE information, the worst or most conservative judgement was used.

Table 24. Safety profile from pooled analysis (safety population; 3-year analysis)

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)				
Outcome	200 mg QD+BAT (n=92)	200 mg BID+BAT (n=84)	400 mg QD+BAT (n=14)	Combined (N=190)
Any AE, n (%)	91 (98.9%)	84 (100.0%)	14 (100.0%)	189 (99.5%)
Any drug-related AE, n (%)	66 (71.7%)	55 (65.5%)	11 (78.6%)	132 (69.5%)
Grade ≥3 AEs, n (%)	59 (64.1%)	48 (57.1%)	10 (71.4%)	117 (61.6%)
Drug-related Grade ≥3 AEs, n (%)	18 (19.6%)	16 (19.0%)	2 (14.3%)	36 (18.9%)
SAE, n (%)	41 (44.6%)	34 (40.5%)	8 (57.1%)	83 (43.7%)
Drug-related SAE, n (%)	7 (7.6%)	5 (6.0%)	0	12 (6.3%)
Fatal AEs, n (%)	5 (5.4%)	5 (6.0%)	4 (28.6%)	14 (7.4%)
Infections and infestations (any grade), n (%)	57 (62.0%)	57 (67.9%)	9 (64.3%)	123(64.7%)

Grade ≥3, n (%)	18 (19.6%)	20 (23.8%)	3 (21.4%)	41 (21.6%)
Cytopenias[a]	16 (17.4%)	20 (23.8%)	1 (7.1%)	37 (19.5%)
Most common AEs (incidence ≥25%)				
Fatigue	39 (42.4%)	26 (31.0%)	8 (57.1%)	73 (38.4%)
Diarrhoea	41 (44.6%)	27 (32.1%)	4 (28.6%)	72 (37.9%)
Upper RTI	26 (28.3%)	28 (33.3%)	4 (28.6%)	58 (30.5%)
Nausea	31 (33.7%)	25 (29.8%)	5 (35.7%)	61 (32.1%)
Dyspnoea	28 (30.4%)	20 (23.8%)	5 (35.7%)	53 (27.9%)
Cough	23 (25.0%)	23 (27.4%)	6 (42.9%)	52 (27.4%)
Oedema peripheral	25 (27.2%)	20 (23.8%)	4 (28.6%)	49 (25.8%)
Headache	26 (28.3%)	24 (28.6%)	3 (21.4%)	53 (27.9%)
Vomiting	25 (27.2%)	14 (16.7%)	2 (14.3%)	41 (21.6%)
Muscle spasms	13 (14.1%)	15 (17.9%)	3 (21.4%)	31 (16.3%)

[a] Anaemia, thrombocytopenia, neutropenia or leukopenia or cytopenia affecting more than 1 cell line.

Abbreviations: AE, adverse events; RTI, respiratory tract infection; SAE, serious adverse events.

Table 25. Grade ≥3 AEs occurring in >5% of patients (3-year analysis)

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)			
	200 mg once daily (n=92)	200 mg twice daily (n=84)	Combined 200 mg (N=176)
Pneumonia	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]
Thrombocytopenia and decreased platelet counts	[redacted]	[redacted]	[redacted]
Hyperglycaemia	[redacted]	[redacted]	[redacted]
Gamma-glutamyl transferase increased	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]

Abbreviations: AE, adverse events; LOT, line of therapy.

3.4 Critique of trials identified and included in the indirect comparison

The belumosudil+BAT trials did not have suitable comparator groups that met the decision problem. The company explored a number of methods to find valid data to enable a comparison of belumosudil+BAT and BAT. The company investigated using a population adjusted indirect comparison (PAIC), [redacted], and using matching-adjusted indirect comparison (MAIC). These approaches are detailed in Sections 3.4.1, 3.4.2, 3.4.3, below.

3.4.1 *Population-adjusted indirect comparisons*

The company assessed the comparability of the studies included in the SLR to the belumosudil+BAT trials (ROCKstar and KD025-208) for the purposes of conducting a PAIC. This assessment included having similar inclusion/exclusion selection criteria, key study design features such as length of follow-up, timing of assessment of outcomes, outcome definitions to ensure the comparability of the endpoints, and having the sufficiently granular data on the patients in the study.

The feasibility of carrying out a PAIC using each of the included studies is reported in Table 8 of Appendix D in the CS. The company concluded that none of the studies included in the SLR were suitable for use in a PAIC. Fourteen studies were not considered suitable due to the study not being limited to patients who have had ≥2 LOT. Six studies were unsuitable as they took place in Asian countries and the company state that inclusion of such studies could create heterogeneity in patient populations and/or health systems. Four studies did not have a population that was comparable. The EAG agree with the company's feasibility assessment of the studies included in the SLR. The EAG's clinical experts were not aware of any studies that were suitable for using in a PAIC analysis with the belumosudil+BAT trials.

3.4.2 **********

[redacted]

3.4.3 Trial-based control arm

While the SLR conducted by the company did not identify a trial suitable for a PAIC, it did find a study the company deemed suitable to provide a trial-based control arm for a naïve comparison with belumosudil+BAT. The company chose the BAT arm from the Phase III, REACH-3 trial of ruxolitinib vs investigator's choice (BAT) after one prior line of therapy to allow comparison to currently available treatments in a naïve direct comparison. The EAG's clinical experts agreed REACH-3 offered a reasonable comparator arm in the absence of head-to-head data.

3.4.3.1 REACH-3

In this section the EAG assesses the comparability of the belumosudil+BAT trials to the REACH-3 BAT arm.

3.4.3.1.1 Eligibility criteria of REACH-3 and the belumosudil trials

Patients eligible for REACH-3 were alloHSCT recipients aged ≥12 years with moderate or severe glucocorticoid-refractory cGvHD. Patients treated previously with 2 or more systemic therapies for cGvHD in addition to corticosteroids with or without CNIs were ineligible.

The eligibility criteria for ROCKstar were alloHSCT recipients aged ≥12 years with persistent cGvHD after receiving 2 to 5 prior systemic LOTs. Only adults were eligible for the KD025-208 trial and they had to have persistent cGvHD after receiving 1-3 LOTs. However, the subgroup used in this analysis were the patients the KD025-208 trial who had ≥2 prior LOTs.

In Section B.3.2.3 of the CS, the company report agreement among their clinical experts that the REACH-3 control arm is a suitable proxy for an established clinical management without belumosudil arm. The company's clinical experts note that the patients in REACH-3 are at an earlier LOT and therefore are likely to have less developed disease that is more treatable. Therefore, the company's experts consider comparison with REACH-3 is a conservative approach.

Based on the baseline characteristics (Table 27) of the patients, the EAG's clinical experts were uncertain if the belumosudil+BAT arm was more complex to treat than the REACH-3 BAT arm (Section 3.4.3.1.2.1).

3.4.3.1.2 Matching-adjusted indirect comparison

After feedback from their clinical experts, the EAG were unclear if the comparison of the belumosudil+BAT trials to the REACH-3 BAT arm is a conservative approach. At the clarification stage, the EAG requested the company address this uncertainty by conducting an unanchored matching-adjusted indirect comparison (MAIC) between the belumosudil+BAT arm and the REACH-3 BAT arm.

For an unanchored indirect comparison, population adjustment methods should adjust for all possible effect modifiers and prognostic variables – limited only by the availability of these data in the published comparator trial. The company has IPD on the patients in the belumosudil+BAT trials and there are 176 patients in the combined treatment arm. The EAG considers combining the two dosing regimens is a valid approach due to the closely matched efficacy reported for the pooled belumosudil 200 mg QD+BAT arm and the pooled belumosudil 200 mg BID+BAT arm. This doubles the number of patients who could be used to match to the REACH-3 BAT arm.

The company had assessed the feasibility of undertaking a PAIC, prior to deciding to use a trial-based control arm in a naïve comparison. In response to the EAG's request for an MAIC, the company reiterated and expanded upon their reasoning why differences in study inclusion criteria and populations did not allow for a robust MAIC. In Table 26 below, reproduced from the clarification response, the company note clinically meaningful prognostic variables and effect modifiers that either, cannot be matched, or for which there is limited overlap between the REACH-3 BAT arm and the belumosudil+BAT arm. The company conclude that critical differences in study inclusion criteria and populations prevented them from attempting to conduct a robust MAIC.

The EAG accept that there is limited overlap between the studies reported in Table 26 and the patients in REACH-3 are at an earlier LOT and have quite different baseline specific organ involvement. The key limiting factor for the EAG is the composition of organs affected within each patient is likely correlated with outcomes. The data reporting in REACH-3 does not allow matching by the composition of organs affected within each patient, and therefore, it may not be feasible for the company to perform a robust MAIC.

Table 26. Comparison of key prognostic baseline characteristics from REACH-3 and the pooled belumosudil studies (≥ 2 lines of prior therapy subgroup) (mITT, September 2022 data cut).

cGvHD therapeutic	BAT (N=164)	Belumosudil 200 mg once daily+BAT (N=92)	Belumosudil 200mg twice daily+BAT (N=84)
Previous aGvHD no.(%)	88 (53.7%)	[redacted]	[redacted]
cGvHD severity no.(%)	Moderate 74 (45.1%) Severe 89 (54.3%)	[redacted]	[redacted]
Median Line of prior Tx	Not reported	[redacted]	[redacted]
Number of involved organs	Not reported	[redacted]	[redacted]
Baseline organ involvement			
Skin	110 (67.1%)	[redacted]	[redacted]
Eye	93 (56.7%)	[redacted]	[redacted]
Mouth	99 (60.4%)	[redacted]	[redacted]
Oesophagus	17 (10.4%)	[redacted]	[redacted]
Upper GI tract	21 (12.8%)	[redacted]	[redacted]
Lower GI tract	10 (6.1%)	[redacted]	[redacted]
Liver	83 (50.6%)	[redacted]	[redacted]
Lung	49 (29.9%)	[redacted]	[redacted]
Joint and fascia	44 (26.8%)	[redacted]	[redacted]

Abbreviations: aGvHD, acute graft-versus-host disease; cGvHD, chronic graft-versus-host disease; GI, gastro-intestinal; Tx, treatment.

3.4.3.1.2.1 Baseline characteristics of REACH-3 and the belumosudil trials

The baseline characteristics of the combined belumosudil+BAT treatment arm and REACH-3 BAT arm are presented below in Table 27. The EAG's clinical experts noted where the baseline characteristics indicate an arm is more difficult to treat. However, baseline characteristics were not consistently reported between the belumosudil+BAT trials and REACH-3 trial.

The belumosudil+BAT arm and REACH-3 BAT arm were consistent in the median age and the sex of the participants. The patients had better HLA matching in the belumosudil+BAT arm, although this is linked to a person's chance of developing cGvHD and not than an indicator of its severity.

As noted in Section 3.2.3, time spent with cGvHD without effective treatment can be a sign of disease that is more developed and more complex to treat. Time from cGvHD diagnosis to recruitment to the trial, number of prior LOTs, and cGvHD severity can be indicators of developed and complex disease.

The belumosudil+BAT arm was recruited a median of 28.14 months after diagnosis with cGvHD, while the REACH-3 BAT arm was median of 4.92 months after diagnosis. Therefore, the patients in the belumosudil+BAT arm had spent a substantially longer period without effective treatment.

The patients in the combined belumosudil+BAT arm had ≥2 LOT and 49% of patients in the REACH-3 BAT arm had received only corticosteroid treatment and 42% had received only corticosteroid with CNIs. National guidance, and the EAG's clinical experts, consider corticosteroid +/- CNIs as the first line treatment (see Section 2.2.1).[3, 4] Therefore 91% of the patients in the REACH-3 BAT arm would be ineligible for third line treatment with belumosudil under its current marketing authorisation.

The severity of a person's cGvHD was assessed using the 2014 NIH consensus criteria and was based on the number of organs involved and the maximum severity in each organ (Table 28). In the combined belumosudil+BAT arm, 69.7% of patients had severe disease compared to 54.3% in the BAT arm of REACH-3.

However, in Section 3.2.3 the EAG's clinical experts noted that different organ systems have different responses to treatment. Lung, liver and gut are known to be more difficult to treat with lower anticipated response rates. These are often the key organs that are associated with the highest disease morbidity and are the targets for treatment. The REACH-3 BAT arm has a substantially higher proportion of patients with liver involved than the belumosudil+BAT arm (50.6% versus 8.5%, respectively) and a higher proportion with lung involved (40.9% versus 33.5%, respectively). The belumosudil+BAT arm had a higher proportion with upper gastrointestinal (21.6% versus 12.8%, respectively) and lower gastrointestinal (9.1% versus 6.1%, respectively) involvement.

The EAG's clinical experts concluded that many factors indicate that patients in REACH-3 had a better prognosis but a key factor, specific organ involvement, indicates that patients in the REACH 3 arm were more complex to treat.

Table 27. Baseline characteristics of combined belumosudil 200 mg+BAT treatment arms and REACH-3 BAT arm

	Combined belumosudil+BAT (N=176) [a]	REACH-3 BAT control arm (N=164)
Median age (range), years	55.0 (18 to 77)	50 (12 to 76)
Males, n (%)	102 (58.0%)	92 (56%)
GVHD prophylaxis after transplant, n (%) None	1 (0.6%)	Not reported

CNI only CNI + methotrexate CNI + methotrexate + other CNI + MMF CNI + MMF + other CNI + MMF + ATG CNI + sirolimus CNI + corticosteroids Other regimen	12 (6.8%) 73 (41.5%) 17 (9.7%) 25 (14.2%) 8 (4.5%) 1 (0.6%) 16 (9.1%) 3 (1.7%) 19 (10.8%)	
HLA matching of donor/recipient, n (%) Matched Partially matched Unknown Missing	 158 (89.8%) 16 (9.1%) 1 (0.6%) 1 (0.6%)	 127 (76%) 35 (21%) 5 (3%) NR
History of acute GVHD, n (%)	124 (70.5%)	88 (53.7%)
Time from chronic GVHD diagnosis to enrolment, median (range), months	28.14 (1.6 to 162.4)	4.92 (0.3 to 64)
NIH chronic GVHD severity[a] n (%) Severe Moderate Mild	 124 (70.5%) 50 (28.4%) 2 (1.1%)	 89 (54.3%) 74 (45.1%) 1 (0.6%)
Organ involvement, n (%) ≥4 organs involved ≥6 organs involved Skin Joints/fascia Eyes Mouth Lungs Oesophagus Upper GI Lower GI Liver No. of organs involved, median (range)	 93 (52.8%) 25 (14.2%) 127 (72.2%) 144 (81.8%) 108 (61.4%) 133 (75.6%) 59 (33.5%) 27 (15.3%) 38 (21.6%) 16 (9.1%) 15 (8.5%) 4.0 (0-7)	 NR NR 110 (67.1%) 44 (26.8%) 93 (56.7%) 99 (60.4%) 67 (40.9%) 17 (10.4) 21 (12.8) 10 (6.1%) 83 (50.6%) NR
Refractory to prior LOT, n (%)	103 (73.0%)	73 (44.5%)
Number or prior lines of therapy, n (%) 1 2 3 4 5	 0 47 (26.7%) 53 (30.1%) 38 (21.6%) 34 (19.3%)	Not reported

≥6 Median	4 (2.3%) 3.0	
Prior systemic chronic GVHD therapies, n (%)b	Prednisone 174 (98.8%) Tacrolimus 109 (61.9%) Sirolimus 84 (47.7%) ECP 9 (5.1%) Ibrutinib 54 (30.7%) MMF 44 (25.0%) Rituximab 39 (22.2%) Ruxolitinib 55 (31.3%)	Glucocorticoid only 81 (49.4%) Glucocorticoid+CNI 69 (42.1%) Glucocorticoid+CNI+other systemic therapy 4 (2.4%) Glucocorticoid+other systemic therapy 9 (5.5%) Missing 1 (0.6%)

[a] September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208).

Abbreviations: Chronic GvHD, cGvHD; CNI, calcineurin inhibitor; ECP, extracorporeal photopheresis; GI, gastrointestinal; graft-versus-host disease, GvHD; HLA, human leukocyte antigen; MMF, mycophenolate mofetil; NR, not reported.

Table 28. Chronic GvHD assessment by the 2014 NIH consensus criteria[12] (reproduced from Table 3, CS)

Category	Organs involved, n	Maximum severity
Mild	≤2	1 (0 for lung)
Moderate (a)	≥3	1 (0 for lung)
Moderate (b)	Any	2 (1 for lung)
Severe	Any	3 (2 for lung)

Maximum severity scale, 0: No clinical manifestations/symptoms; 1: Clinical manifestations with no more than mild disability; 2: Clinical manifestations with moderate disability; 3: Clinical manifestations with severe disability.

Abbreviations: GVHD, graft-versus-host disease; NIH, National Institutes of Health.

3.4.3.1.3 Clinical management in REACH-3, ROCKstar, and KD025-208

The EAG's clinical experts noted that that the three trials primarily used for the CS, ROCKstar, KD025-208 and REACH-3, have established clinical management that is reflective of USA care. Therefore, it is likely that established clinical management was consistent between the three trials despite it being different to UK care. The EAG consider that because established clinical management is likely to be similar across the trials that the estimate of the (relative) benefit of belumosudil over established clinical management alone is likely to be unbiased. However, the absolute is potentially confounded by inconsistencies in the eligibility criteria, baseline characteristics, and outcome reporting between the REACH-3 trial and the belumosudil trials.

The EAG are concerned that the trials are utilising established clinical management that is reflective of USA care, and this limits the generalisability of the results to UK care. For the economic model, the company reweighted the proportions of each treatment included in BAT from REACH-3 to be reflective of UK clinical practice, as presented in Table 12 in Section 2.3.3. However, despite this

reweighting, the AEG are aware that the efficacy data used in the model is based on USA established clinical management.

3.5 Critique of the indirect comparison

In this section the EAG present the results of the naïve comparison between the pooled belumosudil+BAT arms and the BAT arm in REACH-3. In the CS, the company presents a summary of endpoints across different studies and treatment arms (reproduced in Table 29). Unless specifically stated, the data presented from the belumosudil+BAT trials are from the 176 patients in the September 2022 data cut from ROCKstar trial with the ≥2 LOT subgroup in the KD025-208 trial.

Table 29. Summary of endpoints across different studies and treatment arms (reproduced from Table 39, CS)

Outcome	Pooled ROCKstar and KD025-208 – belumosudil+BAT trials	REACH-3 – BAT
OS	✓	✓
FFS	✓	✓
DOR	✓	✓
TTR*	✓	
TTD*	✓	

*While TTR and TTD were not trial endpoints they were derived from the pooled belumosudil studies for the purpose of the economic analysis. For REACH-3, only median statistics for TTR and TTD were available.

Abbreviations: BAT, best available therapy; DOR, duration of response; FFS, failure-free survival; OS, overall survival; TTD, time to treatment discontinuation; TTR, time to response.

3.5.1 Failure-free survival

In the pooled analysis of the belumosudil+BAT arms, median (95% CI) FFS was 15.2 (9.26 to 24.02) months in the once daily arm (n=92) and 16.6 (11.27 to 35.88) months in the twice daily arm (n=84). In the BAT arm (n=164) of REACH-3, median FFS was 5.7 (5.6 to 6.5) months. Therefore, FFS was found to be significantly longer in the belumosudil+BAT arms when naively compared to BAT alone in REACH-3.

3.5.2 Best overall response rate

Best ORR at any time was the response outcome used in the economic model. This is reported at 3 years in the belumosudil+BAT trials and at 24 weeks in REACH-3. The EAG notes that there were potentially patients who had their best response to treatment after 24 weeks who would be observed in the belumosudil+BAT trials and not observed in the REACH-3 BAT arm. The company

provided best ORR at 24 weeks in belumosudil+BAT trials at the clarification stage. The best ORR at 24 weeks and 3 year analysis are presented below in Table 30.

Table 30. Response data for each treatment arm

Outcome	ROCKstar and KD025-208 (≥2 prior lines of therapy subgroup)		REACH-3
	Belumosudil 200 mg QD+BAT (n=92)	Belumosudil 200 mg BID+BAT (n=84)	BAT (n=164)
Best ORR at week 24			
Number of patients with response (%)	[redacted]	[redacted]	99 (60.4%)
CR (%)	[redacted]	[redacted]	11 (6.7%)
PR (%)	[redacted]	[redacted]	88 (53.7%)
Number of patients with no response (%)	[redacted]	[redacted]	65 (39.6%)
Best ORR (3 years analysis)			
Number of patients with response (%)	67 (72.8%)	62 (73.8%)	NR
CR (%)	4 (4.3%)	2 (2.4%)	NR
PR (%)	63 (68.5%)	60 (71.4%)	NR
Number of patients with no response (%)	25 (27.2%)	22 (26.2%)	NR

Abbreviations: BAT, best available therapy; BID, twice daily; CR, complete response; NR, not reported; ORR, overall response rate; QD, once daily; PR, partial response.

3.5.3 *Response at week 24 by organ*

As noted in Section 2.3.5, response by organ at week 24 is presented in table S4 in the REACH-3 trial appendix.[10] This is response at week 24 rather than best ORR by week 24. At the clarification stage, the company provided a table comparing response at week 24 between the REACH-3 BAT arm and the combined belumosudil+BAT arms. The naïve comparison indicates an increased benefit of belumosudil+BAT over BAT for joints/fascia, eyes, upper GI, and lower GI manifestations. BAT has an increased benefit over belumosudil+BAT for liver manifestations. The company caution against evaluating response on an individual organ basis, as in Table 31, as this does not accurately reflect either the baseline manifestation or the individual patient benefit. Neither does it necessarily represent the primary organ impacted by cGvHD.

Table 31. Comparison of response by organ at week 24 from the REACH-3 BAT arm and the combined ROCKstar and KD025-208 arms

Organ	Combined belumosudil+BAT (n=176)[a]		REACH-3 BAT (n=164)	
	Baseline involvement, n	Responders at week 24, n (%)	Baseline involvement, n	Responders at week 24, n (%)
Skin	[redacted]	[redacted]	110 (67.1%)	17/110 (15.5%)
Joints/fascia	[redacted]	[redacted]	44 (26.8%)	7/44 (15.9%)
Eyes	[redacted]	[redacted]	93 (56.7%)	10/93 (10.8%)
Mouth	[redacted]	[redacted]	99 (60.4%)	25/99 (25.3%)
Lungs	[redacted]	[redacted]	49 (29.9%)	3/49 (6.1%)
Oesophagus	[redacted]	[redacted]	17 (10.4%)	5/17 (29.4%)
Upper GI	[redacted]	[redacted]	21 (12.8%)	8/21 (38.1%)
Lower GI	[redacted]	[redacted]	10 (6.1%)	3/10 (30.0%)
Liver	[redacted]	[redacted]	83 (50.6%)	18/83 (21.7%)

[a] September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)

Abbreviations: BAT, best available therapy; GI, gastrointestinal.

3.5.4 *Overall survival*

In the REACH-3 BAT arm, there were 27 (16.5%) deaths after a maximum follow-up of study of approximately 2.4 years, which equates to an OS of 83.5%. At the clarification stage, the company provided the survival rate (95% CI) of the combined belumosudil+BAT arm at two years: 85% (95% CI: 78% to 90%). The OS appears to be broadly similar between the two treatment arms with a small numerical benefit in the belumosudil+BAT arm. However, the EAG notes that these estimates are based on immature data. The company reported marginally more mature data in Section B.3.3.2 of the CS, there were a total of 30 deaths (19%) over a maximum follow-up duration of 4.7 years in the combined analysis of 156 patients in the belumoudil+BAT arm. This analysis uses the data from the August 2021 data cut.

3.5.5 *Duration of response*

The DOR outcomes used in the belumosudil+BAT trials and REACH-3 are presented in Table 32, with the definitions used in each. The company utilise the quaternary DOR in the model and the quaternary DOR is more than twice as long than the primary/secondary DOR. The principal difference between the definitions of quaternary DOR and primary/secondary DOR is the inclusion of summed durations for multiple response/lack of response episodes. This appears to be driving the increase in quaternary DOR.

The DOR used in REACH-3 does not use summed durations for multiple response/lack of response episodes. Given that the benefit in quaternary DOR is largely driven by multiple response/lack of response episodes, it is not comparable with the DOR used in REACH-3. However, the EAG also recognise that the primary/secondary DOR in the belumosudil+BAT trials does not match well with the DOR used in REACH-3. The REACH-3 DOR is taken from time of first response to cGvHD progression, whereas primary/secondary DOR time of first response to lack of response. This comparison favours the BAT arm as cGvHD progression is a category that sits under lack of response (see Section 2.3.4). The EAG is concerned that neither quaternary DOR or primary/secondary DOR are appropriate for comparison with DOR as reported in REACH-3. The EAG recommends that the company utilises the IPD it has for the belumosudil+BAT trials to match the definition of DOR used in REACH-3. Alternatively, the company could conduct scenario analyses to determine how sensitive the results of the economic model are to this outcome.

Table 32. Duration of response data in responders

Treatment arm	Median DOR in responders in weeks (95% CI)
Belumosudil+BAT combined analysis[a]	
Primary/secondary[b]	24.1 (16.14 to 36.14)
Quaternary[c]	69.9 (40.43 to 95.43)
REACH-3 BAT arm[d]	27 (NR)

[a] Combined analysis using the August 2020 data cut in ROCKstar and the ≥2 prior lines of therapy subgroup in KD025-208.

[b] DOR was measured from the time of first documentation of response to the time of first documented deterioration from best response (primary), to the time of first documented lack of response (secondary).

[c] DOR was measured from the time of first documentation of response to the time to the time of first documented lack of response with durations summed for multiple response/lack of response episodes (quaternary)

[d] DOR was measured from time from first documented response cGvHD progression, death, or the additional of new systemic therapies for cGvHD.

Abbreviations: BAT: best available therapy.

3.5.6 Time to treatment discontinuation

The median (range) TTD was 9.2 (0.5 to 64.2) months in the 200 mg QD+BAT arm and 11.8 (0.4 to 39.6) months in the belumosudil 200 mg BID+BAT arm. In the BAT arm of the REACH-3 trial, the median "exposure to therapy" was 5.5 (0.2 to 29.3) months. The TTD was substantially longer in the belumosudil+BAT arms.

3.5.7 Time to response

In the REACH-3 BAT arm, the median (range) time to first response in the responders was 4 (2 to 25) weeks. In the combined belumosudil+BAT analysis using the August 2021 data cut, the median

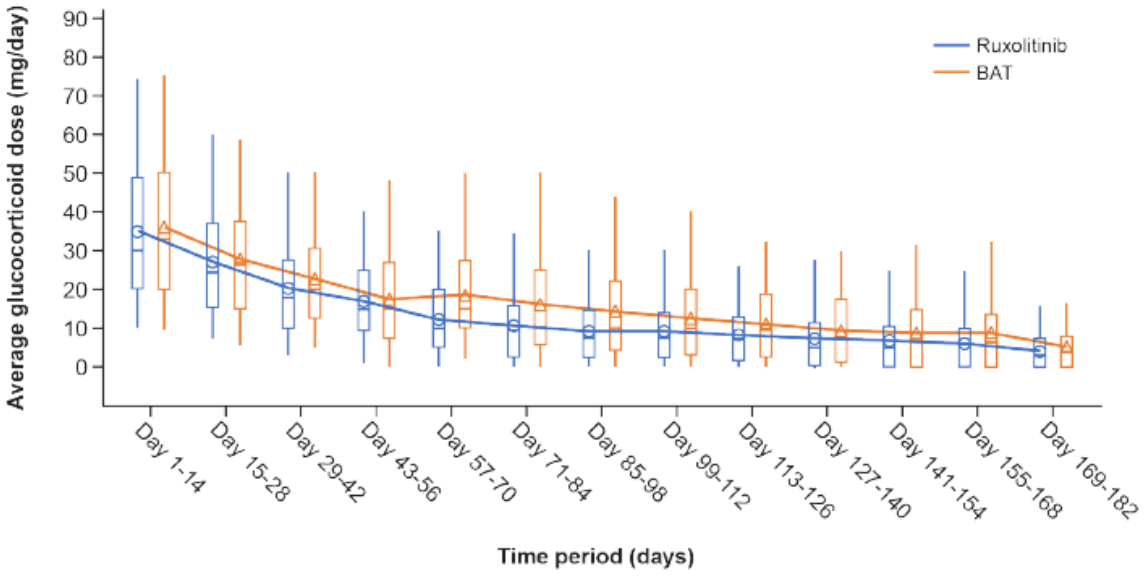
(range) TTR in responders was 7.9 weeks (3.7 to 80.1). Therefore, TTR was found to be sooner in the REACH-3 BAT arm than the belumosudil+BAT arm.

3.5.8 *Discontinuations and reductions of concomitant medications*

REACH-3 reports average corticosteroid use through the trial in a graphical form (Figure 2, below). BAT appears to show a similar steroid sparing effect to ruxolitinib. The average corticosteroid use appears to have been approximately 30 mg/day at baseline and less than 10 mg/day after 26 weeks.

The belumosudil+BAT analysis reported change in corticosteroid use during the trials in the combined analysis. The mean change in corticosteroid dose from baseline was -51.0%, with 26.3% of patients discontinuing treatment. Both BAT alone, and belumosudil+BAT, demonstrate a steroid sparing effect in the trials in this analysis. It is not possible to say which treatment regime is the most steroid sparing from the outcome data provided.

Figure 2. Corticosteroid dose over time up to week 24 (reproduced from Zeiser 2021 appendix)[10]



3.5.9 *Quality of life: modified Lee Symptom Scale;*

REACH-3 reported modified Lee Symptom Scale (LSS) responders, those who had an improvement of ≥7 points on the LSS summary score. Eighteen (11%) of patients in the BAT arm of the REACH-3 trial were LSS responders at 24 weeks. The EAG note that 58.9% of patients in the combined

belumosudil+BAT trials analysis had an improvement of ≥7 points in the LSS summary score. The EAG's clinical experts noted that in routine clinical practice, use of LSS is "aspirational" but not always achievable due to time constraints.

3.5.10 Safety

Adverse events from the belumosudil+BAT trials and the REACH-3 BAT arm are reported in Table 33 and AEs included in the model are reported in Table 34. A higher proportion of people had grade ≥3 AEs, SAEs, and fatal SAEs in the belumosudil+BAT trials. However, the proportions were not substantially different between treatment arms and the EAG notes that safety was assessed over two years in the belumosudil trials and over 24 weeks in the REACH-3 trial.

Table 33. Safety profile in the belumosudil trials and REACH-3 BAT arm

Category, n (%)	BAT (n=158)	Belumosudil+BAT combined safety population[a] (n=190)
	All grades	Grade ≥3
AEs	91 (57.6)	117 (61.6%)
Treatment related	23 (14.6)	36 (18.9%)
SAEs	53 (33.5)	83 (43.7%)
Treatment related	12 (7.6)	12 (6.3%)
Fatal SAEs	8 (5.1)	14 (7.4%)
Treatment related	4 (2.5)	NR
AEs leading to discontinuation	8 (5.1)	Reported separately by trial[b]
Treatment related	5 (3.2)	NR
AEs leading to dose adjustment/interruption	12 (7.6)	Reported in ROCKstar[c]
AEs requiring additional therapy	74 (46.8)	NR

[a] This includes the 400 mg once daily+BAT (n=21) from the KD025-208 trial and the ROCKstar September 2022 data cut.

[b] The frequency of discontinuations due to possible drug-related AEs occurred in 12% of patients in ROCKstar and 5.6% of patients in the KD025-208 (August 2021 data cut).

[c] Two-year safety results (cut-off date: 19 August 2021) for ROCKstar demonstrated that, overall, 20% of patients experienced a dose modification and 10% of patients experienced a dose interruption on account of one or more drug-related AEs.

Abbreviations: AE: adverse event; BAT: best available therapy; SAE: serious adverse event.

Table 34. Grade ≥3 AEs occurring in >5% of patients included in the economic analysis

September 2022 data cut (ROCKstar) and ≥2 LOT subgroup (KD025-208)			
Adverse event	200 mg once daily (n=92)	200 mg twice daily (n=84)	BAT (N=158)
Pneumonia	[redacted]	[redacted]	15 (9.5%)
Hypertension	[redacted]	[redacted]	11 (7.0%)

Anaemia	[redacted]	[redacted]	12 (7.6%)
Thrombocytopenia and decreased platelet counts	[redacted]	[redacted]	16 (10.1%)
Hyperglycaemia	[redacted]	[redacted]	3 (1.9%)
Gamma-glutamyl transferase increased	[redacted]	[redacted]	3 (1.9%)
Diarrhoea	[redacted]	[redacted]	2 (1.3%)

Abbreviations: AE, adverse events; BAT, best available therapy; LOT, line of therapy.

3.6 Conclusions of the clinical effectiveness section

Evidence in support of the clinical effectiveness of belumosudil 200 mg QD/BID+BAT for treating patients aged 12 years and older with cGvHD after ≥2 LOTs is derived from two trials of belumosudil+BAT, ROCKstar and KD025-208. The majority of the data used for the belumosudil+BAT arm comes from the ROCKstar trial and the eligibility criteria and interventions in ROCKstar met the decision problem under review.

Despite the eligibility criteria of ROCKstar trial matching the scope, no data in adolescents (12-18yo) is reported in the ROCKstar trial's September 2022 data cut, and the KD025-208 trial did not recruit adolescents. The company assert that the efficacy is expected to be similar in adolescents as in adults. However, all of the efficacy and safety data in the CS relates to adults and the EAG cannot confirm the same results would be seen in adolescents (12-18yo) (Section 3.2.4).

The comparator in NICE final scope is established clinical management (BAT) without belumosudil and the belumosudil+BAT trials did not have suitable comparator groups that matched the decision problem. The company investigated using a population adjusted indirect comparison (PAIC) and [redacted] to find valid data to enable a comparison of belumosudil+BAT and BAT. However, the company did not find these methods produced a robust comparison (Sections 3.4.1 and 3.4.2).

Instead the company used the BAT arm from the Phase III, REACH-3 trial of ruxolitinib vs investigator's choice (BAT) after one prior line of therapy in a naïve direct comparison with belumosudil+BAT. Given, the eligibility criteria of the REACH-3 trial and the belumosudil+BAT trials, the company concluded that this was a conservative approach. The EAG were uncertain if this was a conservative approach and, at the clarification stage, requested the company perform a MAIC.

However, the population data provided in REACH-3 was insufficiently granular to allow a robust MAIC and the data from the naïve comparison was used in the economic model (Section 3.4.3.1.2).

The EAG's experts assessed the eligibility criteria and baseline characteristics of the patients in the belumosudil+BAT trials and the REACH-3 trial to assess the direction of bias of the naïve comparison. They concluded that many factors indicate REACH-3 BAT as a more treatable arm but a key factor, specific organ involvement, indicates the REACH 3 arm were more complex to treat (Section 3.4.3.1.2.1). Thus, the EAG and its clinical experts consider it impossible to predict the overall likely direction of any bias resulting from differences in the patients recruited to each treatment arm.

The EAG's clinical experts also noted that that the three trials primarily used for this submission, ROCKstar, KD025-208 and REACH-3, have established clinical management that is reflective of USA care. The EAG consider that because established clinical management is likely to be similar across the trials that the estimate of the (relative) benefit of belumosudil over established clinical management alone is likely to be unbiased.

However, the estimate of the absolute benefit in the naïve comparison is potentially confounded by inconsistencies in the eligibility criteria, baseline characteristics, and outcome reporting between the REACH-3 trial and the belumosudil trials (Section 3.4.3.1.3). The EAG considers that with the current clinical evidence available, there is no alternative approach that can resolve the issue. Ideally, the EAG considers that the company should undertake a Phase III RCT of belumosudil+BAT versus BAT alone in the UK.

The results of the naïve comparison are reported in Section 3.5. The key driver for the economic model is FFS, median FFS was significantly longer for belumosudil+BAT arm (15.4 months, 95% CI: 12.42 to 22.74) than the REACH-3 BAT arm (5.7 months, 95% CI: 5.6 to 6.5) (Section 3.5.1). The primary outcome in the belumosudil+BAT trials was best ORR. A higher proportion of patients in the belumosudil+BAT arm (70.0%) compared to REACH-3 BAT (60.4%) responded by 24 weeks.

OS appears to be broadly similar between the two treatment arms with a small numerical benefit in the belumosudil+BAT arm. However, the EAG notes that these estimates are based on immature data. Other outcomes utilised in the economic model were DOR, TTD, and TTR. Quaternary DOR and TTD were substantially higher in the Belumosudil+BAT arm than the REACH-3 BAT arm. However, median TTR was found to be faster in the REACH-3 BAT arm than the belumosudil+BAT arm.

The safety profile reported for the belumosudil+BAT combined safety population and the REACH-3 BAT arm, favoured the BAT arm. A higher proportion of patients in the belumosudil+BAT arm (43.7%) had serious adverse events (SAEs) compared to the REACH-3 BAT arm (33.5%). There were also a higher proportion of Fatal SAEs in the belumosudil+BAT arm (7.4%) compared to the REACH-3 BAT arm (5.1%).

4 Cost effectiveness

Table 35 below presents the incremental cost-effectiveness results of the company's updated (i.e., post clarification) base case results.

Table 35. Company's updated (post clarification) base case results

Interventions	Total Costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
BAT	248,736	[redacted]	[redacted]	-	-	-	-
Belumosudil	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3,571
Probabilistic results							
BAT	250,314	[redacted]	[redacted]	-	-	-	-
Belumosudil	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3,046

Abbreviations: BAT, best available therapy; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year.

4.1 EAG comment on the company's review of cost effectiveness evidence

The company carried out three systematic literature reviews (SLRs) to identify published cost-effectiveness, health-related quality of life (HRQoL) and cost and resource use evidence for chronic graft versus host disease (cGvHD), not limited by intervention. Searches were conducted in December 2022. A summary of the External Assessment Group's (EAG's) critique of the methods implemented by the company to identify relevant evidence is presented in Table 36. Due to time constraints, the EAG was unable to replicate the company's searches and appraisal of identified abstracts.

Table 36. EAG's critique of company's systematic literature review

Systematic review step	Section of CS in which methods are reported			EAG assessment of robustness of methods
	Cost effectiveness evidence	HRQoL evidence	Resource use and costs evidence	
Search strategy	Appendix G.1	Appendix H.1	Appendix I.1	Appropriate
Inclusion/ exclusion criteria	Appendix G.1	Appendix H.2	Appendix I.1	Appropriate
Screening	Appendix G.2	Appendix H.3	Appendix I.2	Appropriate
Data extraction	Appendix G.3 and Company submission B.3.1	Appendix H.4	Appendix I.3	Approach to data extraction was appropriate, but only a summary of included HRQoL and costs studies was provided in the appendices. The EAG considers that

				detailed data extraction tables should have been provided.
Quality assessment of included studies	Appendix G.4.3	N/A – none of the identified studies were suitable for use in the company base case	N/A – none of the identified studies were used to inform the economic model	The EAG considers that even though included HRQoL and costs studies were not used in the economic model, quality assurance should have been performed.

Abbreviations: CS, company submission; EAG, external assessment group; HRQoL, health related quality of life; N/A, not applicable.

The company's cost-effectiveness SLR identified the publications of seven potentially relevant studies, presented in Table 36 of the company submission (CS). Five of the studies adopted a Markov model approach, one study was a microsimulation and one study was cost-effectiveness analysis.[19-25] Interventions assessed were primarily extracorporeal photopheresis (ECP) individually or as part of a basket of treatments and ruxolitinib. The company considered that none of the identified studies fully addressed the decision problem and instead they preferred a partitioned survival modelling approach, which makes use of the belumosudil clinical trial data to estimate the cost-effectiveness of belumosudil for treating cGvHD after ≥2 lines of systemic therapy (≥2 LOT).

The HRQoL SLR identified three studies reporting utility values in patients with cGvHD (Table 45 of the CS) but the company deemed the studies not to be relevant for the economic model as data were not reported for response or failure-free survival (FFS) outcomes. As such, the utilities used in the model are based on pooled PROMIS-GH data from ROCKstar and KD025-208 (the two key belumosudil trials) mapped to EQ-5D. Additionally, for the failure health states, targeted searches to identify published utility data related to the most recent transplants for patients in ROCKstar and KD025-208 were used. Further details of utility data used in the model is presented in Section 4.2.6.

The company's SLR for cost and resource use data identified 23 primary studies but none of which related to data for the UK and so were not relevant for the economic model. As such, the company conducted a *de novo* study of hospital episode statistics (HES) data to estimate health state costs that could be used in the model. Further details of the HES study are provided in Section 4.2.7.

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 37 summarises the EAG's appraisal of the company's economic evaluation against the requirements set out in the NICE reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.

Table 37. NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	All relevant health effects for patients with cGvHD after ≥2 systemic LOT and caregivers have been included in the economic model.
Perspective on costs	NHS and PSS	All relevant costs have been included and are based on the NHS and PSS perspective.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Cost-utility analysis has been provided by the company. Fully incremental analysis not required as there is only one relevant comparator in the analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Lifetime horizon (94 years of age).
Synthesis of evidence on health effects	Based on systematic review	The company performed an appropriate systematic review.
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.	QALYs based on PROMIS-GH data from ROCKstar to EQ-5D-3L for the failure-free health state. Utility values for failure health state, AEs and caregivers based on published EQ-5D data.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	QALYs based on PROMIS-GH data from ROCKstar mapped to EQ-5D-3L for the failure-free health state. Utility values for failure health state and AEs based on published EQ-5D data for ALL, AML, CLL and CML. Caregiver disutility values based on published data for caregivers of patients with multiple sclerosis.

Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Multiple sources for utility values in the model that are not UK specific but are generalisable to the UK cGvHD patients.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The economic evaluation matches the reference case.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Costs included in the analysis have been sourced using NHS reference costs, PSSRU, BNF, eMIT and the NHS Drug tariff.[26-30]
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Discount rate of 3.5% has been used for both costs and health effects.

Abbreviations: ALL, acute lymphoblastic leukaemia; AML, acute myelogenous leukaemia; BNF, British National Formulary; CLL, chronic lymphocytic leukaemia; CML, chronic myelogenous leukaemia; EAG, External Assessment Group; eMIT, Drugs and pharmaceutical electronic market information tool; LOT, lines of therapy; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year.

4.2.2 Modelling approach and model structure

A single *de novo* economic model was developed in Microsoft© Excel to assess the cost-effectiveness of belumosudil compared with best available therapy (BAT) as treatment for cGvHD after two lines of systemic therapy. The model structure is based on a partitioned survival analysis approach, with three main health states: failure-free, failure and death.

Within the failure-free health state, patients are stratified by treatment response status, that is whether they are in response (complete or partial) or have a lack of response. Additionally, within the failure-free health state, patients can be on or off cGvHD treatment. The EAG notes that type of response and treatment status within the failure-free health state are independent of each other. For example, a patient can have a lack of response and remain on their current cGvHD treatment.

The failure health state is subdivided by failure event type: recurrent malignancy or initiation of a new systemic cGvHD therapy. For patients whose failure event is a new systemic cGvHD therapy, they can be on or off treatment. Figure 3 presents the company model schematic and Figure 4 presents the health state occupancy in the model, as well as response over time.

Figure 3. Model structure (reproduced from Figure 14 of the company submission)

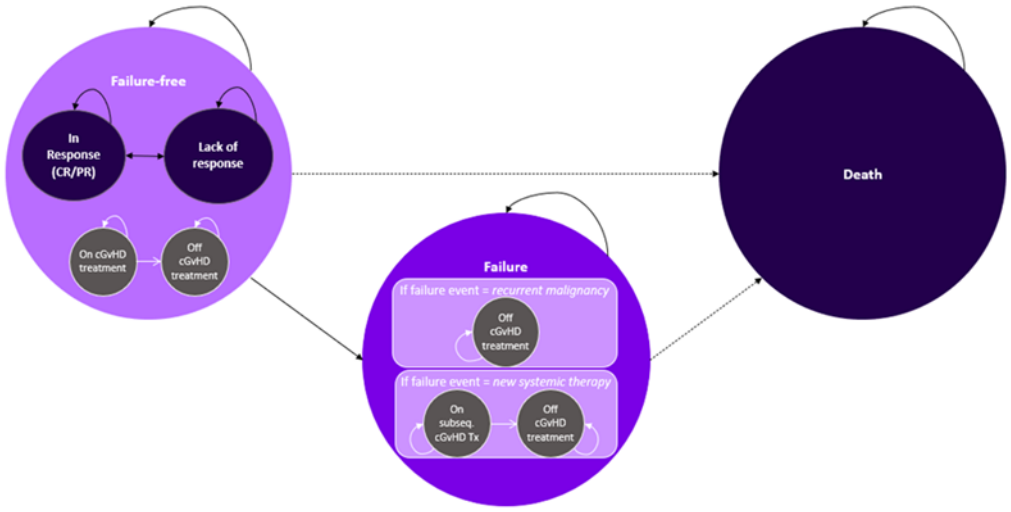
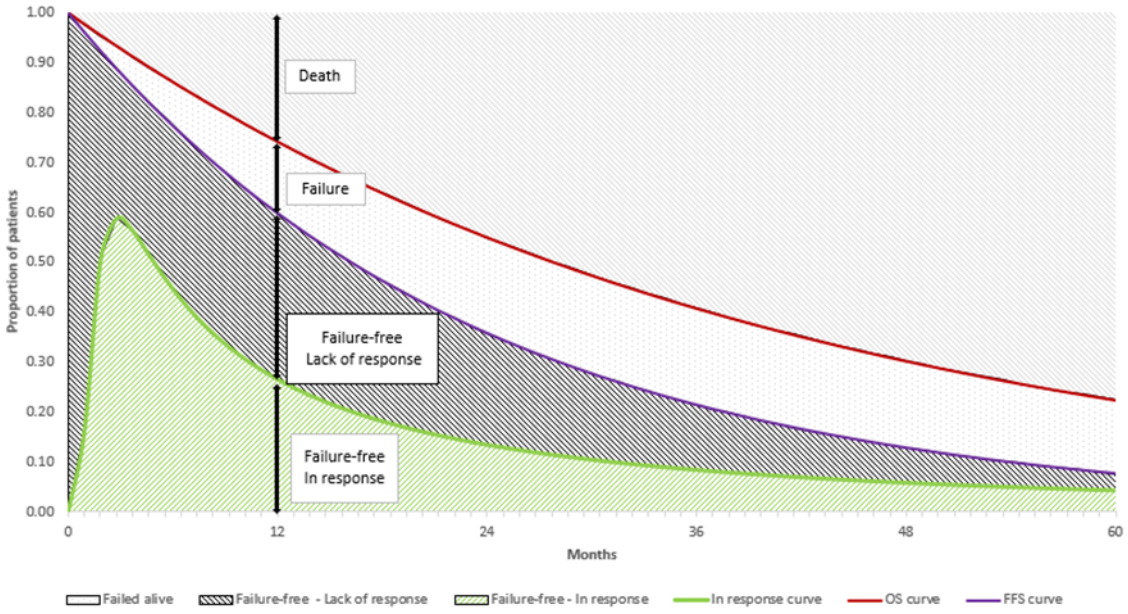


Figure 4. Partitioned survival model approach (reproduced from Figure 15 of the company submission)



All patients enter the model in the failure-free health state and start third-line systemic cGvHD treatment with either belumosudil or BAT. Over time, patients can remain in the failure-free health state or transition to the failure health state or the death state. Once in the failure health state, patients remain there until death.

The proportion of patients occupying a health state during any given cycle is based on treatment-specific parametric survival curves for the clinical outcomes of FFS (used to model the failure-free

health state), overall survival (OS) and time to treatment discontinuation (TTD) (used to estimate the proportion of patients who are failure-free and on cGvHD treatment). Within the failure-free health state, the company produced an "in response" curve based on treatment-specific patient response data, as well as time to response (TTR) and duration of response (DOR) data to estimate the proportion of patients who are failure-free and in response and failure-free with a lack of response. In the model, FFS is capped by OS and TTD is capped by FFS. Additionally, the company assumed that the maximum treatment duration in the failure-free health state is five years.

The proportion of patients occupying the failure health state for any given cycle is calculated as the difference between OS and FFS per cycle. Within the failure health state, patients are stratified by their type of failure event (recurrent malignancy or initiation of a new systemic cGvHD therapy). For patients whose failure event is initiation of a new systemic cGvHD therapy, the company assumed that patients spend 60% of the remainder of their lifetime on subsequent treatment.

Belumosudil is assumed to have an OS benefit compared with BAT, but the magnitude of the benefit is uncertain as comparative, long-term OS data are not available. As such, the company capped the belumosudil OS benefit to five years, such that after year five the risk of death for a belumosudil patient is equal to a BAT patient.

The clinical data informing the model, including response and failure event data, is based on pooled data for the subgroup of patients who have had ≥2 prior lines of therapy (hereafter referred to as ≥2 LOT subgroup) from ROCKstar and KD025-208 for belumosudil and REACH-3 for BAT. A description of how the clinical data are implemented in the model is provided in detail in Section 4.2.4.

The EAG was concerned that the definition of FFS and response in ROCKstar and KD025-208 appeared to differ from REACH-3 (see Table 38). However, in their clarification response, the company confirmed that in ROCKstar and KD025-208, mortality linked to relapse would be categorised as a recurrent malignancy failure event. Additionally, the company confirmed that any change to or introduction of a new systemic cGvHD therapy would be classed as a failure event. As such, the EAG is satisfied the definitions of FFS between ROCKstar, KD025-208 and REACH-3 are aligned. However, as discussed in Section 3.5, the definition of response between ROCKstar, KD025-208 and REACH-3 is different, and this is explored further in Section 4.2.4.

Table 38. Outcome definitions from ROCKstar, KD025-208 and REACH-3

Clinical trial	FFS definition	Response definition
ROCKstar & KD025-208 - belumosudil	The absence of cGvHD treatment change, non-relapse mortality, and recurrent malignancy.	Best response at any post-baseline assessment.
REACH-3 - BAT	Composite time to event endpoint incorporating the following FFS events: i) relapse or recurrence of underlying disease or death due to underlying disease, ii) non-relapse mortality, or iii) addition or initiation of another systemic therapy for cGvHD.	Best response up to week 24.

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; FFS, failure-free survival

4.2.2.1 EAG critique

The EAG's primary concern with the model structure was around the approach to the failure health state. The two failure events in the model (initiation of new systemic cGvHD therapy and recurrent malignancy) are representative of two clinically distinct patient groups who are likely to have different outcomes particularly with regards to OS. Additionally, for patients who initiate a new systemic cGvHD therapy, FFS on their next line of therapy is an important outcome of interest. In the current model structure, time in the failure health state is based on the difference between FFS and OS (as is standard for a partitioned survival model [PSM]), but this does not provide the granularity around survival needed for each type of failure patient within the health state. As such, the EAG considered that each failure event should be a separate health state.

The EAG explored the feasibility of modelling each failure event as a separate health state with the company during the clarification stage. The company explained modelling failure events as separate health states was not possible as data required to model failure events separately is not publicly available for the BAT arm of REACH-3. Additionally, the company advised that stratifying the failure health state to include subsequent FFS and OS related to failure events is not possible with the data available from ROCKstar and KD025-208 and that the current PSM approach makes the best use of the available data from the belumosudil trials and REACH-3.

The company highlighted that within ROCKstar, KD025-208 and REACH-3, recurrent malignancy failure events were low ([redacted] patients in the pooled analysis of ROCKstar and KD025-208 studies for the ≥2 LOT subgroup and 4.3% patients in the BAT arm of REACH-3). Thus, the majority of failure events in the model relate to initiation of a new systemic cGvHD therapy. This is reflected in the

estimated life years associated with initiation of a new systemic cGvHD therapy out of the total estimated life years for the failure health state for belumosudil and BAT ([redacted] years out of [redacted] years and [redacted] years out of [redacted] years, respectively). As such, the company considers that OS in the model implicitly captures the survival outcomes associated with initiation of new systemic cGvHD therapy.

Nonetheless, the company stated that the failure health state in the model captures important costs including subsequent lines of treatment and disease management costs associated with each type of failure event (see Sections 4.2.7.6 and 4.2.7.7). However, the EAG notes that a single utility value for the failure health state is used in the model (see Section 4.2.6.1).

The EAG considers that the company's justifications for not modelling failure events as separate health states is not unreasonable and predominantly hinges on the lack of available data for the BAT arm, which would require several strong assumptions if implemented and limited number of observed recurrent malignancy failure events in both arms of the model.

4.2.3 Perspective, time horizon and discounting

A model cycle length of four weeks (28 days) with half-cycle correction applied was implemented in the model. The time horizon was set to 40 years (lifetime), as the mean age at baseline from the pooled ROCKstar and KD025-208 studies for the ≥2 LOT subgroup was 53.9 years. The model cycle length reflects treatment cycle length in ROCKstar.[7] The perspective of the analysis was based on the UK NHS, with costs and benefits discounted using a rate of 3.5%, as per the NICE reference case.[31] The EAG considers the company's approach to be appropriate.

4.2.4 Treatment effectiveness

Clinical data in the economic model for belumosudil and BAT included response and failure event outcomes as well as FFS, OS, TTD, TTR, DOR and adverse events (AEs). Please refer to Section 3.3.6 and Section 4.2.7.3 for further details of AEs and TTD. No comparative randomised control trials (RCTs) were conducted to compare belumosudil and BAT, thus the key clinical data informing the model for the treatment comparison is derived from a naïve comparison of multiple trials, including ROCKstar, KD025-208 and REACH-3.

ROCKstar is an ongoing Phase II randomised, open-label, multicentre study of belumosudil 200 mg once daily (QD) and belumosudil 200 mg twice daily (BID) in patients with cGvHD who had previously

been treated with at least 2 prior lines of systemic therapy. The company stated that ROCKstar was completed for the adult cohort, but that the study was continuing to recruit adolescent patients.

KD025-208 was a Phase IIa, open-label, dose-escalation, multicentre study of belumosudil 200 mg QD, belumosudil 200 mg BID and belumosudil 400 mg QD in patients with cGvHD. KD025-208 completed in May 2022.

REACH-3 was a Phase III randomised, open-label, multi-centre trial of ruxolitinib versus BAT in patients with corticosteroid-refractory cGvHD after allogeneic stem cell transplantation. In REACH-3, participants who had received 2 or more systemic treatments for cGvHD in addition to corticosteroids ± calcineurin inhibitors (CNIs) for cGvHD were excluded from the study. REACH-3 completed in December 2022.

Clinical data informing the belumosudil arm of the model are derived from a pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, which includes the latest September 2022 data cut from ROCKstar. In the original company submission, the data cut for the pooled analysis informing the model was from August 2021. The Summary of Product Characteristics (SmPC) for belumosudil does not recommend the 400 mg QD dosing regimen and thus is excluded from the company's pooled analysis.[1] However, the SmPC for belumosudil specifies that the dose should be increased to 200 mg BID when co-administered with strong CYP3A inducers or proton pump inhibitors (PPIs).[1] As such, the company's the pooled analysis for belumosudil informing the model is stratified by trial arm (belumosudil 200 mg QD or BID) and the final analysis of total costs and quality-adjusted life-years (QALYs) is a weighted analysis based on the company's assumption of that 5% of patients would be on CYP3A inducers or PPIs.

Clinical data informing the BAT arm of the model is based on the BAT arm from REACH-3. In their submission, the company investigated the feasibility of performing indirect treatment comparisons of belumosudil and BAT, but deemed these were not possible and instead based their clinical data analysis in the model on naïve comparisons of the data (see Section 3.5 for further details).

Table 39 summarises the clinical data used to estimate health-state transitions included in the model, with further detail presented in Sections 4.2.4.2 to 4.2.4.4.

Table 39. Overview of clinical data informing the health state transitions in the model

Health state transition	Clinical data informing the transitions
Failure-free to failure-free	Extrapolated FFS data from pooled ROCKstar and KD025-208 studies (≥2 LOT subgroup) for the belumosudil arms of the model and extrapolated FFS data from REACH-3 for the BAT arm of the model.
Failure-free on treatment	Extrapolated TTD from pooled ROCKstar and KD025-208 studies (≥2 LOT subgroup) for the belumosudil arms of the model. For the BAT arm, estimated HR based on median TTD from REACH-3 applied to belumosudil QD TTD extrapolation. Maximum treatment duration for all arms of the model capped to 5 years.
Failure-free and in response	KM TTR data and extrapolated DOR from pooled ROCKstar and KD025-208 studies (≥2 LOT subgroup) for the belumosudil arms of the model are used in combination to estimate an 'in response' curve. For the BAT arm of the model, only DOR KM data and median TTR were available from REACH-3. Thus, the company extrapolated median TTR using an exponential distribution and combined with extrapolated DOR from REACH-3 to estimate an 'in response' curve for the BAT arm of the model.
Failure-free to failure	The difference between extrapolated FFS and OS for each arm of the model.
Failure to failure	
Failure-free to death	Extrapolated OS data from pooled ROCKstar and KD025-208 studies (≥2 LOT subgroup) for the belumosudil arms of the model and from REACH-3 for the BAT arm of the model.
Failure to death	

Abbreviations: BAT, best available therapy; DOR, duration of response; FFS, failure-free survival; HR, hazard ratio; KM, Kaplan-Meier; LOT, lines of therapy; OS, overall survival; QD, once daily; TTD, time to treatment discontinuation; TTR, time to response.

4.2.4.1 Overview of the company's approach to survival analysis

As clinical data for both arms of the model are not fully mature, extrapolation of the data was necessary. Where Kaplan-Meier (KM) data were available for clinical outcomes, the company conducted parametric survival analysis.

Individual patient-level data (IPD) for all clinical outcomes included in the model were available for the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup. For the BAT arm of REACH-3, published KM curves were available for FFS, OS and DOR, which allowed the company to reconstruct IPD for use in the survival analysis. The company followed the guidelines outlined in the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14 to extrapolate the KM data and select appropriate distributions for the base case.[32]

In the submission, the company stated that a preliminary assessment of curve fit to observed data was made based on diagnostic plots associated with distributions. The company then tested

whether the assumption of proportional hazards (PH) held for FFS, OS, DOR, TTR and TTD (ROCKstar and KD025-208 only) outcomes by producing log-cumulative hazard and Schoenfeld residual plots. The company explained that tests for PHs were conducted for arms within the trials, for example belumosudil 200 mg QD versus belumosudil 200 mg BID and ruxolitinib versus BAT. The company used the outcomes of the PH assessment to decide to either jointly or independently fit survival distributions. The EAG notes that in the model, jointly fitted survival models are only relevant for the belumosudil 200 mg QD and belumosudil 200 mg BID arms, which is not inappropriate as the KM data for clinical outcomes for the two regimens are not substantially different from one another. Survival models for belumosudil and BAT are independent of one another.

Extrapolations of the KM data were then explored using standard parametric survival distributions (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma and generalised gamma). The company assessed the fit of each modelled curve against the observed KM data using statistical goodness of fit statistics, including Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics, visual inspection of the curves and clinical plausibility of the extrapolation over the time horizon of the economic model.

As outlined in Table 39, only median data were available for the outcomes of TTD and TTR from REACH-3 and thus could not be used in the company's survival analysis. Instead, for TTR the company extrapolated median TTR for BAT using an exponential distribution and for TTD, calibrated a hazard ratio (HR) based on median TTD for BAT and applied this to the belumosudil QD TTD extrapolation.

A detailed description of how clinical outcomes are modelled (except TTD) is provided in Sections 4.2.4.2 to 4.2.4.4. For details on TTD, please refer to Section 4.2.7.3.

4.2.4.2 Failure-free survival

In the original company submission, a data cut from August 2021 for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) for belumosudil was used to inform the model. However, the company indicated that a later data cut from September 2022 was available and in their clarification response the company updated the model with these data. Between the August 2021 and September 2022 data cuts, the company observed an improvement in median FFS for the combined belumosudil doses from [redacted] to [redacted]. Median FFS for BAT from REACH-3 was 5.7 months.

Based on the PH assessments for the pooled ROCKstar and KD025-208 studies and REACH-3, the company considered that PH held for each analysis and so decided to jointly fit survival distributions, where treatment arm is a predictor. However, as mentioned previously, jointly fitted survival distributions are only relevant for the belumosudil arms of the model. Based on statistical fit (provided in the company's clarification response to question A1) and clinical plausibility, the company selected generalised gamma distribution for BAT, belumosudil 200 mg QD and belumosudil 200 mg BID.

Figure 5 presents modelled FFS for belumosudil and BAT and Table 40 presents a comparison of observed and modelled FFS over time. The EAG notes that FFS is capped by OS and as can be seen in Figure 5, the cap applies after approximately 15 years. Based on the extrapolations, mean undiscounted FFS for belumosudil and BAT is [redacted] years and [redacted] years, respectively.

Figure 5. Modelled failure-free survival – joint-fit generalised gamma distribution

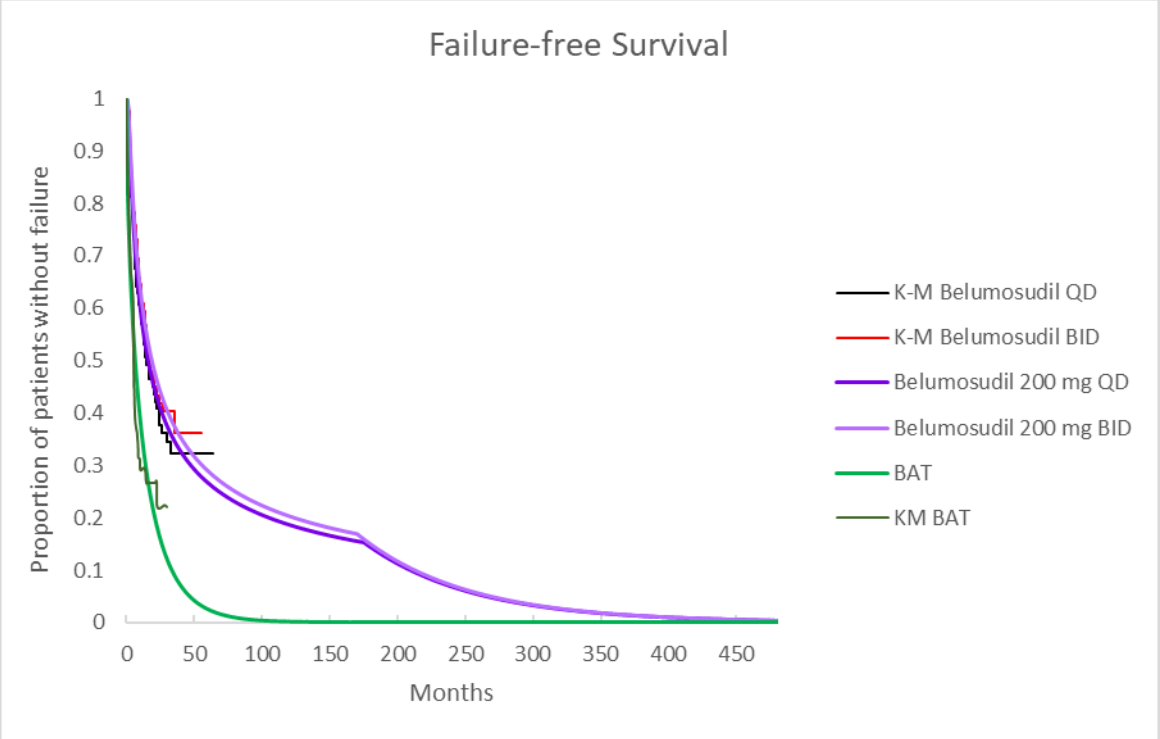


Table 40. Observed versus modelled FFS

Timepoint	Observed FFS			Modelled FFS		
	Belumosudil 200 mg QD	Belumosudil 200 mg BID	BAT*	Belumosudil 200 mg QD	Belumosudil 200 mg BID	BAT

1 year	55.7%	60.9%	30%	56.1%	59.8%	34.1%
2 years	40.8%	41.9%	20%	40.9%	44.0%	15.9%
5 years	32.4%	-	-	26.5%	28.7%	2.4%
10 years	-	-	-	18.6%	20.2%	0.1%
20 years	-	-	-	6.8%	7.0%	0.0%
30 years	-	-	-	1.5%	1.6%	0.0%

Abbreviations: BAT, best available therapy; BID, twice daily; FFS, failure-free survival; QD, once daily.
*Estimates are approximate based on published KM curves in Zeiser *et al*.[10]

4.2.4.3 Treatment response

The company included response outcomes within the failure-free health state to capture the impact on QALYs and costs for patients who achieve complete or partial response or have a lack of response. Table 41 presents the definitions of response used in ROCKstar and REACH-3. The EAG notes that in ROCKstar, KD025-208 and REACH-3, the definition of overall response rate (ORR) was based on the 2014 NIH Consensus Criteria (see Section 3.5 for further details).

Table 41. Definitions of response in ROCKstar and REACH-3.

Response	ROCKstar – belumosudil[7]	REACH-3 – BAT[10]
Complete response	Resolution of all manifestations of cGvHD in each organ or site.	Complete resolution of all signs and symptoms of cGvHD in all evaluable organs without additional therapies.
Partial response	Improvement in at least 1 organ or site without progression in any other organ or site.	An improvement in at least one organ (e.g., improvement of at least one point on a 4- to 7-point scale, or an improvement of at least two points on a 10- to 12-point scale) without progression in other organs or sites or addition/initiation of new systemic treatment.
Lack of response - mixed	Complete or partial response in at least one organ accompanied by progression in another organ.	Complete or partial response in at least one organ accompanied by progression in another organ.
Lack of response - unchanged	Outcomes that do not meet the criteria for complete response, partial response, progression or mixed response.	Stable disease or absence of improvement in any organ involved by cGvHD.

Lack of response - progression	Progression in at least one organ or site without a response in any other organ or site.	Worsening of at least one organ and no improvement (CR or PR) in any other organ.

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; CR, complete response; PR, partial response.

In the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) for belumosudil, best response at any post-baseline assessment was included in the model. However, for the BAT arm, in REACH-3 response was assessed as best response up to Week 24. Table 42 summarises the response data included in the model.

Table 42. Response data included in the economic model

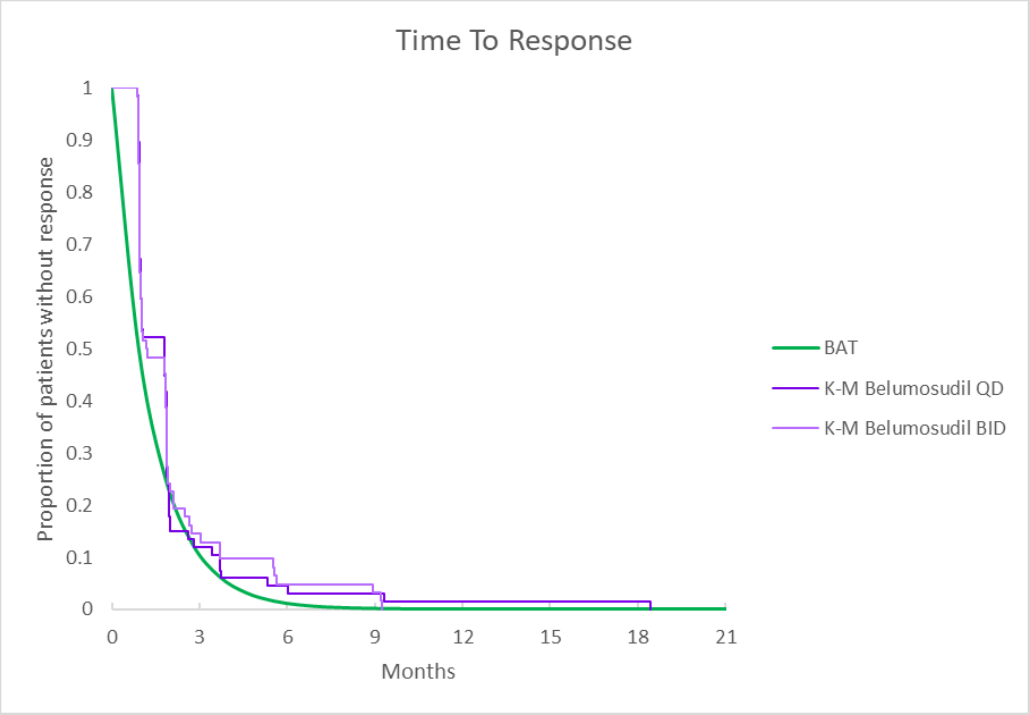
Response	Belumosudil 200 mg QD (n=92)	Belumosudil 200 mg BID (n=84)	BAT (n=164)
Overall - n (%)	67 (72.8%)	62 (73.8%)	99 (60.4%)
Complete - n (%)	4 (6.0%)	2 (3.2%)	11 (11.1%)
Partial - n (%)	63 (94.0%)	60 (96.8%)	88 (88.9%)
Lack of response - n (%)	25 (27.2%)	22 (26.2%)	65 (39.6%)

Abbreviations: BAT, best available therapy; BID, twice daily; QD, once daily

Response data in the model is constant, but the company estimated an 'in response' curve which captures the time to response and the duration of response. Based on data from the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, TTR and DOR KM curves were estimated. From REACH-3, a KM curve for BAT DOR was available but only median TTR was published.[66] Please refer to Section 3.3.1.1 for the definitions of DOR and TTR from the pooled analysis of ROCKstar and KD025-208 and REACH-3.

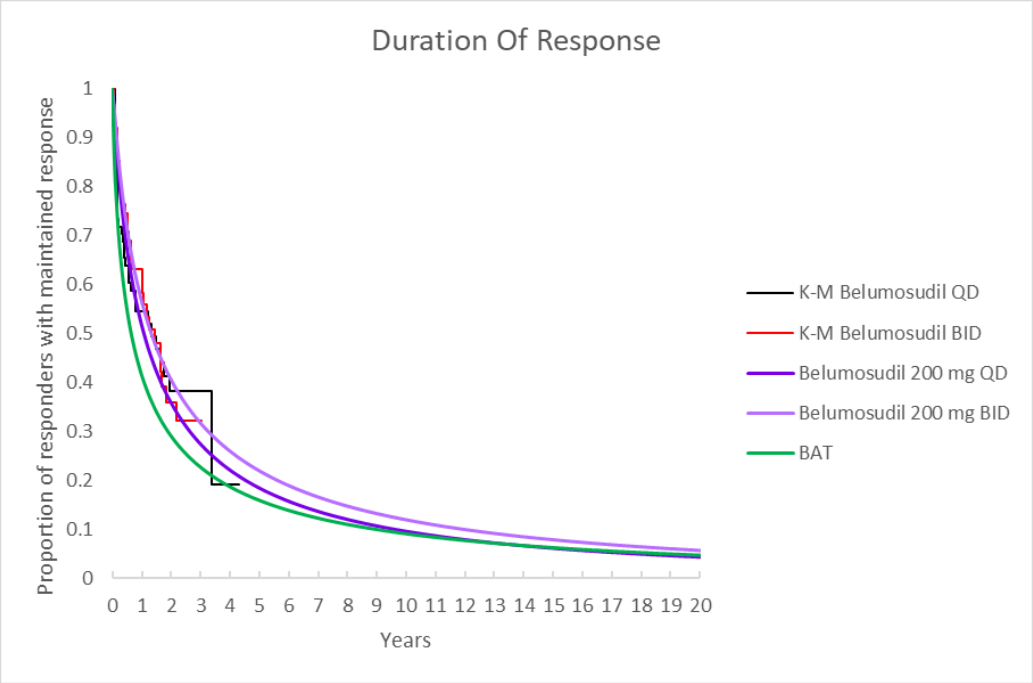
Kaplan-Meier TTR data for belumosudil 200 mg QD and BID were complete, thus the company used this data directly in the model. For the BAT arm of the model, the company estimated a curve using an exponential distribution based on the median TTR (4 weeks) for BAT from REACH-3. Figure 6 presents the TTR curves included in the economic model.

Figure 6. Modelled time to response for belumosudil and BAT



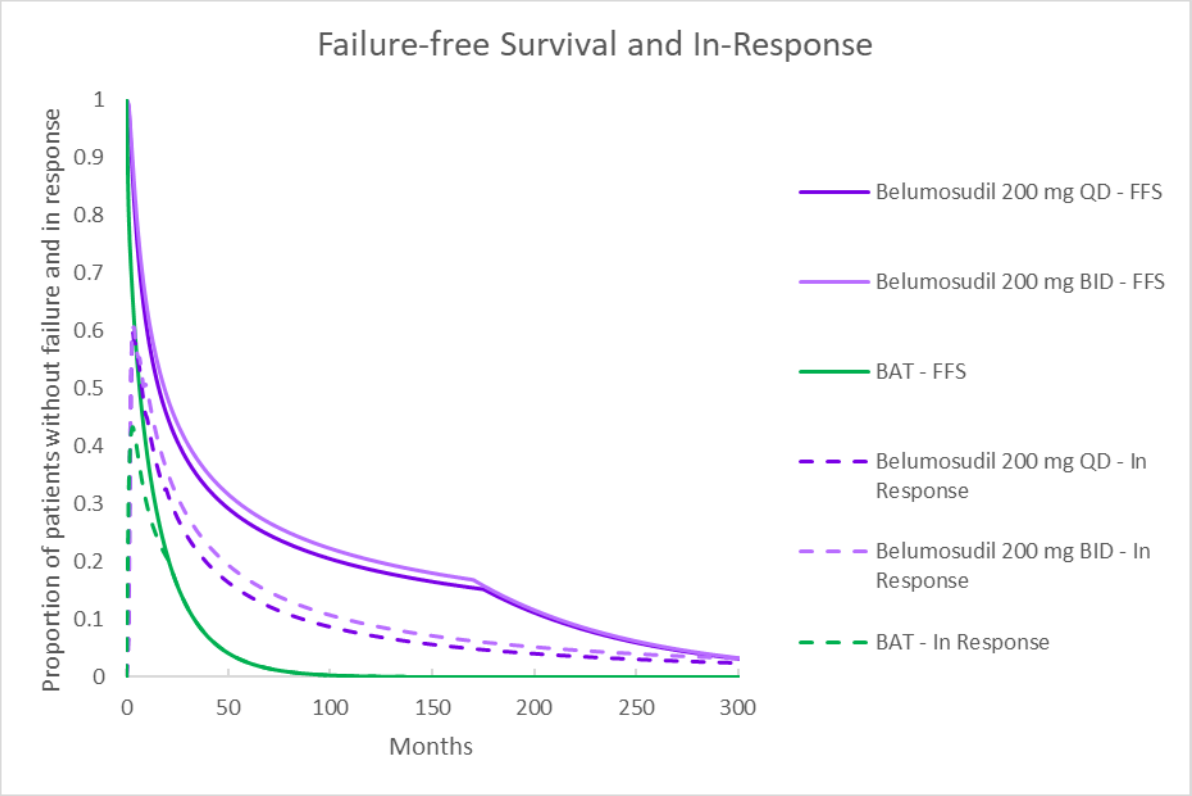
For DOR, based on the PH assessments for the pooled ROCKstar and KD025-208 studies and REACH-3, the company considered that there wasn't enough evidence to reject the PH assumption (see Appendix N of the company submission for PH tests). Therefore, the company decided to jointly fit survival distributions, where treatment arm is a predictor. However, as mentioned previously, jointly fitted survival distributions are only relevant for the belumosudil arms of the model. Based on statistical fit (not provided in the company's clarification response) and clinical plausibility, the company selected the lognormal distribution for BAT, belumosudil 200 mg QD and belumosudil 200 mg BID. Figure 7 presents the DOR curves included in the economic model.

Figure 7. Modelled duration of response for belumosudil and BAT



The company combined the TTR and DOR curves to estimate an 'in response' curve for each arm in the model for patients who have achieved a complete or partial response. In the model, the 'in response curve is capped by FFS. Figure 8 presents the 'in response' curve alongside FFS for each treatment arm in the model. Based on the extrapolations, mean undiscounted 'in response' time for belumosudil and BAT is [redacted] years and [redacted] years, respectively.

Figure 8. Modelled 'in response' and failure-free survival curves (note, OS cap of FFS occurs for belumosudil at ~15 years)



4.2.4.4 Overall survival

The pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup demonstrated that median OS had not been reached for belumosudil 200 mg QD and BID. Based on the August 2021 data cut, 30 deaths were observed in the pooled analysis for belumosudil, and this increased to [redacted] deaths based on the September 2022 data cut. As such, only minor changes in the OS KM estimates were observed as a result of the latest data cut. At one year, OS for belumosudil 200 mg QD and BID was 90.9% and 91.4%, respectively, based on the September 2022 data cut. In REACH-3, median OS was not reached in either arm of the trial and the one-year estimate of OS for BAT was 83.8%.[10] The EAG notes that the HR for ruxolitinib versus BAT was 1.09 (95% CI, 0.65 to 1.82).[10]

Based on the PH assessments for the pooled ROCKstar and KD025-208 studies and REACH-3, the company considered that PH held for each analysis and so decided to jointly fit survival distributions, where treatment arm is a predictor. However, as mentioned previously, jointly fitted survival distributions are only relevant for the belumosudil arms of the model. Based on statistical fit (not provided in the company's clarification response) and clinical plausibility, the company selected

exponential distribution for BAT, belumosudil 200 mg QD and belumosudil 200 mg BID. Figure 9 presents modelled OS for belumosudil and BAT and Table 43 presents a comparison of observed and modelled OS over time. Based on the extrapolations, mean undiscounted OS for belumosudil and BAT is [redacted] years and [redacted] years, respectively.

Figure 9. Modelled overall survival – joint-fit exponential distribution

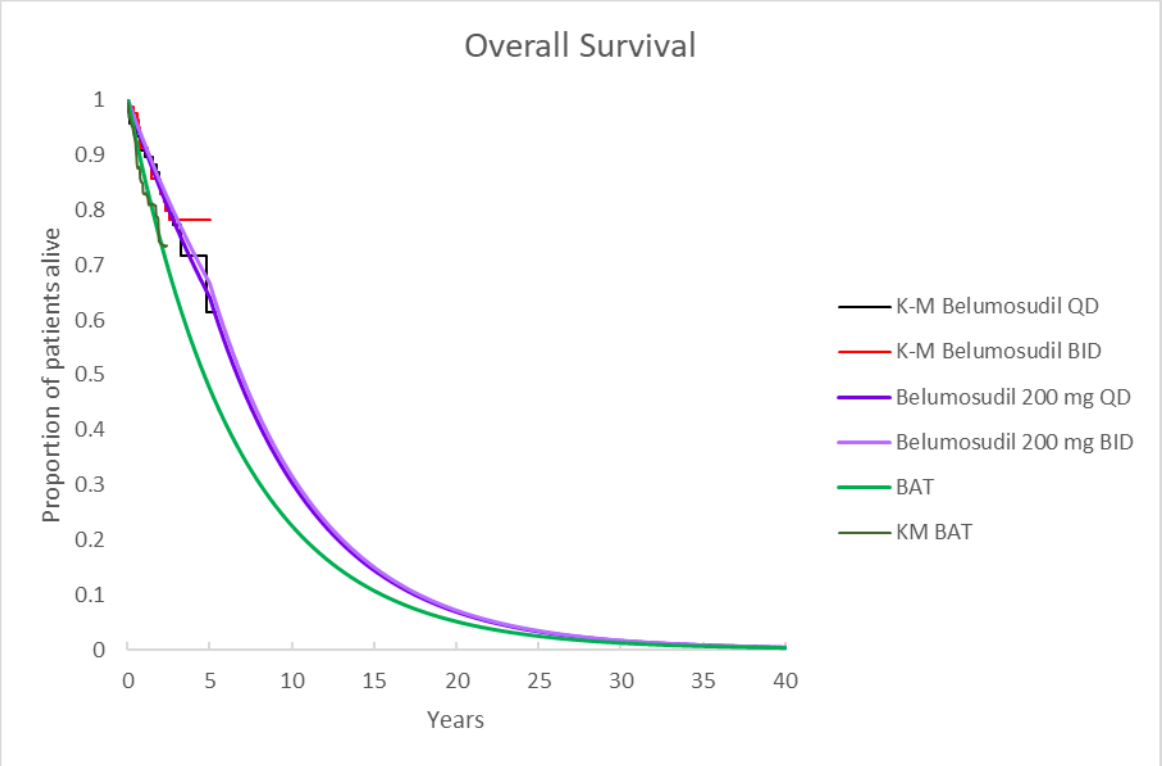


Table 43. Observed versus modelled OS

Timepoint	Observed OS			Modelled OS		
	Belumosudil 200 mg QD	Belumosudil 200 mg BID	BAT*	Belumosudil 200 mg QD	Belumosudil 200 mg BID	BAT
1 year	90.9%	91.4%	83.8%	91.5%	92.2%	86.2%
2 years	85.6%	84.3%	74%	83.1%	84.5%	73.4%
5 years	61.5%	78.2%	-	63.4%	65.9%	46.9%
10 years	-	-	-	30.1%	31.3%	22.3%
20 years	-	-	-	6.8%	7.0%	5.0%
30 years	-	-	-	1.5%	1.6%	1.1%

Abbreviations: BAT, best available therapy; BID, twice daily; OS, overall survival; QD, once daily.
*Estimate for 2 years is approximate based on published KM curves in Zeiser *et al*.[10]

The company acknowledged that long-term OS for belumosudil is uncertain due to the immaturity of the data from the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup. Additionally, OS data for BAT from REACH-3 is immature. In the model, the company implemented an adjustment to the OS curves for belumosudil such that after five years, the risk of death for patients is the same as the risk of death for BAT patients. The EAG notes that based on the updated data cut for the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, which has five years' worth of follow-up, the assumption is implemented after the observed period of OS data.

In the economic model, OS is capped by general population mortality estimates, based on the latest Office for National Statistics (ONS) life tables.[33] The EAG notes that background mortality in the model comes into effect at age 90 years for the BAT arm, age 86 years for the belumosudil 200 mg QD arm and age 85 years for the belumosudil 200 mg BID arm.

4.2.4.5 EAG critique

The updated data cut of the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup only affected the belumosudil arms of the model. Thus, combined with a naïve comparison with BAT (which remained unchanged from the company's original submission), the improvement in FFS with minimal change in OS for belumosudil, in addition with changes to TTD (discussed in Section 4.2.7.3) had a considerable impact on the cost-effectiveness results, with the overall probabilistic incremental cost-effectiveness ratio (ICER) reducing from £15,032 to £3,046.

As discussed in Section 3.5, the naïve comparison of belumosudil versus BAT is a substantial source of uncertainty in the cost-effectiveness analysis. As there is no head-to-head comparative trial, an assessment of belumosudil and BAT can only be conducted with an indirect treatment comparison. Furthermore, the company's submission relies on clinical expert opinion of the differences between ROCKstar, KD025-208 and REACH-3 and the resulting direction of bias with regards to the treatment effect.

Therefore, the EAG considers that there is a substantial amount of uncertainty related to treatment effectiveness in the model as a result of the naïve comparison of belumosudil and BAT and can be considered the primary issue for this single technology appraisal (STA). Nonetheless, the remainder

of this section covers specific issues with the company's approach to inclusion of clinical data in the model.

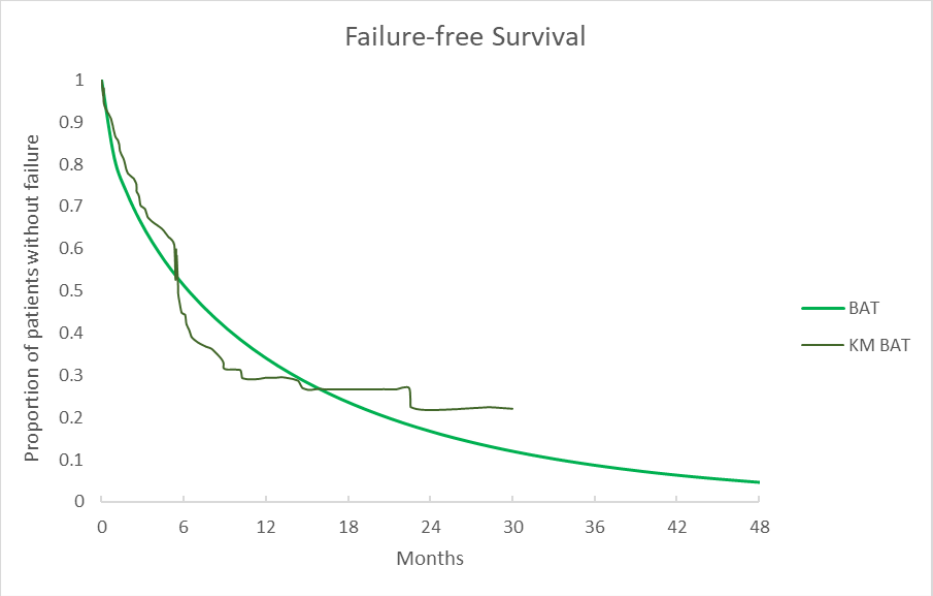
Failure-free survival

Failure-free survival is a key driver of cost-effectiveness in the model and a substantial benefit has been estimated with belumosudil over BAT. As such, the EAG investigated the appropriateness of the company's extrapolation of FFS in the model. During the clarification stage, the EAG requested the company provide the underlying cumulative hazard plots for FFS and OS for belumosudil and BAT (clarification B2), which the company provided in their clarification response (Figure 6 of the company's clarification response).

The EAG considers that for belumosudil 200 mg QD and BID, the cumulative hazards for the company's base case choice of extrapolation for belumosudil FFS and OS closely reflects that of the underlying cumulative hazards for the observed data in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup. However, for BAT from REACH-3 there was a change in the observed cumulative hazards due to the Week 24 assessment point, which the company's chosen extrapolation does not capture (see Figure 10 and Figure 11 below) and is prominent for FFS.

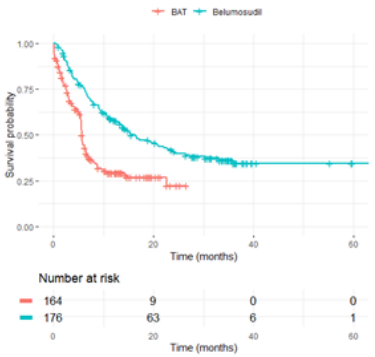
Figure 10. Cumulative hazards plot for BAT from REACH-3 (reproduced from Figure 6 of the company's clarification response)

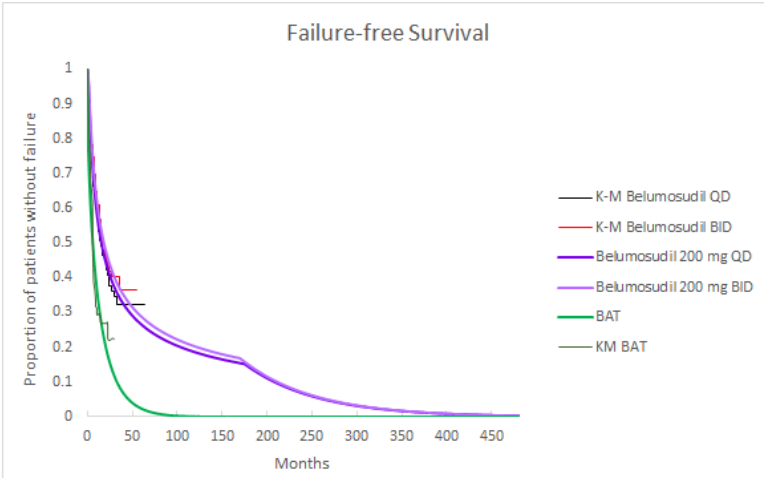
Figure 11. Modelled failure-free survival – BAT



The EAG considers that the change in the hazards in the BAT KM data is simply an artefact of when outcomes were recorded in REACH-3 and is unlikely to be seen in clinically practice. However, the generalised gamma maybe a conservative choice to model BAT, as based on the naïve comparison of KM curves for FFS the company provided in their clarification response (Figure 12 below), the EAG considers that after the Week 24 assessment point, the BAT KM curve plateaus and appears to begin to converge with the belumosudil KM curve. However, the trend towards convergence in FFS between belumosudil and BAT isn't seen in the extrapolation of FFS in the economic model (Figure 12).

Figure 12. Comparison of belumosudil and BAT KM curves and modelled extrapolation





The EAG investigated the other standard parametric distributions included in the company's model and none provided a satisfactory fit to the observed data or produced clinically plausible estimates of FFS. However, the EAG's clinical experts considered that mean FFS for patients on BAT is unlikely to be more than one year. In the model, the company estimated mean FFS to be approximately 1 year. As such, the EAG considers that the company's FFS extrapolation for BAT is the least worst option of the extrapolations explored and is not unreasonable to include in the base case analysis.

Response outcomes

Response in the model is not a primary driver of cost-effectiveness. Changes to the 'in response' curves and response data have limited impact on the ICER. However, the EAG considers that inclusion of response in the model to add granularity to QALYs and costs in the model is potentially adding unnecessary complexity to the analysis. Additionally, given the naïve comparison of belumosudil and BAT, inclusion of response in the model is another source of uncertainty in the model.

The EAG consulted with its clinical experts who advised that in clinical practice, while response to treatment is monitored as it affects how treatments will be delivered or adjusted (thus affecting costs), failure-free survival is a more clinically relevant outcome. The EAG's clinical experts explained that failure events can indicate progression of cGvHD as a change in treatment is required but that treatment for a recurrent malignancy means patients are no longer immunosuppressed and essentially resolves the cGvHD and thus cGvHD treatment would stop. However, recurrent malignancy is a far worse outcome for patients than a change in cGvHD treatment. Additionally, in their submission (page 78) the company states that FFS correlates with overall clinical improvement in cGvHD and incorporates disease control by preventing or delaying the need for cGvHD treatment changes and absence of underlying malignancy.

Therefore, the EAG considers that the company's scenario where response is excluded from the model is a more appropriate approach to the cost-effectiveness analysis and removes a source of uncertainty in the analysis, thus limiting decision risk. The company's scenario removing response has limited impact on the ICER, reducing it from £3,571 to £3,434.

As a secondary issue, the definition of response included in the model is inconsistent between the belumosudil arms and BAT. In the company's base case, the definition of response for belumosudil was best response at any post-baseline assessment. However, for the BAT arm, in REACH-3 response

was assessed as best response up to Week 24. The EAG considers that if the company wanted to include response in the model, then definition of response used in the model should match that of REACH-3. As such, during the clarification stage the EAG requested, and the company provided best response up to Week 24 for the belumosudil arms of the model (clarification question A11).

Table 44 presents the best response up to Week 24 data for all arms of the model. Compared to the company's base case response data for belumosudil, best ORR up to Week 24 is slightly lower (72.8% and 73.8% versus 69.6% and 70.2% for belumosudil 200 mg QD and BID, respectively). The company provided a scenario based on best response up to Week 24 for all treatment arms, which had minimal impact on the ICER, and results are presented in Section 5.2.2.

Table 44. Best response up to Week 24 data included – pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and REACH-3

Response	Belumosudil 200 mg QD (n=92)	Belumosudil 200 mg BID (n=84)	BAT (n=164)
Overall – n (%)	64 (69.6%)	59 (70.2%)	99 (60.4%)
Complete – n (%)	2 (3.1%)	1 (1.7%)	11 (11.1%)
Partial – n (%)	62 (96.9%)	58 (98.3%)	88 (88.9%)
Lack of response – n (%)	28 (30.4%)	25 (29.8%)	65 (39.6%)

Abbreviations: BAT, best available therapy; BID, twice daily; LOT, lines of therapy; QD, once daily

Overall survival

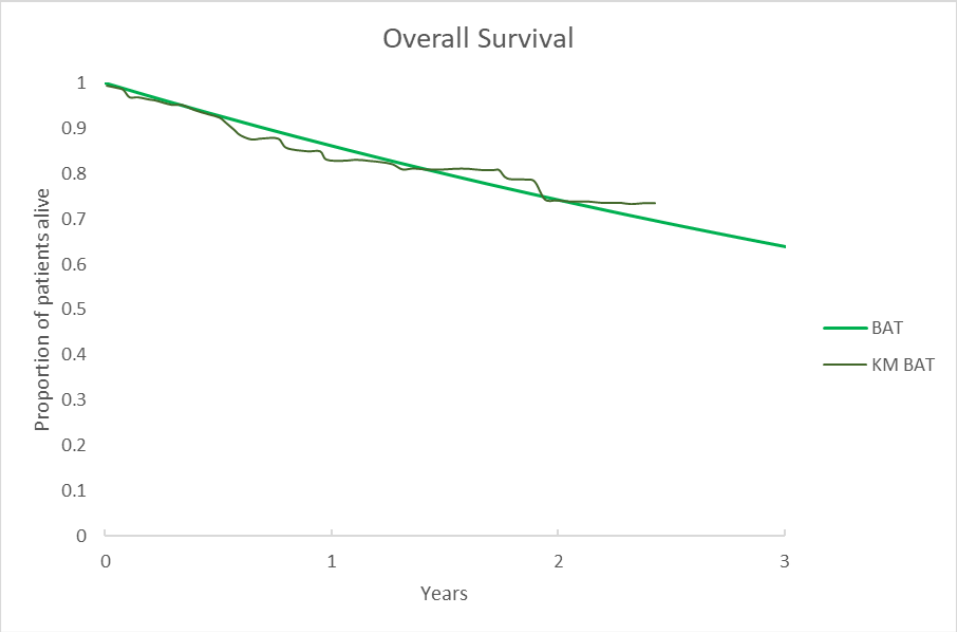
Observed OS for both belumosudil from the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and BAT from REACH-3 is immature, with neither dataset reaching median. As such, the EAG considers there is substantial uncertainty in the estimated OS benefit associated with belumosudil.

The uncertainty in OS due to immature data is further exacerbated for BAT as in REACH-3, patients in the BAT arm could receive ruxolitinib after the Week 24 assessment point if they did not have or maintain a complete or partial response, had unacceptable side effects from a control therapy, or had a flare of cGvHD.[10] Additionally, for BAT patients who did have a complete or partial response, they could only receive ruxolitinib if they had disease progression, mixed response, or unacceptable side effects from the control therapy.[10] In total, 61 out of 164 BAT patients (37.2%) in REACH-3 crossed over to ruxolitinib, which potentially confounds the OS data for BAT. The EAG notes that the OS HR for ruxolitinib versus BAT was 1.09 (95% confidence interval [CI]: 0.65 to 1.82). The EAG

considers that the OS data for BAT are overestimated from REACH-3, which potentially introduces bias in the comparison with belumosudil; however the EAG notes that this exacerbates the existing uncertainty in any naïve indirect treatment comparison.

Due to the Week 24 assessment in REACH-3, there is a there is a change in the observed cumulative hazards (see Figure 10). However, based on Figure 13, the EAG considers the company's choice of exponential extrapolation for BAT does not seem unreasonable.

Figure 13. Modelled overall survival compared with KM curve for BAT (REACH-3)



As with FFS, the cumulative hazards for the company's base case choice of extrapolation for belumosudil OS closely reflects that of the underlying cumulative hazards for the observed data in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (Figure 6 of the company's clarification response). Therefore, the EAG considers the company's choice of exponential distribution for the belumosudil treatment arms to be reasonable.

The company acknowledged that there is substantial uncertainty around long-term OS for the belumosudil arms due to the immaturity of the data. Thus, to limit the uncertainty, the company implemented an assumption where the risk of death for belumosudil patients is equal to the risk of death for BAT patients after five years (which is after the observed period of data for the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup). In the model, this results in a gradual

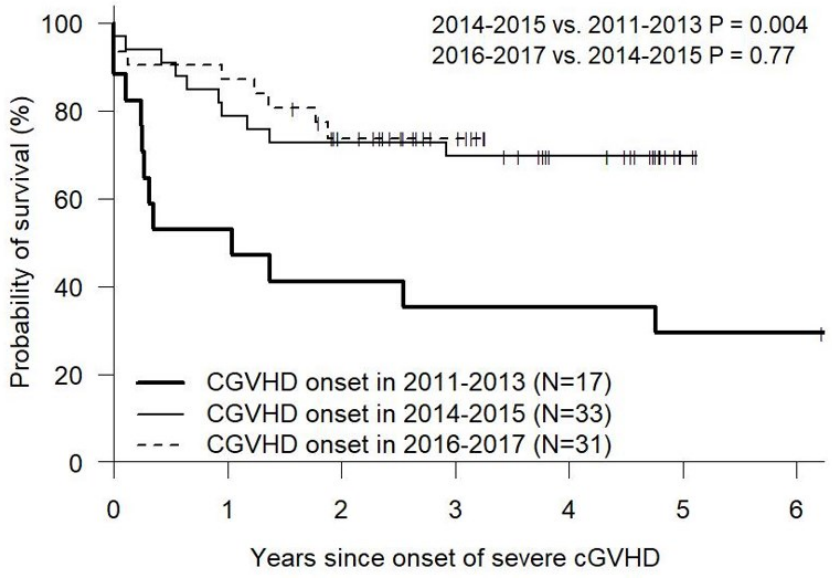
decline of the belumosudil OS extrapolations towards the BAT OS curve, with convergence after approximately 25 years (Figure 9).

The EAG notes that as a result of the post five-year risk of death assumption, OS for BAT is intrinsically linked to OS for belumosudil, such that changes to OS modelling assumptions for BAT affect belumosudil. The EAG's clinical experts considered that OS for BAT is likely to be low as at this point in the treatment pathway (third line of systemic treatment for cGvHD), as there are limited effective treatment options.

The EAG's clinical experts were concerned that a substantial proportion of time was spent in the failure health state for BAT patients ([redacted] years in the failure state out of [redacted] years alive). However, the company explained that in the failure health state most patients' failure event was initiation of new systemic cGvHD treatment (see Section 4.2.4.2 for further details). Additionally, in their clarification response to question B5, the company provided evidence to demonstrate that survival of cGvHD patients has improved over the last 10 years (Figure 14 below).[34]

Together with the evidence from REACH-3, where median OS was not reached for the BAT arm, the EAG considers it is not unreasonable that BAT patients will fail on third line treatment and move the next line of therapy quickly and spend most of their estimated mean life years in the failure health state.

Figure 14. Probability of survival after the onset of cGvHD for cohorts diagnosed over the last years. Reproduced from Bashey *et al*.[34] (reproduced from Figure 8 of the company clarification response)



With regards to an OS, the EAG's clinical experts advised that treatments which are effective at treating cGvHD will improve survival, but that if a patient survives beyond five years, then OS between treatments are likely to be similar. Additionally, the company explained in its response to clarification A16, that there is no strong rationale for belumosudil to be life extending beyond the benefit of treating cGvHD and keeping patients' failure-free. Including the company's assumption of risk of death equal to BAT after five years (which is notably after the observed period of data from the pooled analysis), the estimated life years gained was [redacted] life years between belumosudil and BAT.

Nonetheless, given the uncertainty around OS, the EAG considers that a scenario which removes the OS benefit for belumosudil might be useful for the committee to consider (see Section 6.2 for scenario results) and has included it as part of its preferred base case. The scenario resulted in a dominant ICER for belumosudil, which is driven by a reduction in disease management and subsequent treatment costs associated with a shorter duration in the failure health state ([redacted] years versus [redacted] years). The EAG notes that by removing the OS benefit, belumosudil patients are dying quicker in the model, but spend more of their time in the failure-free health state. The EAG explored a scenario around its base case where the OS benefit for belumosudil is not removed from the model for committee consideration, also presented in Section 6.3. The EAG notes that the inclusion of an OS benefit for belumosudil+BAT for the EAG base case has a substantial impact on the ICER, changing it from dominant to £28,943. Therefore, the EAG recommends that the committee obtain advice from its clinical experts on the clinical plausibility of an OS benefit for belumosudil+BAT.

4.2.5 Failure events

Within the failure health state, patients are stratified by their failure event (initiation of a new systemic cGvHD treatment or recurrent malignancy) and are assigned failure-specific costs and an overall utility value for the failure health state (see Section 4.2.6.1). Kaplan-Meier plots for the distribution of failure events by cause were available from both the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (Figure 9 and 10, Appendix N of the company submission) and REACH-3 (Figure S5 in Zeiser *et al.*).[10] Table 74 in Appendix 9.1 of this report presents the distribution of failure events by cause, including non-relapse mortality, from the updated pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and REACH-3.

In their clarification response to questions B3 and B4, the company explained that in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and REACH-3, most failure events were

attributed to initiation of a new cGvHD treatment rather than a recurrent malignancy. In the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, [redacted] patients had a recurrent malignancy and in REACH-3 only 4.3% of BAT patients had a relapse of their underlying disease.

The company explained that as OS in the model includes non-relapse mortality, this was not modelled separately as a failure event and so proportions presented in Table 74, Appendix 9.1 are not used in the economic model. Instead, the company reweighted the proportions of initiation of new systemic cGvHD therapy and recurrent malignancy to add up to 100% (presented in Table 45 below) and it is these proportions that inform the failure health state in the economic model.

The company assumed that from 36 months onwards, all new failure events were due to initiation of a new systemic cGvHD treatment. Additionally, if no failure events were observed during a period or all failure events were due to non-relapse mortality, the company assumed for that period failure events were due to initiation of a new systemic cGvHD treatment.

The EAG notes that for the belumosudil 200 mg BID arm in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, no recurrent malignancies were observed. Additionally, published failure event data for REACH-3 were presented in six-month periods, thus the company assumed the same proportion of failure events for the 0-3 months period as for the 3-6 month period.

Table 45. Distribution of failure events included in the economic model

Time period	Belumosudil 200 mg QD		Belumosudil 200 mg BID		BAT	
	Initiation of new systemic cGvHD therapy	Recurrent malignancy	Initiation of new systemic cGvHD therapy	Recurrent malignancy	Initiation of new systemic cGvHD therapy	Recurrent malignancy
0-3 months	71.05%	28.95%	100.00%	0.00%	94.97%	5.03%
3-6 months	85.11%	14.89%	100.00%	0.00%	94.97%	5.03%
6-9 months	88.78%	11.22%	100.00%	0.00%	82.11%	17.89%
9-12 months	83.08%	16.92%	100.00%	0.00%	82.11%	17.89%
12-18 months	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
18-24 months	83.33%	16.67%	100.00%	0.00%	100.00%	0.00%
24-30 months	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
30-36 months	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
36+ months	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%

Abbreviations: BAT, best available therapy; BID, twice daily; cGvHD, chronic graft versus host disease; QD, once daily.

4.2.5.1 EAG critique

Most failure events for all treatment arms in the model related to initiation of a new systemic treatment for cGvHD, which the EAG considers reflects what was observed in ROCKstar, KD025-208 and REACH-3. Notably, the proportion of recurrent malignancies was lower for BAT patients compared to belumosudil patients, based on data from the respective trials.

The EAG was concerned that after 36 months, the risk of recurrent malignancy was zero. However, the EAG's clinical experts advised that the longer a patient remains relapse-free, the risk of recurrent malignancy is reduced. Furthermore, this was observed in the clinical trial data for belumosudil and BAT, with low numbers of patients experiencing a relapse in their underlying disease.

Nonetheless, the EAG's clinical experts advised that risk of recurrent malignancy is unlikely to be zero at three years. As such, during the clarification stage the EAG requested, and the company provided a scenario exploring risk of recurrent relapse of 5% per year after 36 months. The results of the scenario are provided in Section 5.2.2 and demonstrates that including a risk of recurrent malignancy after 36 months for the remainder of the time horizon of the model had limited impact on the ICER.

4.2.6 Health-related quality of life

Quality-adjusted life-years accrued by the patient cohort in each model cycle are dependent on the following:

- utility attributable to each model health state including;
    - response status within the failure-free health state;
- the partial loss of utility due to adverse events and intravenous (IV) infusions;
- impact on caregiver HRQoL; and
- an age- and sex-related reduction in quality of life.

Table 46 summarises the utility values informing the economic model. The estimates for the failure-free health state were updated during the clarification stage to use the September 2022 data cut from ROCKstar.

Table 46. Summary of utility values used in the model

Health state/ parameter	Utility value	SE	Source/assumption
Failure-free – complete response	[redacted]	0.007	PROMIS-GH utility data from ROCKstar mapped to EQ-5D-3L.
Failure-free – partial response	[redacted]	0.007	
Failure-free – lack of response	[redacted]	0.007	
Failure – new cGvHD systemic therapy	0.479	0.036	Assumed to be the same as failure – recurrent malignancy
Failure – recurrent malignancy	0.479	0.036	Weighted average of published utility values for AML, ALL, CML and CLL.
AE Disutilities			
Pneumonia	-0.195	-0.039	NICE TA359[35] (SE not reported; assumed 20% of the mean)
Hypertension	-0.020	-0.004	NICE TA689[36] (SE not reported; assumed 20% of the mean)
Anaemia	-0.090	-0.018	
Thrombocytopenia	-0.110	-0.022	
Hyperglycaemia	0.000	0.000	Assumption
Gamma-glutamyl transferase increased	0.000	0.000	Assumption
Diarrhoea	-0.176	-0.035	NICE TA827[37] (SE not reported; assumed 20% of the mean)
Central line-related infections	0.000	0.000	Assumption
IV infusion	-0.037	0.010	Matza *et al.* 2013[38]
Caregiver disutilities			
Failure-free - In response - PR	-0.045	0.057	Acaster *et al.* 2013[39]
Failure-free - Lack of response	-0.045	0.057	
Failure – New cGvHD Systemic Therapy	-0.142	0.062	
Failure – Recurrent Malignancy	-0.142	0.062	

Abbreviations: AE, adverse events; ALL, acute lymphoblastic leukaemia; AML, acute myelogenous leukaemia; cGvHD, chronic graft versus host disease; CLL, chronic lymphocytic leukaemia; CML, chronic myelogenous leukaemia; IV, intravenous; PR, partial response; SE, standard error.

The details of each utility category are given in the following subsections.

4.2.6.1 Health state utility values

Utility values based on response status for the failure-free health state are derived from PROMIS-GH utility data from ROCKstar (described in Section 3.3.4) mapped to EQ-5D-3L using an algorithm by Thompson *et al.*[13] PROMIS-GH data were not collected in KD025-208. During the clarification stage, the company updated the utility analysis to be based on the September 2022 data cut from ROCKstar. Additionally, the company confirmed in their clarification response that the Oxford

Population Health HERC database of mapping studies was searched to find a suitable mapping algorithm for PROMIS-GH to EQ-5D-3L and Thompson *et al*.[13] was selected as it was the most recently developed algorithm.

In ROCKstar, PROMIS-GH data were collected in the modified intention to treat population (mITT) on day one of cycle one to five and then on day one of every other cycle until end of treatment. Figure 15 presents the mean mapped EQ-5D-3L score per treatment arm and overall in ROCKstar, including number of responses collected during each cycle.

Figure 15. Mean mapped EQ-5D-3L per cycle (reproduced from Figure 10 of the company clarification response)

The company used mixed-effect repeated linear regression models to analyse mapped EQ-5D-3L data from ROCKstar. All patients with a non-missing baseline utility value and at least one non-missing post-baseline utility value were included in the analysis. Six models were explored and considered time-varying covariates including treatment failure, response, and lack of response. Specifications of the utility models are presented in Table 12, Appendix N of the company submission.

In their clarification response, the company confirmed that utility values that inform the response categories within the failure-free health state were based on model six, which included failure and response as covariates (without granularity for type of failure event or depth of response). The company explained that their choice of regression model was based primarily on face validity of estimates and model structure rather than model fit statistics.

Based on the company's utility regression analysis, the utility values in the failure-free health state for complete response (CR), partial response (PR) and lack of response (LR) were [redacted], [redacted] and [redacted]. The company assumed the utility value for CR to be equal to PR as there were only a small number of utility observations specific to CR.

The company provided a scenario where response is excluded from the model and as part of this scenario, estimated a single utility value for the failure-free health state. In their clarification response, the company explained that mean utility value for the failure-free health state ([redacted]) was estimated using model one, which only included failure as a covariate (Appendix N of the company submission).

Even though some of the company's regression models included failure by cause as a covariate, the company considered estimates of utility related to failure lacked face validity. Based on the models that included type of treatment failure as a covariate, utility values estimated for failure-new therapy ([redacted]) and failure-recurrent malignancy ([redacted]) were higher than for failure-free ([redacted]) (estimates taken from Table 40 of the company clarification response). The company explained that only 25 patients with failure had utility measurements recorded in 74 visits.

As such, the company estimated a utility value for the failure health state from published data. The company conducted literature searches in related disease areas (indications for the most recent transplants in ROCKstar). Table 16 in Appendix N of the company submission presents a summary of the indications for most recent transplant in ROCKstar. The company identified utility data for acute myelogenous leukaemia (AML), acute lymphoblastic leukaemia (ALL), chronic myelogenous leukaemia (CML) and chronic lymphocytic leukaemia (CLL) and these malignancies represented 64.4% of the population in ROCKstar.

The company estimated a weighted average utility for recurrent malignancy by reweighting the proportion of patients with AML, ALL, CML and CLL from ROCKstar to equate to 100% and applying the weights to the associated utility value for each malignancy. Table 47 presents an overview of the

utility data and sources used to estimate the weighted utility value for failure – recurrent malignancy.

Table 47. Utility values used to estimate utility for recurrent malignancy in the model (reproduced from Table 17, Appendix N of the company submission)

Indication	Overall weight in ROCKstar	Reweighted proportion	Mean Utility value	SE	Source	Description
AML	40.9%	63.5%	0.51	0.032*	TA642 (based on Joshi *et al.*)[40, 41]	Utility of treatment failure/relapse/refractory disease in AML in Joshi *et al.*[40] Used as utility value for post-event with HSCT in relapsed or refractory AML in TA642[41]
ALL	14.4%	22.4%	0.30	0.04	Aristides *et al.*[42]	Utility of progressive disease in relapsed or refractory B-precursor ALL
CML	6.1%	9.4%	0.59	0.059*	TA813 (derived from TA451)[43, 44]	Utility of relapse for stem cell transplant patients with progressive disease in third line CML in TA813[44]
CLL	3.0%	4.7%	0.68	0.021*	Beusterien *et al.*[45]	Utility of progressive disease in CLL
Total / weighted average	**64.4%**	**100%**	**0.479**	**0.036***	**-**	**-**

Abbreviations: ALL, acute lymphoblastic leukaemia; AML, acute myelogenous leukaemia; CLL, chronic lymphocytic leukaemia; CML, chronic myelogenous leukaemia; HSCT, haematopoietic stem cell transplantation; SE, standard error; TA, technology appraisal

*Company calculated values

For patients whose failure event was a new cGvHD systemic treatment, the company were unable to identify published utility data related to progression to next line of systemic therapy to inform the model. Thus, the company assumed the utility value for failure – new cGvHD systemic treatment was equal to the estimated weighted utility for failure – recurrent malignancy. As such, the EAG considers that the utility value of 0.479 represents the utility value for failure.

4.2.6.2 Adverse event and IV infusion disutilities

The company obtained AE disutility values and duration of each AE from relevant technology appraisals in indications related to the underlying disease of patients in ROCKstar (AML and CLL).[35-37]

The incidence of each AE by treatment arm included in the model is presented in Sections 3.3.6 and 3.5.10.

The EAG notes that the company assumed that central line-related infections would only apply to patients on ECP in the BAT arm of the model. Based on clinical expert opinion, the company assumed 20% of patients treated with ECP would experience a central-line-related infection.

The company estimated the QALY loss associated with each AE by multiplying the disutility for each AE by the associated duration (in days) of the AE. Table 47 of the CS presents the disutility, duration and QALY loss associated with each AE included in the model.

A one-off AE related QALY loss for each treatment arm was estimated by multiplying the QALY loss of each AE by the treatment-specific incidence of the AE (Table 48). The company assumed that impact of AEs was assumed to occur in the first four weeks of treatment and as such the one-off AE QALY loss was applied in the first model cycle only.

Table 48. One-off AE related QALY loss by treatment

Treatment arm	One-off QALY loss
Belumosudil 200 mg QD	-0.001
Belumosudil 200 mg BID	-0.001
BAT	-0.002

Abbreviations: AE, adverse event; BAT, best available therapy; BID, twice daily; QALY, quality-adjusted life-year; QD once daily.

Based on feedback from the company's clinical experts, the company included a disutility associated with treatments that are administered via intravenous (IV) infusion (ECP and rituximab). A disutility value associated with IV infusion of -0.037 (SE 0.010) with an assumed duration of 28 days, was based on published data from Matza *et al.*,[38] identified in the company's HRQoL SLR.

The company assumed that the IV infusion disutility value would be applied to all patients receiving ECP and rituximab for the duration of time patients are on these treatments.

4.2.6.3 Caregiver disutility

For the base case, the company included the impact on HRQoL of informal carers of patients with cGvHD. The NICE methods guide states that, "*evaluations should consider all health effects for patients, and, when relevant, carers. When presenting health effects for carers, evidence should be*

provided to show that the condition is associated with a substantial effect on carer's health-related quality of life and how the technology affects carers".[31]

The company stated that there is a lack of published information on the burden and associated disutilities for caregivers of patients with cGvHD. A Sanofi sponsored disease specific programme (DSP) analysis found that a substantial amount of time is spent on care by informal carers of patients with cGvHD and that anxiety increase with the disease severity of the patient. Additionally, the company's clinical experts advised that there are impacts (emotional, financial, social) on the quality of life of informal carers of patients with cGvHD and these are akin to the impact on caregivers of patients with multiple sclerosis (MS).

As such, the company identified a study by Acaster *et al.* which estimated the utility decrements for carers of patients with MS stratified by self-reported disability as measured by the patient determined disease steps (PDDS) scale.[39] The PDDS scale is a self-assessment scale that assesses functional disability in people with MS and ranges from level (normal) to level 8 (bedridden).

The company assumed that for failure-free patients with a partial or lack of response, the utility decrement for caregivers of cGvHD patients would be akin to carers of MS patients with a PDDS level of 2 (moderate disability) or 3 (gait disability). For patients in the failure health state, regardless of failure event, the utility decrement for caregivers of cGvHD patients would be akin to carers of MS patients with a PDDS level of 4 (early cane). Table 49 presents an overview of the caregiver disutilities included in the economic model.

Table 49. Description of carer disutilities included in the economic model.

Parameter	Disutility (SE)	Description and source
Failure-free partial response and lack of response	-0.045 (0.057)	Acaster 2013.[39] Relates to MS patient PDDS level 2 (moderate disability) & level 3 (gait disability).
Failure – new cGvHD systemic therapy and recurrent malignancy	-0.142 (0.062)	Acaster 2013.[39] Relates to MS patient PDDS level 4 (early cane).

Abbreviations: cGvHD, chronic graft versus host disease; MS, multiple sclerosis; PDDS, patient determined disease steps; SE, standard error

4.2.6.4 Age- and sex-related utility adjustment

Utilities in the model were adjusted for age and sex, as per the NICE methods guide.[31] The multiplicative approach was used as recommended by the DSU TSD 12.[46] General population utility

values adjusted for age and sex were obtained from the HSE 2014 dataset, as recommended by the DSU.[47]

4.2.6.5 EAG critique

The EAG considers that overall, the company's approach to estimating utility values for the model was thorough. In addition to the regression analyses of utility data from ROCKstar, the company performed a utility elicitation study (reported in Appendix N.4 of the company submission), although this did not produce clinical plausible results, which the EAG agrees is true. While not discussed in the submission, the company also derived utility estimates from their Adelphi disease specific programme (DSP) study.[48, 49] However, the Adelphi DSP study only had two self-reported UK patient responses.[48]

As mentioned in Section 4.2.4.5, the EAG prefers to remove response from the model such that patients in the failure-free health state only have a single utility value. In the company's scenario, the failure-free utility value was estimated to be [redacted]. The EAG considers the company's approach to estimate the failure-free utility value was reasonable.

However, with regards to response utility values included in the base case, the company assumed the utility value for CR to be equal to PR as there were only a small number of utility observations specific to CR. In ROCKstar, only six patients had a complete response. As such, the EAG agrees that utility observations for complete responders maybe subject a high degree of uncertainty. Nonetheless, in their clarification response (Table 34) the company provided mapped EQ-5D-3L utility values for CR ([redacted]), PR ([redacted]) and LR ([redacted]). The EAG ran a scenario with these utility values, but as they are similar to the company base case values, the ICER remained the same (£3,571).

The EAG considers that the utility value for failure – new cGvHD systemic therapy is key driver of QALYs in the model, as patients in the BAT arm spend the majority of their time in this health state. In the company's base case, the utility value for failure – new cGvHD systemic therapy is assumed to be equal to the utility value for failure – recurrent malignancy (0.479). Thus, there is a single utility value of 0.479 for the failure health state. The utility value for failure – recurrent malignancy was estimated as a weighted average of utility values for progressed/relapsed disease for AML, ALL, CML, CLL and thus does not reflect patients who change treatment for cGvHD rather than experience a recurrence in their underlying malignancy.

The EAG's clinical experts explained that a recurrent malignancy means patients are no longer immunosuppressed, which essentially resolves the cGvHD, and thus cGvHD treatment would stop. However, recurrent malignancy is a far worse outcome for patients than a change in cGvHD treatment. As such, the EAG was concerned that utility value for failure - new cGvHD systemic therapy was too low. The EAG investigated the company's utility regression models and found that some of the regression models estimated a utility value for based on type of treatment failure.

During the clarification stage, the EAG requested, and the company provided, utility data for new cGvHD systemic therapy and recurrent malignancy (Table 40 of the company clarification response). The resulting utility value of [redacted], based on 69 observations from 22 patients, for failure – new cGvHD systemic therapy was higher than the utility value for failure-free ([redacted]), which was based on 1,197 observations from 140 patients. The company stated that analysis demonstrated that the utility value for failure – new cGvHD systemic therapy lacked face validity.

The EAG considers that there is a high degree of uncertainty around the utility value for failure – new cGvHD systemic therapy due to the limited number of observations. Additionally, as advised by its clinical experts, the EAG considers that patients who require a change in treatment for their cGvHD represent a sicker population as the failure event indicates more advanced cGvHD.

The company provided several scenarios around its base case in its original submission exploring alternative utility values for failure – new cGvHD systemic therapy. However, the EAG notes that two scenarios were particularly useful and included one scenario using a utility value obtained from the company's Adelphi DSP study (0.52) and one from a study by Crespo *et al*. ([redacted]).[20] The EAG investigated the sources of the alternative utility values and considered both had flaws. The utility value from the Adelphi DSP study was based on sample of 10 cGvHD patients from Europe, but only two were from the UK. It is unclear how the utility values were derived as, most UK patients in the Adelphi DSP study were in remission for their disease (98%) but the patient reported utility value used for the failure – new cGvHD systemic therapy was for treatment failure, but not specific cause of failure.

With regards to the utility value from Crespo *et al*.,[20] the company estimated a utility decrement based on the difference between the utility values for stable disease (0.736) and progression (0.696) and applied the utility decrement to the base case utility value for failure-free – lack of response ([redacted]). Progression in the Crespo *et al*.[20] study was defined as progression of cGvHD for specific

organs. The source study of utility values from Crespo *et al*.,[20] was Pidala *et al*.[50] and reported utility values associated with severity of cGvHD using the SF-36.

The EAG considers that both of the company's alternative utility values for failure – new cGvHD systemic therapy better reflect the EAG's clinical experts' view, but that there is uncertainty around the methods from the source studies. As a pragmatic approach, the EAG has run two scenarios (presented in Section 6.2), one using the progression utility value from Crespo *et al*.[20] (0.696) and another scenario exploring the midpoint of the Adelphi DSP study utility value for treatment failure and progression utility value from Crespo *et al*.,[20] estimated to be 0.608. As a pragmatic approach, the EAG included the estimated midpoint utility value of 0.608 in its base case, and with scenarios explored the Adelphi DSP study utility value for treatment failure and progression utility value from Crespo *et al*.,[20] presented in Section 6.3.

With regards to the inclusion of caregiver disutility values, the EAG's clinical experts advised that patients with cGvHD may have significant disability depending on the organs affected and this will likely have a negative impact on the HRQoL of informal caregivers. The EAG's clinical experts considered the company's assumption that HRQoL evidence for caregivers of patients with multiple sclerosis would be akin to caregivers of cGvHD patients, in lieu of any specific cGvHD evidence, is reasonable. However, the EAG's clinical experts advised that impact on caregiver HRQoL would not be the same for patients who initiate a new systemic cGvHD therapy compared with patients who have a recurrent malignancy. Recurrent malignancy would be associated with a greater psychological and physical burden for caregivers. Instead, the EAG's clinical experts advised that it might be reasonable to assume the caregiver impact for patients who initiate a new systemic cGvHD therapy would be akin to patients who have had a partial or lack of response.

During the clarification stage, the EAG requested, and the company provided, a scenario where caregiver disutility for failure – new cGvHD systemic therapy was the same as failure-free PR/LR. The scenario has minimal impact on the ICER, increasing it from £3,571 to £4,065 but the EAG considers it is a more appropriate assumption and has therefore included it in the EAG base case, presented in Section 6.3.

The EAG highlights that in the company's Adelphi DSP study, UK physicians (n=40) reported that 52% of their cGvHD patients had an informal caregiver, but that 95% of the UK physicians considered

their patients did not need a caregiver.[48] As such, the EAG has included a scenario around its base case where caregiver disutilities are removed from the model (see Section 6.3).

The EAG notes some secondary issues concerning disutilities associated with AEs and IV infusions, which are not primary drivers but that should be changed in the model. The EAG was concerned that an assumption of no disutility for central line-related infections was inconsistent with the high-cost nature of treating this AE (see Section 4.2.6.2). Instead, the EAG considers that the disutility value and duration associated with infections and infestations from TA689 (-0.22, 14 days) should have been used.[36] The EAG ran a scenario using the TA689 infection and infestations disutility for central line-related infections, which had minimal impact on the ICER (see Section 6.2), but for completeness, has included it in the EAG base case, presented in Section 6.3.

Lastly, during the clarification stage, the EAG requested the company to explore the inclusion of concomitant medications for the belumosudil arms of the model, as these were excluded in the company base case (see Section 4.2.7.9). Concomitant medications in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (August 2021 data cut) included ECP, which in the BAT arm of the model incurs a disutility of -0.037 related to treatment with IV infusion. However, the company did not include the IV infusion disutility for belumosudil in their concomitant medication scenario, stating in the model that they assumed the impact of this would be captured as part of the trial-based utility values used in the failure-free health state.

The EAG agrees with the company's approach to exclude IV infusion disutility for belumosudil but considers that this extends to BAT as the utility values used in the model are not treatment specific. The EAG acknowledges usage of treatments administered via IV infusions would be different for the BAT arm in clinical practice, but that it is a conservative assumption to remove the IV infusion disutility for BAT, which has minimal impact on the ICER (see Section 6.2). As such, the EAG prefers to remove the IV infusion disutility from its base case, presented in Section 6.3.

4.2.7 Resource use and costs

The company included the following costs in the economic model: drug acquisition, administration, accommodation, disease management, adverse events, subsequent treatments, and recurrent malignancy. The details for each of these are given in the following subsections.

Unit costs used in the model reflect 2021 prices and where necessary, published costs for previous years were inflated using the Personal Social Services Research Unit (PSSRU) hospital and community health services pay and prices index.[51]

4.2.7.1 Drug acquisition costs

Belumosudil is a fixed dose drug given as a 200 mg tablet taken orally QD. The SmPC for belumosudil recommends that the dose of belumosudil should be increased to 200 mg BID when co-administered with strong CYP3A inducers or PPIs.[1] The list price per box of 30 x 200 mg tablets is £6,708. There is an approved patient access scheme (PAS) discount in place for belumosudil of [redacted] and this has been included in all analyses presented in this report. The discounted pack price of belumosudil is [redacted], resulting in a cost per cycle (28 days) for the QD regimen of [redacted] and for the BID regimen of [redacted].

The comparator in the model, BAT, is comprised of several treatment options, most of which are used off-label to treat cGvHD. As mentioned in Section 2.3.3, the basket of treatments included as part of BAT was sourced from REACH-3 and adjusted to remove treatments that are not used in the UK (based on clinical expert opinion obtained by the company). Table 50 presents the distribution of BAT components included in the model. The unit costs of each treatment in BAT are presented in Appendix K of the CS and were sourced from the British National Formulary (BNF), Drugs and pharmaceutical electronic market information tool (eMIT), the NHS Electronic Drug Tariff and for ECP, a study by Button *et al*.[26, 27, 29, 52]

Table 50. Distribution and dosing regimen of BAT components included in the economic model.

Treatment	Proportion used in the company base case	Dosing regimen and administration
ECP	64.6%	IV administration – one cycle (two treatments on consecutive days) every 2 weeks for the first 12 weeks. Reduced to one cycle every 4 weeks for partial responders after 12 weeks.[52]
Mycophenolate mofetil	22.2%	1000 mg taken orally twice daily.[53]
Imatinib	5.1%	100 mg daily for first 4 weeks, 200 mg daily in weeks 5-12, 400 mg daily after 12 weeks, taken orally.[54]
Sirolimus	4.4%	6 mg loading dose on first day, followed by 2 mg QD, taken orally.[55]
Rituximab	3.8%	IV administration – 375 mg/m^2 per week for 4 consecutive weeks.[56]

Abbreviations: BAT, best available therapy; IV, intravenous; QD, once daily.

The maximum treatment duration for all treatments, except rituximab, was assumed to be five years. The maximum treatment duration for rituximab was assumed to be four weeks and this was based on the dosing regimen used in the Phase II study of rituximab for the treatment of steroid-refractory cGvHD.[56]

The number of administrations per four-week (28-day) cycle is presented in Table 2 Appendix K of the CS. For ECP, the company assumed that after 12 weeks, treatment is reduced to one cycle every four weeks for partial responders. The company assumed that in week 13-24, 40% of the patients, and after 24 weeks, 50% of the patients are assumed to be partial responders.

In the model, the company assumed that there is no vial sharing for rituximab and adopted an approach to estimate the minimum number of vials needed based on dose according body surface area (BSA) (minimum wastage approach). In their clarification response, the company explained that based on a normal distribution of BSA from the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup, the minimum number of vials needed for the relevant BSA based dose was estimated and used to estimate a weighted average cost of rituximab per administration.

The company calculated a weighted cost of BAT over time by estimating the cost of each treatment based on its dosing regimen and acquisition cost per cycle and weighting the cost by the proportion of each treatment assumed. Table 51 presents the costs of BAT included in the model.

Table 51. BAT drug acquisition costs

Model cycle	Cost per cycle
1st cycle (weeks 1-4)	£4,285.38
2nd to 3rd cycle (weeks 5-12)	£4,265.47
4th to 6th cycle (weeks 13-24)	£3,415.04
7th cycle onwards (weeks 25+) – maximum treatment duration of 5 years.	£3,202.32

Abbreviations: BAT, best available therapy.

A confidential discount is available for rituximab. The source of the confidential price for rituximab is the commercial medicines unit (CMU). As such, the EAG has produced a confidential appendix to the EAG report. Analyses included in the confidential appendix include the company base case results, scenario analyses and EAG base case and scenario analyses.

4.2.7.2 Drug administration costs

Administration costs for oral drugs, which includes belumosudil, mycophenolate mofetil, imatinib and sirolimus, was assumed to be zero. For ECP, the company estimated the administration cost of treatment to be the cost of two hours of specialist nurse time (£110), based on PSSRU.[51]

Rituximab has a maximum treatment duration of four weeks, equivalent to four IV administrations. The company estimated rituximab administration costs based on assumptions in TA627 (lenalidomide with rituximab for previously treated follicular lymphoma).[57] In TA627, NHS reference codes SB13Z (deliver more complex parenteral chemotherapy at first attendance) and SB15Z (deliver subsequent elements of a chemotherapy cycle) were used.[30] As a simplification for the base case, the company used the cheaper of the two costs (SB13Z, £426.80) as the administration cost for rituximab, which is considered by the EAG to be a conservative assumption. Table 52 presents the BAT administration costs per cycle included in the model.

Table 52. BAT drug administration costs*

Model cycle	Cost per cycle
1st cycle (weeks 1-4)	£445.72
2nd to 3rd cycle (weeks 5-12)	£395.92
4th to 6th cycle (weeks 13-24)	£316.74
7th cycle onwards (weeks 25+) – maximum treatment duration of 5 years.	£296.94

Abbreviations: BAT, best available therapy.

*Costs include accommodation costs assumed for ECP, discussed in Section 4.2.7.4.

4.2.7.3 *Time to treatment discontinuation*

In the original company submission, a data cut from August 2021 for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) for belumosudil was used to inform the model. However, the company indicated that a later data cut from September 2022 was available and in their clarification response the company updated the model with this data. Figure 16 presents the TTD KM curves based on the August 2021 and September 2022 data cuts from the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup).

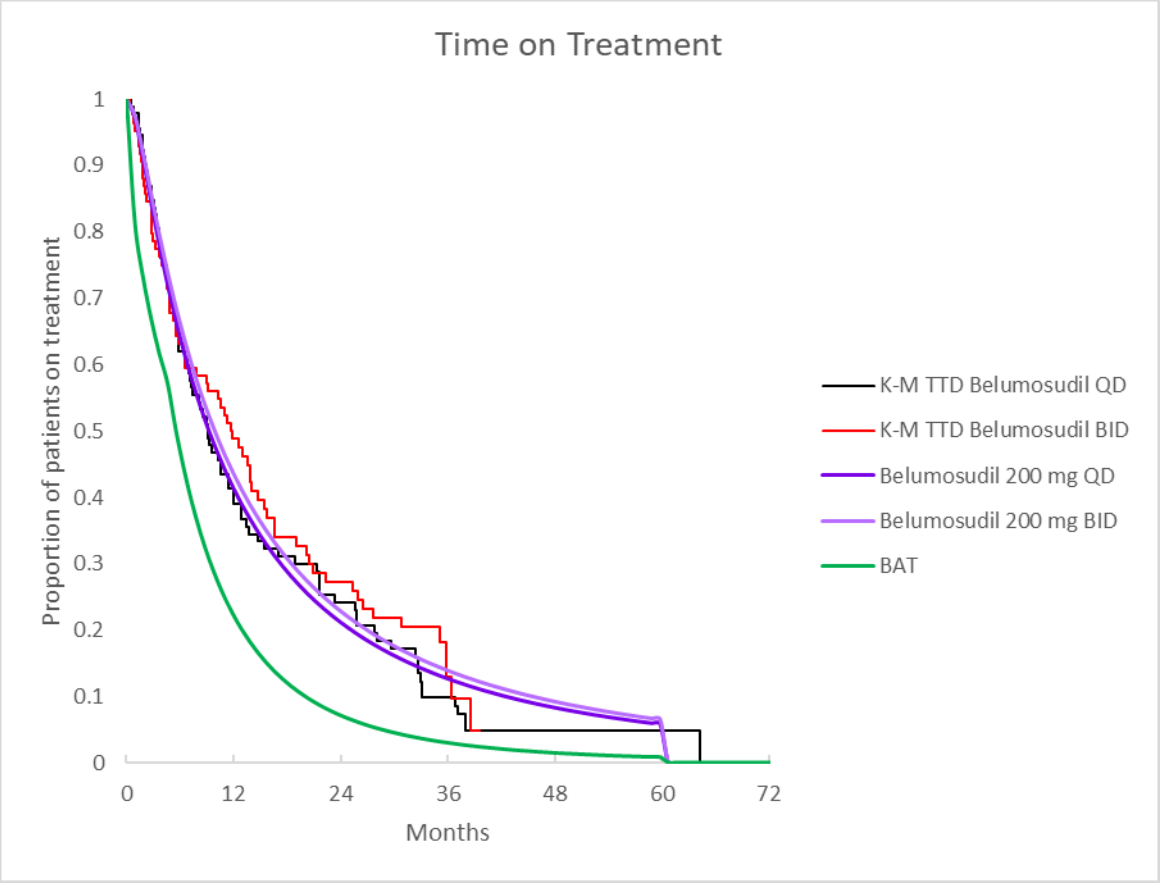
The company presented two TTD curves for the September 2022 data cut based on different definitions of treatment discontinuation. The company explained that ROCKstar had completed for the adult cohort but was still recruiting adolescent patients, thus adult patients who were still on treatment in the 2021 data cut would be classed as having discontinued treatment due to study termination in the 2022. The company produced an adjusted TTD curve where patients who discontinued treatment due to study termination ([redacted] patients) were censored at the time they received their last dose of study drug.

Figure 16. Kaplan-Meier time to treatment discontinuation curves for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) for belumosudil – August 2021 and September 2022 data cuts (reproduced from Figure 1 of the company clarification response)

Based on the PH assessments for the pooled ROCKstar and KD025-208 studies, the company considered that PH held and so decided to jointly fit survival distributions, where treatment arm is a predictor. Based on statistical fit (not provided in the company's clarification response) and clinical plausibility, the company selected the lognormal distribution for belumosudil 200 mg QD and belumosudil 200 mg BID.

For BAT, only median treatment duration (5.54 months) from REACH-3 was available.[10] Therefore, to estimate a TTD curve for BAT, the company calibrated a HR to the belumosudil 200 mg QD TTD curve to estimate the reported median TTD for BAT from REACH-3. The estimated TTD HR for BAT was 1.71. As such, TTD for BAT is intrinsically linked to modelling assumptions for the belumosudil TTD curves. Figure 17 presents modelled TTD for belumosudil and BAT.

Figure 17. Modelled time to treatment discontinuation



The EAG notes that TTD is capped by FFS. Additionally, the company assumed a maximum treatment duration of five years for all treatment arms.

4.2.7.4 Accommodation costs associated with ECP

According to clinical experts (both company's and EAG's), ECP is can only be administered at specialist centres, which are limited in number around England. For the base case, the company assumed that for patients that require ECP and are not local to a specialist centre, overnight accommodation would be required for treatment and this cost would be reimbursed by NHS England. The company assumed that 50% of patients on ECP require overnight accommodation (one night stay). The company assumed an accommodation cost of £150 per night, based assumptions included in a previous CAR-T technology assessment (TA559).[58]

4.2.7.5 Disease management costs

The company used HES secondary care data to estimate the disease management costs in the model.

The HES study population included patients aged ≥12 years with an allogenic haematopoietic stem cell transplant (alloHSCT) between 1 April 2017 and 31 December 2020 and data included in the study were up until 31 March 2022.[59] The HES database contained information on reimbursed diagnoses and procedures from all NHS inpatient admissions, outpatient appointments and emergency care attendances in England. From the HES data, the company estimated the mean costs of attendances for inpatient, outpatient, A&E and ICU for non-GvHD patients, cGvHD patients with non-high-cost therapy, cGvHD patients with first high-cost therapy and cGvHD patients with at least 2 high cost therapies. Treatments considered as high-cost therapy in the analysis included ECP, rituximab and protein tyrosine kinase inhibitors (e.g., ruxolitinib and imatinib). The EAG notes that the HES study did not include the cost of treatments.

The company's assumptions for the disease management of patients included in the model are as follows:

- **Patients in the failure-free health state with CR**: assumed to be the mean cost incurred by haematopoietic stem cell transplant (HSCT) patients without cGvHD in the HES study throughout the time horizon of the model.
- **Patients in the failure-free health state with PR and LR**: assumed to be the mean cost incurred by all HSCT patients with cGvHD in the HES study in the first year, with a linear decrease in each year to reach the disease management cost of patients with CR in the fifth

year. The model assumes that patients remaining failure-free incur the same costs regardless of response status after the fifth year.

- **Patients in the failure state with a new systemic therapy**: assumed to incur the mean cost of HSCT patients with two or more records of high-cost therapy in the HES study.
- **Patients in the failure state with recurrent malignancy**: These were not available from the HES study and so were sourced from TA642[41] that included the total costs incurred by patients with AML-related inpatient admissions, ICU, emergency department, outpatient visits, diagnostic procedures, lab tests, and blood transfusions. Acute myeloid leukaemia was the most common underlying malignancy in ROCKstar (40.9%).

Table 54 of the CS presents the annual disease management costs by health state for years one to five and beyond. For patients with partial or lack of response, the company assumed that the longer these patients occupied the failure-free health state, disease management costs over time would reduce. In the model, the company assumed that disease management costs for PR and LR patients would decline linearly over five years until they reached the same disease management costs as complete responders for years five and beyond. The EAG's clinical experts agreed that, irrespective of response, if patients remain failure-free disease management is likely to reduce over time.

For the company's scenario which excludes response from the model, disease management costs the failure-free health state is assumed to be the same as partial and lack of response.

Table 53. Disease management costs per cycle per year (reproduced from Table 55 of the CS)

Health states	Mean cost per cycle per year					Source
	1st year	2nd year	3rd year	4th year	≥5th year	
Failure-free						
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Partial response and Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Failure						
New cGvHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Recurrent malignancy	£2,719.46	£2,719.46	£2,719.46	£2,719.46	£2,719.46	NICE TA642[41]

Abbreviations: cGvHD, chronic graft versus host disease; CS, company submission.

4.2.7.6 *Cost of subsequent treatments*

For patients whose failure event was initiation of a new cGvHD treatment, costs of subsequent treatment were included in the model. The company stated that there was a lack of evidence for fourth line or later cGvHD treatment, as the treatment pathway is patient/case dependent. Thus, as a simplifying assumption, the company assumed that basket of subsequent treatments was the same as BAT, regardless of the treatment patients received at third-line. Furthermore, the company assumed that patients spend 60% of their remaining lifetime on subsequent treatment but also assume that these treatments are given for a lifetime. Thus, the EAG considers that the practical implementation of subsequent treatments in the model is that 60% of patients whose failure event is initiation of new cGvHD treatment receive subsequent treatments for life (as stated in Appendix K of the CS) and 40% do not go on to have any further treatment. The proportion of patients on subsequent treatment (60%) are evenly distributed amongst the treatments in the basket, except for rituximab (Table 54).

As a scenario, the company explored subsequent treatment durations, presented in Table 54, based clinical expert feedback and implemented as a one-off cost to incident subsequent treatment cGvHD starters.

Table 54. Distribution and duration of subsequent treatments

Subsequent treatment	Proportion on treatment	Duration of treatment	
		Base case	Scenario
ECP	14.5%	Lifetime	24 weeks
Mycophenolate mofetil	14.5%	Lifetime	24 weeks
Sirolimus	14.5%	Lifetime	16 weeks
Rituximab	2.0%	Lifetime	4 weeks
Imatinib	14.5%	Lifetime	24 weeks
No subsequent treatment	40%	-	-

Abbreviations: ECP, extracorporeal photopheresis.

The unit costs of each subsequent treatment are presented in Appendix K of the CS and were sourced from the BNF, Drugs and eMIT, the NHS Electronic Drug Tariff and for ECP, a study by Button *et al*.[26, 27, 29, 52] The number of administrations for each treatment was assumed to be the same as for BAT from cycle 7 onwards (week 25+) (except for rituximab, which was assumed to be four administrations per cycle), presented in Table 2, Appendix K of the CS. The per cycle of drug

acquisition and administration cost for subsequent treatments was estimated to be £808.81 and £81.99, respectively.

4.2.7.7 Cost of recurrent malignancy

To inform the cost of treatment for a recurrent malignancy, the company performed a targeted search to identified data from published literature. A one-time cost for post-progression treatment of AML was sourced from TA642, based on AML being the most common malignancy in ROCKstar (40.9%).

The original source of the AML post-progression unit cost per cycle (£5,179.09) was Wang *et al.*,[60] which estimated an average cost of second-line treatment regimens for adult AML patients in the UK Haematological Malignancy Research Network (HMRN). In TA642, the company assumed the average number of cycles was 2.6, resulting in a post-progression cost of £8,264.47. The cost from TA642 was inflated to 2021 prices, resulting in a one-off cost of £8,908 for recurrent malignancy.

4.2.7.8 Costs of adverse events

The costs of AEs in the model were based on the probability of each AE multiplied by the unit cost of the AE. Unit costs of AEs were sourced from NHS references costs and were based on the outpatient setting, except for central line-related infections.

The company assumed central line-related infections were treated in an inpatient setting with the costs obtained from a study by Manoukian *et al.*[61] The unit cost of central line-related infections was based on the direct cost of bloodstream infections in NHS Scotland (£5,917) and inflated to 2021 prices. The EAG notes that the company assumed that central line-related infections would only apply to patients on ECP in the BAT arm of the model. Based on clinical expert opinion, the company assumed 20% of patients treated with ECP would experience a central-line-related infection.

Unit costs of AEs are presented in Table 57 of the CS and incidence of each AE by treatment arm included in the model is presented in Sections 3.3.6 and 3.5.10. Table 55 presents the one-off cost of AEs by treatment arm included in the model. The company assumed that impact of AEs was assumed to occur in the first four weeks of treatment and as such the one-off AE cost was applied in the first model cycle only.

Table 55. One-off AE cost by treatment arm

Treatment	One-off AE Management Cost
Belumosudil 200 mg QD	£ 186.55
Belumosudil 200 mg BID	£ 144.36
BAT	£ 987.35

Abbreviations: AE, adverse event; BAT, best available therapy; BID, twice daily; QD, once daily.

4.2.7.9 EAG critique

The EAG considers there are several key issues with resource use and costs, which are as follows:

- Exclusion of concomitant medication costs from the belumosudil arms of the economic model – primary driver of cost-effectiveness in the model.
- Use of observed KM TTD data to estimate drug acquisition costs for belumosudil.
- Exponential distribution for BAT TTD.
- Inclusion of accommodation costs for patients receiving ECP, assumed to be reimbursed by the NHS.
- Exclusion of monitoring costs from the economic model.
- Maximum subsequent treatment duration – primary driver of cost-effectiveness in the model.

The EAG notes that disease management costs in the model are a primary driver of cost-effectiveness in the model but considers the company's HES study to be thorough, with data reflecting the UK population. Additionally, the EAG considers that disease management cost for the failure-free health state used for the company's scenario, which removes response from the model, is appropriate.

Secondary issues with regards to costs include lack of data around composition of subsequent treatments and second alloHSCT for patients with recurrent malignancy, but the EAG considers these have minimal impact on the ICER. Each of the key issues are discussed in detail in the following subsections.

Concomitant medications

The NICE final scope specifies that the intervention for this appraisal is belumosudil with established clinical management (hereafter referred to as belumosudil+BAT).[5] However, the company included

belumosudil in the model as a monotherapy and costs reflect this assumption. In ROCKstar and KD025-208, concomitant medications were permitted. The EAG considers the exclusion of concomitant medications, which would be considered akin to established clinical management, a significant omission in the cost-effectiveness analysis which was not discussed or justified in the company's original submission. Additionally, clinical outcomes in the model include the efficacy of concomitant medications (the composition of which is similar to the basket of treatments included in BAT but usage of each treatment may be different) and so the EAG considers the costs of these treatments should be accounted for in the economic model as these costs would be incurred in UK clinical practice.

Therefore, the EAG requested the company to justify the exclusion of concomitant medications as part of clarification question B23. The company explained that based on guidance from their clinical trials team who considered that efficacy of a patient's existing treatment package would not "boost" the efficacy of belumosudil.

The EAG considers the company's rationale is not satisfactory as it ignores the fact that belumosudil was given in addition to established clinical management in ROCKstar and KD025-208 (which aligns with the NICE final scope) and this is how it will likely be provided to patients in the NHS. As such, the EAG considers that total costs for belumosudil should include the costs of concomitant medications as these will still be incurred by the NHS, but the usage of these treatments is likely to differ to what is assumed for BAT as a comparator. Thus, the intervention in the model should reflect belumosudil+BAT, rather than belumosudil monotherapy. The EAG notes that in REACH-3, concomitant medications were not permitted, thus ruxolitinib was given only as monotherapy.

Nonetheless, in their clarification response to question B23, the company provided the breakdown of concomitant medications for the treatment of cGvHD used for at least five trial subjects in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (August 2021 data cut). The company explained that due to paucity of time, they were unable to obtain concomitant medication data for the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (September 2022 data cut). Table 56 outlines the concomitant medications for the treatment of cGvHD used for at least five trial subjects reported in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (August 2021 data cut).

Table 56. Concomitant medication use based on the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (August 2021 data cut)

Concomitant treatment	Dosing regimen**	Proportion on concomitant medication		
		Belumosudil 200 mg QD (n=81)	Belumosudil 200 mg BID (n=75)	BAT*
Prednisolone n (%)	1 mg/kg QOD	77 (95.1%)	73 (97.3%)	95.2%
Tacrolimus n (%)	1 mg BID	28 (34.6%)	28 (37.3%)	34.7%
ECP n (%)	3.2 sessions per 28-day cycle	20 (24.7%)	26 (34.7%)	0%
Sirolimus n (%)	2 mg QD	17 (21.0%)	18 (24.0%)	21.1%
MMF n (%)	1000 mg BID	11 (13.6%)	2 (2.7%)	13.0%
Budesonide n (%)	3 mg TDS	6 (7.4%)	3 (4.0%)	7.2%
Montelukast n (%)	10 mg QD	4 (4.9%)	4 (5.3%)	5.0%
Azithromycin n (%)	250 mg TIW	4 (4.9%)	4 (5.3%)	5.0%

Abbreviations: BAT, best available therapy; BID, twice daily; MMF, mycophenolate mofetil; QD, once daily; QOD, once every other day; TDS, three times per day; TIW, three times per week.

*Weighted average of belumosudil 200 QD and BID based on proportions on each regiment (base case values of 95% for QD regimen and 5% for BID regimen)

**Dosing regimens may not reflect the concomitant treatment usage in the ROCKstar trial.

The company provided a scenario exploring concomitant medications in the model, but considered that even though concomitant medications were not permitted in REACH-3, for "fairness" they should be included in the BAT arm of the model. As such, they estimated a weighted average use of each treatment based on the usage in the belumosudil 200 mg QD and BID regimens. As total drug acquisition costs in the model are estimated based TTD, the company assumed that [redacted] of time on treatment with belumosudil would include ECP treatment. Additional scenarios exploring 100% and 50% time on treatment with ECP relative to belumosudil was also provided by the company. As with BAT, the company included the cost of central line-related infections for 20% of belumosudil patients on concomitant ECP.

Table 44 in the company's clarification response presents the pack size and costs of each treatment per model cycle. The overall cost per model cycle of concomitant medications for belumosudil 200 mg QD, belumosudil BID and BAT was estimated to be £1,071.72, £1,480.52 and £70.04, respectively. The company's concomitant medication scenario, assuming 80% time on ECP, increased the ICER from £3,571 to £16,674 (see Section 5.2.2).

The EAG notes that prednisolone, budesonide and montelukast are corticosteroids and azithromycin is an antibiotic. Based on advice from the EAG's clinical experts, treatment with corticosteroids and tacrolimus is ongoing throughout the treatment of cGvHD. Additionally, the NICE final scope states

that antibiotics are used in the management of cGvHD.[5] Therefore, the EAG considers tacrolimus, corticosteroids and antibiotics are background treatments and would be equally given in all treatment arms. However, as a survival benefit is estimated with belumosudil, the total cost of background treatments will likely exceed that of BAT. As mentioned in Section 4.2.4.5, the EAG considers there is substantial uncertainty with the estimated survival benefit for belumosudil and to limit the decision risk, prefers to remove this assumption. Thus, the EAG considers that removing the cost of background therapies in combination with the removal of a survival benefit with belumosudil is an appropriate scenario.

The EAG considers that the company's reason for inclusion of concomitant medications for the BAT arm of the model for the scenario analysis is not clinically valid. By definition, BAT is a composition of treatments that reflects established clinical management and thus concomitant medications are implicitly part of the basket. Additionally, the usage of concomitant medications for BAT is not based on evidence as it was not permitted in REACH-3. As such, the EAG considers that concomitant medications for BAT should be excluded from the scenario analysis. Furthermore, concomitant medications should be included as part of total costs for belumosudil.

Table 57 presents the EAG's preferred assumptions for concomitant medication use. The EAG notes that the concomitant medications presented in the below table reflects established clinical management outlined in the NICE final scope, with the omission of imatinib and rituximab. The EAG's scenario resulted in a dominant ICER for belumosudil. Detailed results of the scenario are presented in Section 6.2. For the EAG base case, concomitant medications for belumosudil have been included, thus the modelled intervention is now belumosudil+BAT, presented in Section 6.3.

Table 57. Concomitant medication usage – EAG preferred assumptions

Concomitant treatment	Dosing regimen*	Proportion on concomitant medication		
		Belumosudil 200 mg QD (n=81)	Belumosudil 200 mg BID (n=75)	BAT
ECP n (%)	3.2 sessions per 28-day cycle	20 (24.7%)	26 (34.7%)	0%
Sirolimus n (%)	2 mg QD	17 (21.0%)	18 (24.0%)	0%
MMF n (%)	1000 mg BID	11 (13.6%)	2 (2.7%)	0%

Abbreviations: BAT, best available therapy; BID, twice daily; EAG, External Assessment Group; MMF, mycophenolate mofetil; QD, once daily; QOD, once every other day; TDS, three times per day; TIW, three times per week.

*Dosing regimens may not reflect the concomitant treatment usage in the ROCKstar trial.

Time to treatment discontinuation

Based on the September 2022 data cut for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup) for belumosudil, KM data for TTD is nearly mature. At approximately [redacted] for the belumosudil 200 mg QD arm and at approximately [redacted] of patients are on belumosudil 200 mg BID.

As can be seen in Figure 17, the lognormal distribution does not provide a good visual fit to the observed KM data for both belumosudil arms. Notably, for belumosudil 200 mg BID, the lognormal distribution underestimates the KM data between approximately seven to 36 months and between approximately 17 to 34 months for belumosudil 200 mg QD. Additionally, the lognormal overestimates TTD after months 34 and 36 for the QD and BID regimens, respectively. The EAG considers that the poor fit of the lognormal distribution compared with KM data is likely to provide an inaccurate estimation of the drug acquisition costs for belumosudil. The EAG investigated if the other standard distributions provided a better visual fit to the KM, but none were satisfactory.

Given the maturity of the KM TTD data for both belumosudil arms, the EAG considers the observed data can be used directly to estimate belumosudil drug acquisition costs up until the five-year maximum treatment duration. The EAG ran a scenario using the TTD KM data for belumosudil 200 mg QD and BID to estimate drug acquisition costs. As KM data for the belumosudil 200 mg BID arm is only available up to three years, and the KM curves converge for both belumosudil doses at approximately month 40 (Figure 17), the EAG assumed that TTD for the BID regimen would be equal to the QD regimen between years three and five (up to the maximum treatment duration cap). The scenario reduced the ICER from £3,571 to £2,047 (see Section 6.2). Use of KM TTD data for belumosudil is included in the EAG base case presented in Section 6.3.

The EAG considers that company's approach to calibrate a TTD HR for BAT based on median TTD presented in REACH-3 and applied to the TTD curve for belumosudil 200 mg QD is inappropriate. The company's approach assumes there is a relationship between the rate of discontinuation for patients on belumosudil and the treatments included in BAT but the EAG considers that this is a strong assumption that is currently not supported by any evidence.

The EAG explored whether the assumption of treat until failure was appropriate but as most treatments are used off-label, there is not specific guidance around treatment discontinuation for patients with cGvHD. As such, assuming treatment until failure for BAT may overestimate costs.

Nonetheless, the company's scenario exploring treatment until failure for BAT resulted in a dominant ICER for belumosudil.

The company provided a scenario using an exponential distribution to extrapolate median TTD for BAT (presented in Section 5.2.1), which the EAG considers is more appropriate as it removes the link with belumosudil, and extrapolation is based entirely on the BAT data. The EAG acknowledges there is uncertainty around TTD for BAT, but without KM data from REACH-3 use of the exponential distribution based on median TTD is the best approach given the limited data available and has been included in the EAG base case, presented in Section 6.3.

A secondary issue with TTD is around the relative dose intensity (RDI) for patients on treatment. In the company's base case, RDI for both belumosudil and BAT is 100%. However, the EAG's clinical experts considered that patients on treatment and failure free will likely be weaned off treatment over time. In their clarification response, the company considered that treatment weaning for belumosudil can be captured by RDI reported in ROCKstar ([redacted] and [redacted] for belumosudil 200 mg QD and BID, respectively). Equivalent RDI data are unavailable for BAT from REACH-3. The company provided a scenario implementing RDI for belumosudil, which reduced the ICER from £3,571 to £1,598 (see Section 5.2.2 for results). However, the EAG considers the company's base case approach assuming 100% RDI is more appropriate as without RDI data for BAT (where patients are also likely to be tapered off treatment), the scenario is biased in favour of belumosudil.

Accommodation costs for patients receiving ECP

The company included the costs of accommodation for 50% of patients receiving ECP in the BAT arm of the model. The EAG consulted its clinical experts to understand if the NHS funds overnight stays for patients travelling to specialist centres for treatment. The EAG's clinical experts advised that the NHS does not fund overnight stays, but that some specialist centres may have facilities for patients to stay overnight on the premises. The EAG notes that the NHS does provide helps with costs for patients on low incomes (NHS low-income scheme [LIS])[62] but that it does not include help with accommodation costs. As such, the EAG considers accommodation costs should be excluded from the cost-effectiveness analysis and has implemented this assumption in its preferred base case, presented in Section 6.3.

Exclusion of monitoring costs

In the company's original submission, monitoring costs were excluded from the economic model without justification. As such, the EAG requested the company to explain their decision to exclude monitoring costs from the model during the clarification stage. The company explained that monitoring is likely to be the same for all cGvHD patients who are on treatment irrespective of what type of treatment they receive and this view was also echoed by the EAG's clinical experts. As such, even for patients whose failure event is initiation of a new treatment for cGvHD, monitoring would be the same irrespective of whether they failed on belumosudil or BAT. Therefore, the EAG considers the company's approach to exclude monitoring costs from the model is not unreasonable.

Subsequent treatment costs

In the model, subsequent treatments for patients whose failure event was initiation of new systemic cGvHD treatment was assumed to be the same as BAT, regardless of the treatment patients received at third-line. However, the proportion receiving each type of treatment was based on an assumption that 60% of patients on a new systemic cGvHD therapy would receive treatment and this was evenly distributed among the subsequent treatment options (except for rituximab).

The company stated that assumptions around subsequent treatments was made due to a paucity of data for fourth-line treatment and beyond. During the clarification stage, the EAG asked the company to provide data on subsequent treatments given in ROCKstar and KD025-208, but they advised that no details of treatments provided beyond FFS were collected in the trials.

The EAG consulted with its clinical experts who agreed with the composition of subsequent treatments included in the economic model, but considered that the proportions of each type of treatment may not reflect UK clinical practice. As such, during the clarification stage the EAG requested, and the company provided, as scenario exploring the EAG's clinical expert assumptions of subsequent treatment (presented in Table 58).

The scenario resulted in a dominant ICER for belumosudil. However, the company noted that pulsed methylprednisolone was removed as a part of established clinical management in the NICE final scope after consultation as it was rarely used in clinical practice. Additionally, cyclosporine is considered a background therapy, thus the EAG considers that the company's approach to its exclusion as part of subsequent treatment may not be unreasonable and that overall the company's

base case approach to subsequent treatments and proportions of use can be considered conservative.

Table 58. EAG clinical expert subsequent treatment assumptions

Subsequent treatment	Proportion on treatment	
	Company base case	EAG clinical expert assumptions
ECP	14.5%	14.5%
Mycophenolate mofetil	14.5%	14.5%
Sirolimus	14.5%	10%
Rituximab	2.0%	5%
Imatinib	14.5%	10%
Cyclosporine	-	15%
Pulsed methylprednisolone	-	10%
No subsequent treatment	40%	21%

Abbreviations: EAG, External Assessment Group; ECP, extracorporeal photopheresis.

The EAG was concerned with the company's assumption that duration of subsequent treatments was lifetime, especially as for patients in the failure free health state, a maximum treatment duration of five years was assumed. In their clarification response, the company explained that it was their intention to capture the costs of multiple further lines of treatment over a patient's lifetime, thus implementing a proportion of 60% of patients on treatment allows for time off treatment.

The EAG considers that an alternative approach to estimating the duration of subsequent treatments would be to include a maximum subsequent treatment duration, akin to the assumption used in the failure free health state. The EAG ran a scenario exploring maximum subsequent treatment duration of five years. In the model, the company included a scenario where a single cost of subsequent treatments based on mean duration of treatment (in weeks), informed by the company's clinical experts, was applied to incident patients entering the failure state because of initiation of new systemic cGvHD treatment. The proportion of each subsequent treatments reflected the company base case proportions, presented in Table 58. The EAG adjusted the company's scenario to adapt the mean duration of each subsequent treatment to 260 weeks (except for rituximab, which remained as four weeks as per recommended treatment guidelines). The EAG maintained the assumption that 60% of patients would be on subsequent treatment, to account for various changes in treatment over the five years (such as treatment weaning, treatment pauses and treatment discontinuations).

The EAG's five-year maximum subsequent treatment duration scenario increased the ICER from £3,571 to £7,638 (see Section 6.2) and has included it in the EAG base case (see Section 6.3). The EAG acknowledges that the maximum subsequent treatment duration scenario is a simplification of time on subsequent treatment, but considers it to be a more plausible approach than assuming subsequent treatment costs are applied for the remainder of the model time horizon.

5 Cost effectiveness results

5.1 Company's cost effectiveness results

Table 59 presents the cost-effectiveness results of the company's updated (i.e., post clarification) base case deterministic and probabilistic analyses. The company performed probabilistic sensitivity analysis (PSA) to assess the joint parameter uncertainty around base case results. Incremental results from the company's PSA are based on 5,000 simulations.

In the base case probabilistic analysis, an incremental quality-adjusted life-year (QALY) gain of [redacted] over best available therapy (BAT) along with additional costs of [redacted] for belumosudil, generates an incremental cost-effectiveness ratio (ICER) of £3,046 per QALY. Using the £20,000 and £30,000 threshold, the incremental net monetary benefit (INMB) is [redacted] and [redacted] and the incremental net health benefit (INHB) is [redacted] and [redacted]. A positive NHB implies that overall population health would be increased as a result of the new intervention

A proposed confidential patient access scheme (PAS) discount for belumosudil is applied in the company's base case and is therefore reflected in the results presented in this report. A confidential discount is available for rituximab, which is included in BAT. The source of the confidential price for rituximab is the commercial medicines unit (CMU). As such, the External Assessment Group (EAG) has produced a confidential appendix to the EAG report. Analyses included in the confidential appendix include the company base case results, scenario analyses and EAG base case and scenario analyses.

Table 59. Company's updated (post clarification) base case results

Interventions	Total Costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
BAT	248,736	[redacted]	[redacted]	-	-	-	-
Belumosudil	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3,571
Probabilistic results							
BAT	250,314	[redacted]	[redacted]	-	-	-	-
Belumosudil	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3,046

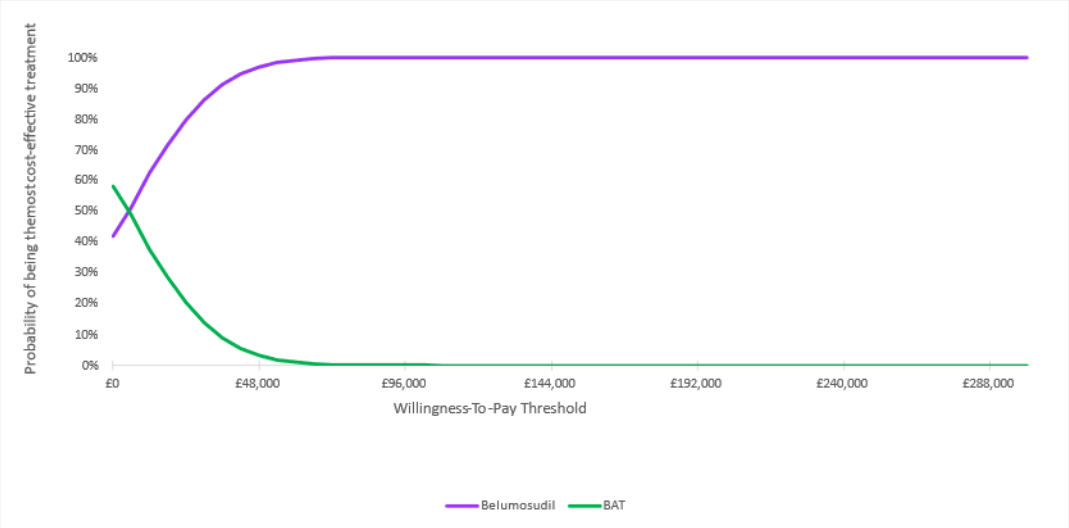
Abbreviations: BAT, best available therapy; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year.

A PSA scatterplot is presented in Figure 18 and a cost-effectiveness acceptability curve (CEAC) is presented in Figure 19. Based on these analyses, the probability that belumosudil is cost effective versus BAT is 86.2% at a willingness to pay (WTP) threshold of £30,000.

The EAG considers the parameters and respective distributions chosen for PSA to be generally sound (see Table 19, Appendix N of the company submission [CS] for PSA inputs). The EAG also considers the probabilistic results to be comparable to the deterministic results.

Figure 18. Cost-effectiveness plane – PSA scatterplot: belumosudil versus best available therapy (reproduced from the company's clarification response, Figure 2)

Figure 19. Cost-effectiveness acceptability curve: belumosudil versus best available therapy (reproduced from the company's clarification response, Figure 3)

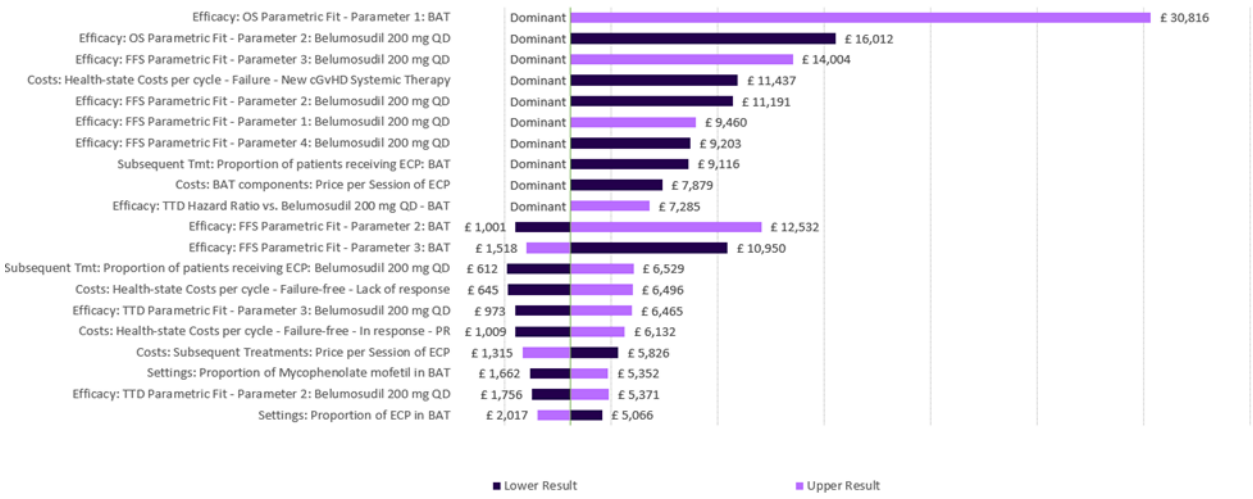


5.2 Company's sensitivity analyses

5.2.1 One-way sensitivity analysis

The company conducted one-way sensitivity analyses (OWSAs) to assess the impact, on the ICER, of varying specific parameters in isolation and to identify the main model drivers. The results are illustrated using the tornado diagram in Figure 20. The ICER was most sensitive to variation in the parameters used to estimate the overall survival (OS) curves for belumosudil 200 mg once daily (QD) and BAT and the failure-free survival (FFS) curves for belumosudil 200 mg QD.

Figure 20. Tornado plot (reproduced from the company's clarification response appendix A, Figure 11)



5.2.2 Scenario analysis

The company undertook a series of scenario analyses to assess the impact of applying alternative assumptions to key model parameters. Details of each scenario are provided in Appendix N.9 of the company submission. In addition, the company conducted several additional scenario analyses requested by the EAG. Results of all the scenario analyses conducted by the company are presented in Table 60. The EAG notes that it was unable to replicate Scenario 5 and the scenario provided in response to clarification B28 and the results provided in the below table are taken from the company's clarification response. Several requested scenarios were not provided by the company, as such the EAG have conducted these additional scenario analyses and provided the results in Section 6.3.

Table 60. Company scenario analyses

	Results per patient	Belumosudil (1)	Best available therapy (2)	Incremental value (1-2)
0	**Company updated base case – post clarification**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,571
1	**Time horizon – 20 years**			
	Total costs (£)	[redacted]	242,869	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	6,436
2	**Time horizon – 5 years**			
	Total costs (£)	[redacted]	155,797	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	53,692
3	**Discount rate – 0%**			
	Total costs (£)	[redacted]	297,234	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
4	**Discount rate – 1.5%**			
	Total costs (£)	[redacted]	273,811	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
5	**Alternative distribution of BAT components***			
	Total costs (£)	N/A	N/A	N/A
	QALYs	N/A	N/A	N/A

	ICER (£/QALY)	-	-	8,827
6	**FFS for all treatments – Joint fit gamma**			
	Total costs (£)	[redacted]	248,581	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	64,208
7	**FFS for belumosudil QD and BID: Joint Fit – Log-normal; FFS for BAT: Joint Fit – Weibull**			
	Total costs (£)	[redacted]	242,898	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	34,893
8	**FFS for belumosudil QD and BID: Independent Fit – Log-normal for QD and Generalised Gamma for BID; FFS for BAT: Independent Fit – Gamma**			
	Total costs (£)	[redacted]	250,197	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	25,163
9	**OS long-term assumption for belumosudil: Do not assume same probability of death as for BAT after 5 years**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	23,754
10	**OS for all treatments: Joint Fit – Log-normal**			
	Total costs (£)	[redacted]	465,948	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
11	**OS for all treatments: Joint Fit – Weibull**			
	Total costs (£)	[redacted]	341,371	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
12	**OS for all treatments: Joint Fit – Gamma**			
	Total costs (£)	[redacted]	317,109	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
13	**OS for belumosudil QD and BID: Independent Fit – Gamma for QD; Log-normal for BID; OS for BAT: Independent Fit – Log-logistic**			
	Total costs (£)	[redacted]	423,014	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
14	**TTD for BAT: exponential curve fitted to median**			
	Total costs (£)	[redacted]	246,376	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	5,174

15	**Treat until failure (all treatments)**			
	Total costs (£)	[redacted]	261,824	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	27,792
16	**Treat until failure (BAT only)**			
	Total costs (£)	[redacted]	261,824	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
17	**Alternate distribution of subsequent treatments (applied for all initial treatments)**			
	Total costs (£)	[redacted]	280,793	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
18	**Alternate approach to costing of subsequent treatments**			
	Total costs (£)	[redacted]	202,925	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	17,158
19	**Maximum duration of treatment for all treatments (except rituximab): 3 years**			
	Total costs (£)	[redacted]	247,430	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
20	**No maximum duration of treatment for all treatments (except rituximab)**			
	Total costs (£)	[redacted]	249,340	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	10,634
21	**Alternate proportions of responders to ECP assumed for drug cost calculations**			
	Total costs (£)	[redacted]	242,580	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	7,742
22	**Disease management costs for all Failure-free health states follow the decrease observed in Schain *et al.* 2021[63]**			
	Total costs (£)	[redacted]	245,607	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
23	**Disease management costs for all Failure-free health states reduced in Years 5+**			
	Total costs (£)	[redacted]	247,459	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
24	**Health state utility for Failure – New cGvHD Systemic Therapy: Crespo *et al.* 2012[20]**			
	Total costs (£)	[redacted]	248,736	[redacted]

	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,703
25	**Value of health state utility for Failure – Recurrent Malignancy and for Failure – New cGvHD Systemic Therapy: upper bound of range**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,586
CQ	**EAG requested scenarios**			
B7	**Belumosudil treatment weaning**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	1,598
B9	**Ongoing risk of recurrent malignancy after 36 months**			
	Total costs (£)	[redacted]	248,640	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,704
B10	**Best response for belumosudil measured up to 24 weeks**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,787
B11a	**Removal of response, maintenance of utility value for failure-free health state (0.741)**			
	Total costs (£)	[redacted]	249,493	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,434
B11b	**Removal of response, maintenance of utility value for failure-free health state (0.741), disease management costs weighted by response**			
	Total costs (£)	[redacted]	249,066	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,237
B11c	**Removal of response, maintenance of utility value for failure-free health state (0.741), disease management costs weighted by response, treatment weaning for belumosudil (B7)**			
	Total costs (£)	[redacted]	249,066	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	1,279
B17	**Alternative utility values for CR ([redacted])**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,564
B18	**Alternative utility values for 'failure – new treatment' estimated from ROCKstar ([redacted])**			

	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	5,186
B20	**Carer disutility for failure – new cGvHD systemic therapy equal to carer disutility for failure-free PR & LR**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,065
B21	**Disutility for IV infusion limited to the day of infusion**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,609
B23.1	**Inclusion of concomitant medications for belumosudil and BAT – ECP usage** [redacted]			
	Total costs (£)	[redacted]	249,373	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	16,674
B23.2	**Inclusion of concomitant medications for belumosudil and BAT – ECP usage 100%**			
	Total costs (£)	[redacted]	249,373	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	19,860
B23.3	**Inclusion of concomitant medications for belumosudil and BAT – ECP usage 50%**			
	Total costs (£)	[redacted]	249,373	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	11,893
B24	**Composition of BAT based on company NICE advisory board report**			
	Total costs (£)	[redacted]	247,392	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,484
B28	**EAG clinical experts' basket of treatments assumption***			
	Total costs (£)	N/A	N/A	[redacted]
	QALYs	N/A	N/A	[redacted]
	ICER (£/QALY)	-	-	Dominant

Abbreviations: BAT, best available therapy; BID, twice daily; cGvHD, chronic graft versus host disease; CQ, clarification question; CR, complete response; EAG, External Assessment Group; ECP, extracorporeal photopheresis; FFS, failure-free survival; ICER, incremental cost effectiveness ratio; IV, intravenous; LR, lack of response; N/A, not available; OS, overall survival; PR, partial response; QALY, quality adjusted life year; QD, once daily; TTD, time to treatment discontinuation.

*EAG unable to replicate scenario. Results taken from the company's clarification response.

5.3 Model validation and face validity check

For the model validation, the company stated that guidelines from the International Society for Pharmacoeconomics and Outcomes Research-Society of Medical Decision Making (ISPOR-SMDM) were followed. An external vendor was used to validate the company's conceptual model, primarily to assess if the model approach reflected the underlying disease course, available evidence and if it addressed the decision problem. In addition, the company set up an advisory board comprised of English clinical and economic experts to discuss details on the target patient population, including characteristics, clinical management, clinical outcomes and resource use. Lastly, the company's external vendor sought validation on model inputs from clinical key opinion leaders.

Quality assurance of the model was performed internal company peer reviewer not involved in the model development.

The EAG considers the company's model validation and face validity check to be robust and has not identified any obvious errors in the model. However, the majority of the company's deterministic scenarios required manual changes to the model, which may result in errors running the scenarios. Nonetheless, with instructions provided by the company during the clarification stage, the EAG was able to replicate most of company's scenario analyses.

6 Additional economic analysis undertaken by the EAG

6.1 Exploratory and sensitivity analyses undertaken by the EAG

In Section 4 of this report, the External Assessment Group (EAG) has described several scenarios that warrant further exploration in addition to the company's own sensitivity and scenario analyses to ascertain the impact of these changes on the incremental cost-effectiveness ratio (ICER). The scenarios that the EAG has performed are as follows:

- Removal of overall survival (OS) benefit – Section 4.2.4.5.
- Concomitant medication costs only for belumosudil – Section 4.2.7.9.
- Concomitant medication costs only for belumosudil only and removal of cost of background therapies – Section 4.2.7.9
- Removal of OS benefit, inclusion of concomitant medications for belumosudil only and removal of cost of background therapies – Section 4.2.7.9.
- Kaplan-Meier (KM) time to treatment discontinuation (TTD) data for belumosudil – Section 4.2.7.9.
- Removal of accommodation costs for patients on extracorporeal photopheresis (ECP) – Section 4.2.7.9.
- Maximum subsequent treatment duration of five years (except for rituximab) – Section 4.2.7.9.
- Utility value for failure – new chronic graft versus host disease (cGvHD) systemic therapy of 0.696 from Crespo *et al*.[20]- Section 4.2.6.5.
- Midpoint utility value of 0.608 for failure new cGvHD systemic therapy – Section 4.2.6.5.
- Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689 – Section 4.2.6.5.
- Removal of intravenous (IV) disutility for best available therapy (BAT) – Section 4.2.6.5.

6.2 EAG scenario analysis

Table 61 presents the results of the EAG exploratory analyses described in Section 6.1. Results reported include the company's proposed patient access scheme (PAS) discount of [redacted]. A confidential discount is available for rituximab, which is included in best available therapy (BAT). The source of the confidential price for rituximab is the commercial medicines unit (CMU). As such, the EAG has produced a confidential appendix to the EAG report. Analyses included in the confidential

appendix include the company base case results, scenario analyses and EAG base case and scenario analyses.

Table 61. Results of the EAG's scenario analyses

	Results per patient	Belumosudil	BAT	Incremental value
0	**Company base case – post clarification**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,571
1	**Removal of OS benefit**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
2	**Concomitant medication costs only for belumosudil**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	17,105
3	**Concomitant medication costs only for belumosudil + removal of cost of background therapies for belumosudil**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	16,716
4	**Scenario 1+3**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
5	**KM TTD data for belumosudil**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	2,047
6	**Removal of accommodation costs for patients on ECP**			
	Total costs (£)	[redacted]	247,984	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,080
7	**Maximum subsequent treatment duration of five years (except for rituximab)**			
	Total costs (£)	[redacted]	238,019	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	7,638

8	**Utility value for failure – new cGvHD systemic therapy of 0.696 from Crespo *et al.*[20]**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,800
9	**Midpoint utility value of 0.608 for failure new cGvHD systemic therapy utility value**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	4,213
10	**Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689[36]**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,568
11	**Removal of IV disutility for BAT**			
	Total costs (£)	[redacted]	248,736	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	3,613

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; KM, Kaplan-Meier; OS, overall survival; QALY, quality adjusted life year; TTD, time to treatment discontinuation.

6.3 EAG preferred assumptions

In this section, the EAG presents its preferred base case for the cost-effectiveness of belumosudil for the treatment of cGvHD after two or more lines of systemic therapy. The EAG notes that one of its key assumptions, which is a primary driver of cost-effectiveness in the model, is the inclusion of concomitant medication costs for belumosudil. By including concomitant medication costs for belumosudil, the intervention in the EAG preferred base case is belumosudil in addition to BAT and hereafter will be referred to as belumosudil+BAT. The EAG highlights that the usage of each treatment in BAT for the belumosudil arm of the model is different to the BAT comparator arm and is based on the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup (August 2021 data cut).

The EAG's preferred base case is a simplified version of the company's base, as the naïve comparison of clinical outcomes for belumosudil+BAT with BAT from REACH-3 introduced uncertainty in the analysis and the direction of bias was problematic to identify. Thus, to aid committee decision making, the EAG preferred to remove the key sources of uncertainty, linked to response outcomes and OS. The assumptions that form the EAG's preferred base case are listed below.

- Removal of response outcomes – company scenario.
- Removal of OS benefit for belumosudil+BAT.
- Concomitant medication costs for belumosudil only.
- Removal of cost of background therapies.
- KM TTD data for belumosudil.
- Exponential distribution for BAT TTD.
- Removal of accommodation costs for patients on ECP.
- Maximum subsequent treatment duration of five years (except for rituximab).
- Midpoint utility value of 0.608 for failure – new cGvHD systemic therapy.
- Caregiver disutility for failure – new cGvHD systemic therapy equal to failure-free (partial response [PR]/lack of response [LR])
- Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689.[36]
- Removal of IV disutility for BAT.

Results of the EAG's preferred base case are presented in Table 62 and detailed results presented in Table 63. The EAG tested the following scenarios around its base case, presented in Table 64:

- Inclusion of the OS benefit for belumosudil+BAT.
- Adelphi disease specific programme (DSP) treatment failure utility value (0.52) for the failure – new cGvHD systemic therapy.
- Utility value of 0.696 from Crespo *et al.*[20] for the failure – new cGvHD systemic therapy.
- Removal of caregiver disutility for all health states.
- Combined scenario of Inclusion of the OS benefit for belumosudil+BAT and Removal of caregiver disutility for all health states.

Table 62. EAG's preferred model assumptions (deterministic) – belumosudil+BAT versus BAT

Preferred assumption	Section in EAG report	Incremental costs	Incremental QALYs	ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case – post clarification	-	[redacted]	[redacted]	3,571	-
Removal of response outcomes – company scenario	4.2.4.5	[redacted]	[redacted]	3,434	£3,434
Removal of OS benefit	4.2.4.5	[redacted]	[redacted]	Dominant	Dominant

Concomitant medication costs for belumosudil only	4.2.7.9	[redacted]	[redacted]	17,105	Dominant
Removal of cost of background therapies*	4.2.7.9	[redacted]	[redacted]	Dominant	Dominant
KM TTD data for belumosudil	4.2.7.9	[redacted]	[redacted]	2,047	Dominant
Exponential distribution for BAT TTD	4.2.7.9	[redacted]	[redacted]	5,174	Dominant
Removal of accommodation costs for patients in ECP	4.2.7.9	[redacted]	[redacted]	4,080	Dominant
Maximum subsequent treatment duration of five years (except for rituximab)	4.2.7.9	[redacted]	[redacted]	7,638	Dominant
Midpoint utility value of 0.608 for failure new cGvHD systemic therapy utility value	4.2.6.5	[redacted]	[redacted]	4,213	Dominant
Caregiver disutility for failure – new cGvHD systemic therapy equal to failure-free (PR/LR)	4.2.6.5	[redacted]	[redacted]	4,065	Dominant
Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689[36]	4.2.6.5	[redacted]	[redacted]	3,568	Dominant
Removal of IV disutility for BAT	4.2.6.5	[redacted]	[redacted]	3,613	Dominant

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; KM, Kaplan-Meier; LR, lack of response; OS, overall survival; PR, partial response; QALY, quality adjusted life year; TTD, time to treatment discontinuation

*Scenario combines the assumptions of removal of OS benefit, concomitant medication costs only for belumosudil and removal of cost of background therapies.

Table 63. EAG's base case results

Interventions	Total Costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
BAT	235,716	[redacted]	[redacted]	-	-	-	-
Belumosudil+ BAT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Probabilistic results							
BAT	236,410	[redacted]	[redacted]	-	-	-	-
Belumosudil+ BAT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: BAT, best available therapy; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year.

Table 64. Deterministic scenario analyses around the EAG base case

	Results per patient	Belumosudil+BAT	BAT	Incremental value
0	**EAG base case**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
1	**Inclusion of OS benefit**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	28,449
2	**Adelphi DSP treatment failure utility value (0.52)[49] for the failure – new cGvHD systemic therapy**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
3	**Utility value of 0.696 from Crespo *et al* [20]for the failure – new cGvHD systemic therapy**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
4	**Removal of caregiver disutility**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
5	**Scenario 1 + 4**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	26,749

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; DSP, disease specific programme; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; OS, overall survival; QALY, quality adjusted life year.

6.4 Conclusions of the cost effectiveness sections

Generally, the EAG considers the company's submitted cost-effectiveness analysis adheres to the decision problem defined in the NICE final scope.[5] However, one key departure from the NICE final scope, was the company's approach to model belumosudil as a monotherapy, when in ROCKstar, KD025-208 and in the company submission, belumosudil is considered as an add-on treatment to a patient's existing treatment package (BAT). Thus, the EAG considers that the intervention in the model should be belumosudil+BAT, which aligns with the NICE final scope.[5]

The fundamental issue with the company's cost-effectiveness analysis is the naïve comparison of clinical outcomes for belumosudil+BAT versus BAT. As there is no head-to-head comparative trial, an assessment of belumosudil and BAT can only be conducted with an indirect treatment comparison. Furthermore, the company's submission relies on clinical expert opinion of the differences between ROCKstar, KD025-208 and REACH-3 and the resulting direction of bias with regards to the treatment effect. Therefore, the EAG considers that there is a substantial amount of uncertainty related to treatment effectiveness in the model due to the naïve comparison of belumosudil+BAT and BAT.

The EAG considers that it is appropriate to simplify the approach to the cost-effectiveness analysis by removing key uncertainty associated with response outcomes and OS benefit in the model. However, the inclusion of an OS benefit for belumosudil+BAT in the model has a substantial impact on the ICER. Therefore, the EAG recommends that the committee obtain advice from its clinical experts on the clinical plausibility of an OS benefit for belumosudil+BAT.

The EAG notes that issues around preferred alternative assumptions around costs and utility values are also important for the committee to consider, but the impact of these assumptions become secondary when the assumption of no OS benefit is employed in the model (ICER becomes dominant). However, the choice of preferred utility values informs the application of the severity modifier, discussed in Section 7. The company's base case is associated with a severity modifier of 1.2, but the EAG's preferred base case estimated a severity modifier of 1. However, the EAG notes that the both the company's and EAG's base case ICERs are well below the NICE cost-effectiveness threshold of £20,000 per QALY.

7 Severity modifier

As outlined in the National Institute for Health and Care Excellence (NICE) methods guide,[31] "*the committee will consider the severity of the condition, defined as the future health lost by people living with the condition with standard care in the NHS*". The thresholds of quality-adjusted life-year (QALY) weightings for severity are reported in Table 65.

Table 65. QALY weighting for severity

QALY weight	Proportional QALY shortfall	Absolute QALY shortfall
1	Less than 0.85	Less than 12
x1.2	0.85 to 0.95	12 to 18
x1.7	At least 0.95	At least 18.

Abbreviations: QALY, quality-adjusted life-year

The company calculated the absolute and proportional QALY shortfall using a published calculator by Schneider *et al*. coded in their economic model.[64] The tool calculates the expected total QALYs for the general population matched to baseline age and sex distribution included in the economic model. The source of the general population EQ-5D data used in the calculator is from a study by Hernandez *et al*. 2020.[65] Table 66 presents the company's preferred assumptions for the general population QALY shortfall estimates.

Table 66. Summary of preferred assumptions for general population QALY shortfall estimates

Factor	Value or source	Reference to section in submission or rationale
Sex distribution - male	58.0%	Post clarification economic model – Pooled data for the ≥2 LOT subgroup of ROCKstar and KD025-208 (September 2022 data cut).
Starting age (mean)	53.9 years	Post clarification economic model – Pooled data for the ≥2 LOT subgroup of ROCKstar and KD025-208 (September 2022 data cut).
Expected total QALYs for the general population	14.613	Schneider *et al*. 2021.[64] Estimated based on starting age and sex distribution at baseline
Discount rate	3.5%	NICE reference case[31]

Abbreviations: LOT, lines of therapy; QALY, quality-adjusted life-year

To calculate the absolute and proportional QALY shortfall using the calculator, the company used the base case total QALYs estimated for the best available therapy (BAT) arm, estimated to be [redacted]. The results of the company's QALY shortfall analysis is presented in Table 67 and Table 68 presents a summary of health state benefits and utility values for the QALY shortfall analysis. Table 69 presents a summary of the company's preferred assumptions for the BAT QALY shortfall estimates.

Table 67. Summary of QALY shortfall analysis

Expected total QALYs for the general population	Expected total QALYs that people living with a condition would be expected to have with BAT	QALY shortfall	
		Proportional shortfall	Absolute shortfall
14.61	[redacted]	[redacted]	[redacted]

Abbreviations: BAT, best available therapy; QALY, quality-adjusted life-year.

Table 68. Summary of health state benefits and utility values for QALY shortfall analysis

State	Discounted QALYs
Failure-free – 'in response' (complete/partial response)	[redacted]
Failure-free – lack of response	[redacted]
Failure – New cGvHD Systemic Therapy	[redacted]
Failure – Recurrent Malignancy	[redacted]
One-off AE-related QALY loss	[redacted]
Disutility associated with IV infusion	[redacted]
Caregiver disutility – failure-free (PR/LR)	[redacted]
Caregiver disutility - failure	

Abbreviations: AE, adverse events; cGvHD, chronic graft versus host disease; IV, intravenous; LR, lack of response; PR, partial response; QALY, quality-adjusted life-year.

Table 69. Summary of company preferred assumptions for BAT QALY shortfall estimates

Modelled input	Assumption or value (reference to appropriate table or figure in submission)	Rationale or justification
'In response' curve for BAT	Section B.3.3.3 to B.3.3.5 of the CS. Inclusion of response data, time to response and duration of response to calculate and 'in response' curve used to estimate QALYs in the failure-free health state	Company considers that inclusion of response in the model is important to estimate costs and QALYs.
TTD curve for BAT	Section B.3.3.6.2 of the CS. Calibrated HR based on median TTD for BAT from REACH-3, applied to belumosudil QD TTD curve. Choice of TTD approach has an impact on the estimation of total QALY loss related to IV infusions.	Calibrated HR applied to belumosudil QD curve estimates the median BAT TTD from REACH-3
Failure-free – 'in response' (complete/partial response) utility value	Section B.3.4.3 of the company submission. [redacted]	Based on mapped PROMIS-GH data from ROCKstar.
Failure-free – lack of response utility value	Section B.3.4.3 of the company submission. [redacted]	Based on mapped PROMIS-GH data from ROCKstar.
Failure – New cGvHD Systemic Therapy utility value	Section B.3.4.3 of the company submission. 0.479 (0.036)	Assumed to be the same as failure – recurrent malignancy due to lack of published data.
Failure – Recurrent Malignancy utility value	Section B.3.4.3 of the company submission. 0.479 (0.036)	Based on a weighted average of progression utilities for AML, ALL, CLL, CML from TA642 (based on Joshi *et al.*), Aristides *et al.*, TA813 (derived from TA451) and Beusterien *et al.*[40-42, 44, 45]
One-off AE-related QALY loss	Section B.3.4.3 of the company submission. -0.002	Weighted average disutility based on published disutility values for Grade ≥3 AEs occurring in at least 5% of patients in the pooled analysis of ROCKstar and KD025-208 or REACH-3.
Disutility associated with IV infusion	Section B.3.4.3 of the company submission. -0.037 (0.010)	Inclusion of IV infusion disutility based on clinical expert opinion obtained by the company.
Caregiver disutility – failure-free (PR/LR)	Section B.3.4.3 of the company submission. -0.045 (0.057)	Inclusion of caregiver disutility based on clinical expert opinion obtained by the company which suggested high caregiver burden for patients with cGvHD who are on their third line of systemic treatment.
Caregiver disutility - failure	Section B.3.4.3 of the company submission. -0.142 (0.062)	

Abbreviations: AE, adverse event; ALL, acute lymphoblastic leukaemia; AML, acute myelogenous leukaemia; cGvHD, chronic graft versus host disease; CLL, chronic lymphocytic leukaemia; CML, chronic myelogenous leukaemia; CS, company submission; HR, hazard ratio; IV, intravenous; LR, lack of response; NICE, National Institute for Health and Care Excellence; PR, partial response; QALY, quality-adjusted life-year; QD, once daily; TA, technology appraisal; TTD, time to treatment discontinuation.

7.1 Cost-effectiveness estimates

Based on the QALY shortfall analysis, the company estimated that a severity modifier of 1.2 should be considered by the committee. Table 70 presents the company's and the External Assessment Group's (EAG's) preferred cost-effectiveness results without the severity weighting. Table 71 presents the company's preferred cost-effectiveness results with the severity modifier of 1.2 applied to the incremental QALYs. However, the severity modifier of 1.2 does not apply to the EAG's preferred cost effectiveness results, as the absolute QALY shortfall is less than 12 and the proportional QALY shortfall is less than 0.85 (see Section 7.2).

Table 70. Cost-effectiveness results without severity weighting

Scenario	Incremental costs (£)	Incremental QALYs	ICER (£/QALYs)
Company base case	[redacted]	[redacted]	3,571
EAG base case	[redacted]	[redacted]	Dominant

Abbreviations: EAG, external assessment group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 71. Cost-effectiveness results with severity weighting (x1.2)

Scenario	Incremental costs (£)	Incremental QALYs (x1.2)	ICER (£/QALYs)
Company base case	[redacted]	[redacted]	[redacted]

Abbreviations: EAG, external assessment group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 72. Cost-effectiveness results with severity weighting (x1.2) presented using NHB

Scenario	ICER (£/QALYs)	Incremental NHB £20,000 threshold	Incremental NHB £30,000 threshold
Company base case	[redacted]	[redacted]	[redacted]

Abbreviations: EAG, external assessment group; ICER, incremental cost-effectiveness ratio; NHB, net health benefit; QALY, quality adjusted life year

7.2 EAG critique

The EAG considers the Schneider *et al*. calculator an appropriate tool to estimate absolute and proportional QALYs. The key assumptions preferred by the EAG are presented in Table 73 and application of these in the EAG base case results in an absolute QALY shortfall of [redacted] and a proportional QALY shortfall of [redacted], resulting in a severity modifier of 1.

Table 73. Summary of EAG preferred assumptions for BAT QALY shortfall estimates

Modelled input	EAG preferred value/ assumption	EAG rationale
'In response' curve for BAT	Excluded.	Section 4.2.4.5 – The EAG prefers to exclude response from the model. Response in the model is not a primary driver of cost-effectiveness. Changes to the 'in response' curves and response data have limited impact on the ICER. Inclusion of response in the model to add granularity to QALYs and costs in the model is potentially adding unnecessary complexity to the analysis. Additionally, given the naïve comparison of belumosudil and BAT, inclusion of response in the model is another source of uncertainty.
TTD curve for BAT	Exponential distribution to extrapolate median TTD for BAT.	Section 4.2.7.9 - The company provided a scenario using an exponential distribution to extrapolate median TTD for BAT which the EAG considers is more appropriate that the company's base case approach as it removes the link with belumosudil, and extrapolation is based entirely on the BAT data.
Failure-free utility value	[redacted]	Section 4.2.6 – As response is removed from the model, a single utility value for the failure-free health state is appropriate. Utility value based on mapped PROMIS-GH data from ROCKstar.
Failure – New cGvHD Systemic Therapy utility value	0.608	Section 4.2.6.5 - Utility value for failure – new cGvHD systemic therapy is key driver of QALYs in the model, as patients in the BAT arm spend the majority of their time in this health state. The EAG prefers to use an estimated midpoint utility value based on the company's Adelphi DSP study and a publication by Crespo *et al*.[20]
Inclusion of disutility value for central line-related infections based on TA689.	-0.22, 14 days	Section 4.2.6.5 - The EAG was concerned that an assumption of no disutility for central line-related infections was inconsistent with the high-cost nature of treating this AE. Instead, the EAG prefers the disutility value and duration associated with infections and infestations from TA689.
Disutility associated with IV infusion	Excluded.	Section 4.2.6.5 - In the company's concomitant medication scenario, they assumed that the impact of IV infusions is captured as part of the trial-based utility values used in the failure-free health state. The EAG agrees with the company's approach to exclude IV infusion disutility for belumosudil but considers that IV infusion disutility should be removed for all treatment arms as the utility values used in the model are not treatment specific.
Caregiver disutility – failure – initiation of a new systemic cGvHD therapy	-0.045	Section 4.2.6.5 - Assumed to be the same as caregiver disutility – failure-free (PR/LR). The EAG's clinical experts advised that impact on caregiver HRQoL would not be the same for patients who initiate a new systemic cGvHD therapy compared with patients who have a

		recurrent malignancy. Instead, the EAG's clinical experts advised that it might be reasonable to assume the caregiver impact for patients who initiate a new systemic cGvHD therapy would be akin to patients who have had a partial or lack of response.

Abbreviations: AE, adverse event; ALL, acute lymphoblastic leukaemia; AML, acute myelogenous leukaemia; cGvHD, chronic graft versus host disease; CLL, chronic lymphocytic leukaemia; CML, chronic myelogenous leukaemia; CS, company submission; EAG, external assessment group; HR, hazard ratio; IV, intravenous; LR, lack of response; NICE, National Institute for Health and Care Excellence; PR, partial response; QALY, quality-adjusted life-year; QD, once daily; TA, technology appraisal; TTD, time to treatment discontinuation.

8 References

1. Electronic Medicines Compendium (EMC). REZUROCK 200 mg film-coated tablets. Summary of Product Characteristics (SmPC), 2022. Available from: https://www.medicines.org.uk/emc/product/14659/smpc#companyDetails. Date accessed: April 2023.
2. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. *Biol Blood Marrow Transplant* 2015; **21**: 389-401 e1.
3. Dignan FL, Amrolia P, Clark A, Cornish J, Jackson G, Mahendra P, et al. Diagnosis and management of chronic graft-versus-host disease. *Br J Haematol* 2012; **158**: 46-61.
4. Specialised Commissioning Team. Clinical Commissioning Policy: Treatments for Graft versus Host Disease (GvHD) following Haematopoietic Stem Cell Transplantation. UK: 2017 NHS England: 16069 /P.
5. National Institute for Health and Care Excellence (NICE). Final scope: Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021], 2023. Available from: https://www.nice.org.uk/guidance/gid-ta11136/documents/final-scope. Date accessed: April 2023.
6. Zanin-Zhorov A, Blazar BR. ROCK2, a critical regulator of immune modulation and fibrosis has emerged as a therapeutic target in chronic graft-versus-host disease. *Clin Immunol* 2021; **230**: 108823.
7. Kadmon Corporation. Clinical study report: A Phase 2, Randomized, Multicenter Study to Evaluate the Efficacy and Safety of KD025 in Subjects With Chronic Graft Versus Host Disease (cGVHD) After at Least 2 Prior Lines of Systemic Therapy (The ROCKstar Study). 2020.
8. Kadmon Corporation. Clinical study report: A Phase 2a, Dose-Escalation, Open-Label Study to Evaluate the
Safety, Tolerability, and Activity of KD025 in Subjects With Chronic Graft Versus Host Disease. 2020.
9. National Institue for Health and Care Excellence (NICE). Ruxolitinib for treating chronic graft versus host disease refractory to corticosteroids (terminated appraisal), 2022. Available from: https://www.nice.org.uk/guidance/ta840. Date accessed: April 2023.
10. Zeiser R, Polverelli N, Ram R, Hashmi SK, Chakraverty R, Middeke JM, et al. Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. *The New England journal of medicine* 2021; **385**: 228-38.
11. Sanofi. Belumosudil Virtual Advisory Board Meeting Report. 2023.
12. Lee SJ, Wolff D, Kitko C, Koreth J, Inamoto Y, Jagasia M, et al. Measuring therapeutic response in chronic graft-versus-host disease. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: IV. The 2014 Response Criteria Working Group report. *Biol Blood Marrow Transplant* 2015; **21**: 984-99.
13. Thompson NR, Lapin BR, Katzan IL. Mapping PROMIS Global Health Items to EuroQol (EQ-5D) Utility Scores Using Linear and Equipercentile Equating. *Pharmacoeconomics* 2017; **35**: 1167-76.
14. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.
15. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions. 2019. Available from: https://training.cochrane.org/handbook/current. Date accessed: 23 July 2020.
16. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366**: l4898.
17. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998; **52**: 377-84.

18. Lee SJ, Onstad L, Chow EJ, Shaw BE, Jim HSL, Syrjala KL, et al. Patient-reported outcomes and health status associated with chronic graft-versus-host disease. *Haematologica* 2018; **103**: 1535-41.
19. CADTH. CADTH reimbursement review: Ruxolitinib (Jakavi). *Canadian Journal of Health Technologies* 2022; **2022**: 1-220.
20. Crespo C, Pérez-Simón JA, Rodríguez JM, Sierra J, Brosa M. Development of a population-based cost-effectiveness model of chronic graft-versus-host disease in Spain. *Clin Ther* 2012; **34**: 1774-87.
21. de Waure C, Capri S, Veneziano MA, Specchia ML, Cadeddu C, Di Nardo F, et al. Extracorporeal Photopheresis for Second-Line Treatment of Chronic Graft-versus-Host Diseases: Results from a Health Technology Assessment in Italy. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2015; **18**: 457-66.
22. HAS. Jakavi (ruxolitinib) - Acute (GvHD) and chronic (GvHD) graft versus host disease when there is an inadequate response to corticosteroids or other therapies. *Assess HEALTH TECHNOLOGIES* 2022.
23. Okumura L, Holanda P, Riveros BS, Lucchetta RC, Rosim MP, Dierick K, et al. PRO4 Cost-effectiveness analysis of an integrated extracorporeal photopheresis system versus selected treatments for refractory chronic graft versus host disease. *Value in Health Regional Issues* 2019; **19**: S74.
24. Peacock ADF, Taylor C. P615 Cost-effectiveness of extracorporeal photopheresis for the treatment of chronic graft versus host disease in the australian setting. The 48th Annual Meeting of the European Society for Blood and Marrow Transplantation: Patient Advocacy – Patient Advocacy – Poster Session (P613-P616). *Bone Marrow Transplantation* 2022; **57**: 468-70.
25. Yalniz FF, Murad MH, Lee SJ, Pavletic SZ, Khera N, Shah ND, et al. Steroid Refractory Chronic Graft-Versus-Host Disease: Cost-Effectiveness Analysis. *Biol Blood Marrow Transplant* 2018; **24**: 1920-7.
26. BNF. British National Formulary. Available from: https://bnf.nice.org.uk. Date accessed.
27. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT), 2021. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. Date accessed: April 2023.
28. Jones K, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A, et al. Unit Costs of Health and Social Care 2022
Manual. 2022. Available from: https://kar.kent.ac.uk/100519/1/Unit_Costs_of_Health_and_Social_Care_2022%20%287%29.pdf. Date accessed: May 2023.
29. NHS Business Services Authority (NHSBSA). Electronic Drug Tariff Database. 2023. Available from: https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff. Date accessed.
30. NHS England. National Schedule of NHS Costs 2020/21, 2022. Available from: https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/. Date accessed: April 2022.
31. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. Process and methods [PMG36]. 2022. Available from: https://www.nice.org.uk/process/pmg36/ Date accessed: May 2023.
32. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. NICE Decision Support Unit, 2011.
33. Office for National Statistics (ONS). National life tables: UK, 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables. Date accessed: April 2023.

34. Bashey A, Zhang X, Morris LE, Holland HK, Solh M, Solomon SR. Improved Survival of Patients Diagnosed with Severe (Grade 3-4) Acute GVHD or Severe NIH Grade Chronic GVHD in the Current Era Compared to Historic Controls. *Blood* 2019; **134**: 2006-.
35. National Institute for Health and Care Excellence. Idelalisib for treating chronic lymphocytic leukaemia (TA359), 2015. Available from: https://www.nice.org.uk/guidance/ta359. Date accessed: April 2023.
36. National Institute for Health and Care Excellence. Acalabrutinib for treating chronic lymphocytic leukaemia (TA689), 2021. Available from: https://www.nice.org.uk/guidance/ta689. Date accessed: April 2023.
37. National Institute for Health and Care Excellence. Oral azacitidine for maintenance treatment of acute myeloid leukaemia after induction therapy (TA827), 2022. Available from: https://www.nice.org.uk/guidance/ta827. Date accessed: April 2023.
38. Matza LS, Cong Z, Chung K, Stopeck A, Tonkin K, Brown J, et al. Utilities associated with subcutaneous injections and intravenous infusions for treatment of patients with bone metastases. *Patient Prefer Adherence* 2013; **7**: 855-65.
39. Acaster S, Perard R, Chauhan D, Lloyd AJ. A forgotten aspect of the NICE reference case: an observational study of the health related quality of life impact on caregivers of people with multiple sclerosis. *BMC Health Serv Res* 2013; **13**: 346.
40. Joshi N, Hensen M, Patel S, Xu W, Lasch K, Stolk E. Health State Utilities for Acute Myeloid Leukaemia: A Time Trade-off Study. *PharmacoEconomics* 2019; **37**: 85-92.
41. National Institute for Health and Care Excellence. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia (TA642), 2020. Available from: https://www.nice.org.uk/guidance/ta642. Date accessed: April 2023.
42. Aristides M, Barlev A, Barber B, Gijsen M, Quinn C. Population preference values for health states in relapsed or refractory B-precursor acute lymphoblastic leukemia in the United Kingdom. *Health and quality of life outcomes* 2015; **13**: 181-.
43. National Institute for Health and Care Excellence. Ponatinib for treating chronic myeloid leukaemia and acute lymphoblastic leukaemia (TA451), 2017. Available from: https://www.nice.org.uk/guidance/ta451. Date accessed: April 2023.
44. National Institute for Health and Care Excellence. Asciminib for treating chronic myeloid leukaemia after 2 or more tyrosine kinase inhibitors (TA813), 2022. Available from: https://www.nice.org.uk/guidance/ta813. Date accessed: April 2023.
45. Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health and quality of life outcomes* 2010; **8**: 50-.
46. Ara R, Wailoo AJ. NICE DSU Technical Support Document 12: The use of health state utility values in decision models., 2011. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD12-Utilities-in-modelling-FINAL.pdf. Date accessed: April 2023.
47. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report. 2022.
48. Sanofi. GVHD Disease Specific Programmes (DSP) Analysis Slide Deck (data on file). 2022.
49. Sanofi. Disease Specific Programmes (DSP) Analysis Tables (data on file). 2022.
50. Pidala J, Kurland, B.,Chai, X.,Majhail, N.,Weisdorf, D. J.,Pavletic, S.,Cutler, C.,Jacobsohn, D.,Palmer, J.,Arai, S.,Jagasia, M.Lee, S. J. Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. *Blood* 2011; **117**: 4651-7.
51. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care. 2021.
52. Button PM, Child F, Mootien M, Saglam S, Magsino J, Dhanoia H. Service improvement in a nurse-led clinic for extracorporeal photopheresis. *Br J Nurs* 2021; **30**: S10-s4.

53. Martin PJ, Storer BE, Rowley SD, Flowers ME, Lee SJ, Carpenter PA, et al. Evaluation of mycophenolate mofetil for initial treatment of chronic graft-versus-host disease. *Blood* 2009; **113**: 5074-82.
54. Olivieri A, Locatelli F, Zecca M, Sanna A, Cimminiello M, Raimondi R, et al. Imatinib for refractory chronic graft-versus-host disease with fibrotic features. *Blood* 2009; **114**: 709-18.
55. Couriel DR, Saliba R, Escalon MP, Hsu Y, Ghosh S, Ippoliti C, et al. Sirolimus in combination with tacrolimus and corticosteroids for the treatment of resistant chronic graft-versus-host disease. *Br J Haematol* 2005; **130**: 409-17.
56. Cutler C, Miklos D, Kim HT, Treister N, Woo SB, Bienfang D, et al. Rituximab for steroid-refractory chronic graft-versus-host disease. *Blood* 2006; **108**: 756-62.
57. National Institute for Health and Care Excellence. Lenalidomide with rituximab for previously treated follicular lymphoma (TA627), 2020. Available from: https://www.nice.org.uk/guidance/ta627. Date accessed: April 2023.
58. National Institute for Health and Care Excellence. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies (TA559), 2019. Available from: https://www.nice.org.uk/guidance/ta559. Date accessed: April 2023.
59. Sanofi. Data on File. HES Report - The clinical and economic burden associated with chronic graft-versus-host disease following allogeneic haematopoietic stem cell transplant in England; Version No.: 01.00. 2022.
60. Wang HI. Cancer treatment pathways: AML a detailed costing study example. 2011 [PhD]: University of York; 2011.
61. Manoukian S, Stewart S, Graves N, Mason H, Robertson C, Kennedy S, et al. Bed-days and costs associated with the inpatient burden of healthcare-associated infection in the UK. *J Hosp Infect* 2021; **114**: 43-50.
62. Service NH. NHS Low Income Scheme (LIS). 2023. Available from: https://www.nhs.uk/nhs-services/help-with-health-costs/nhs-low-income-scheme-lis/. Date accessed: May 2023.
63. Schain F, Batyrbekova N, Liwing J, Baculea S, Webb T, Remberger M, et al. Real-world study of direct medical and indirect costs and time spent in healthcare in patients with chronic graft versus host disease. *Eur J Health Econ* 2021; **22**: 169-80.
64. Schneider P, McNamara S, Love-Koh J, Doran T, Gutacker N. Quality-adjusted life expectancy norms for the English population. 2021.
65. Hernandez Alava M, Pudney S, Wailoo A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study. 2020.
66. Le RQ, Wang X, Zhang H, Li H, Przepiorka D, Vallejo J, et al. FDA Approval Summary: Ruxolitinib for Treatment of Chronic Graft-Versus-Host Disease after Failure of One or Two Lines of Systemic Therapy. *The oncologist* 2022; **27**: 493-500.

9 Appendices

9.1 Distribution of Failure events

Table 74. Distribution of failure events in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and REACH-3

Period	Belumosudil 200 mg QD			Belumosudil 200 mg BID			BAT		
	New cGvHD Systemic Therapy	Recurrent Malignancy	Non-Relapse Mortality	New cGvHD Systemic Therapy	Recurrent Malignancy	Non-Relapse Mortality	New cGvHD Systemic Therapy	Recurrent Malignancy	Non-Relapse Mortality
0-3 months	49.54%	20.18%	30.28%	92.26%	0.00%	7.74%	87.34%	4.63%	8.03%
3-6 months	85.11%	14.89%	0.00%	100.00%	0.00%	0.00%	87.34%	4.63%	8.03%
6-9 months	80.56%	10.19%	9.26%	85.71%	0.00%	14.29%	65.52%	14.28%	20.20%
9-12 months	83.08%	16.92%	0.00%	56.63%	0.00%	43.37%	65.52%	14.28%	20.20%
12-18 months	100.00%	0.00%	0.00%	62.77%	0.00%	37.23%	100.00%	0.00%	0.00%
18-24 months	83.33%	16.67%	0.00%	75.00%	0.00%	25.00%	0.00%	0.00%	100.00%
24-30 months	66.67%	0.00%	33.33%	100.00%	0.00%	0.00%	100.00%	0.00%	0.00%
30-36 months	50.00%	0.00%	50.00%	100.00%	0.00%	0.00%	100.00%	0.00%	0.00%
36+ months	100.00%	0.00%	0.00%	100.00%	0.00%	0.00%	100.00%	0.00%	0.00%
Source	Pooled ROCKstar (mITT population) and KD025-208 (mITT population on 200 mg QD or 200 mg BID and with ≥2 prior lines of systemic cGvHD therapy) *post-hoc* analyses – September 2022 data cut						Zeiser *et al.* 2021[10]		

Abbreviations: BAT, best available therapy; BID, twice daily; cGvHD, chronic graft versus host disease; mITT, modified intention-to-treat; QD, once daily.

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG 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 Tuesday 13 June 2023** 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 Concomitant medication data

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 21: "The company provided the breakdown of concomitant medications for the treatment of cGvHD used for at least five trial subjects in ROCKstar only based on the 2021 data cut" The data on concomitant medication use in the company's model (and presented in the company's response to clarification questions) were actually based on the pooled analysis of ROCKstar and KD025-208 (August 2021 data cut).	"The company provided the breakdown of concomitant medications for the treatment of cGvHD used for at least five trial subjects in the pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup based on the 2021 data cut"	Factual accuracy. We would like to acknowledge that this was mistakenly described as being based on ROCKstar alone in two places in the company's response to clarification questions (Question B23, page 67 ["In lieu of this, we have tabulated the therapies recorded in the ROCKstar study (2021 data cut)"] and page 69 [in the caption for Table 44]) and therefore in the EAG's report.	Thank you for clarifying the data analysis used. This has been updated throughout the EAG report.
Please note that this issue also affects the following places in the EAG report: • Page 116: "Concomitant medications in ROCKstar included ECP, which in the BAT arm of the model incurs a disutility of -0.037 related to treatment with IV infusion."			

- Page 127: “Nonetheless, in their clarification response to question B23, the company provided the breakdown of concomitant medications for the treatment of cGvHD used for at least five trial subjects in ROCKstar only based on the 2021 data cut.”
- Page 127: “Table 56 outlines the concomitant medications for the treatment of cGvHD used for at least five trial subjects reported in ROCKstar.”
- Page 127: caption for Table 56 (“Table 56. Concomitant medication use based on ROCKstar (2021 data cut)”)

Page 145: “The EAG highlights that the usage of each treatment in BAT for the belumosudil arm of the model is different to the BAT comparator arm and is based on data from ROCKstar.”

Issue 2 Incorrectly named clinical trial

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**EAG response**
Page 24: “The three trials primarily included in the company’s submission, ROCKstar, KD052-208...” KD025-208 is inaccurately named. Please note that this issue also affects the following places in the EAG report: p36 (x2), p38 (x2), p40, p41, p47, p70 (x2), and p78	“The three trials primarily included in the company’s submission, ROCKstar, KD025-208...”	Typographical error	Thank you for highlighting this error. The EAG report has been corrected throughout.

Issue 3 Analysis including concomitant medication costs

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 21: "The company's ICER post clarification changed from £3,571 to £17,085." In this analysis, the EAG proposed to include concomitant medication costs (for belumosudil only) but excluding the costs of concomitant tacrolimus, corticosteroids, and antibiotics. However, it appears that in the model, the EAG only excluded the costs of concomitant tacrolimus, corticosteroids, and antibiotics for belumosudil 200 mg BID. These should have also been excluded for belumosudil 200 mg QD. Therefore, we believe the correct ICER for this	"The company's ICER post clarification changed from £3,571 to £16,716."	Correction of results	Thank you for highlighting the error in the scenario. This has been corrected in the EAG versions of the model and throughout the EAG report.

scenario should have been £16,716 instead of £17,085.			

Please note that this issue also affects the following places in the EAG report:

- Table 7 (page 26): affects the results of the 'Removal of cost of background therapies' scenario and of all scenarios further down the table
- Table 8 (pages 26-27): affects the results of the EAG base case and of all scenarios in this table
- Table 61 (page 144): this would affect the results of Scenario 3 ('Concomitant medication costs only for belumosudil + removal of cost of background therapies for belumosudil') and of Scenario 4 ('Scenario 1+3'). However, as noted under Issue 53 and Issue 54, the values currently displayed for Scenarios 3 and 4 are not the intended ones.
- Table 62 (page 147): affects the results of the 'Removal of cost of background therapies' scenario. (Note: The cumulative ICER column for all scenarios further down the table would also be affected if the ICER was numerical. However, as the cumulative ICER is 'Dominant' for all these scenarios, no correction is needed for those.)
- Table 63 (page 147): affects both deterministic and probabilistic results in this table
- Table 64 (pages 147-148): affects the results of the EAG base case and of all scenarios in this table

Table 70 (page 153): affects results of the EAG base case

Issue 4 Analysis without response

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 23: "The company's scenario removing response has limited impact	"The company's scenario removing response has limited impact on the ICER, reducing it from £3,571 to £3,434."	Correction of typographical error	Thank you for highlighting this error. The EAG report has been corrected.

on the ICER, reducing it from £3,517 to £3,434." The value of the base-case ICER in the company's model was £3,571, not £3,517.			
Please note that this issue also affects the following place in the EAG report: Page 100: "The company's scenario removing response has limited impact on the ICER, reducing it from £3,517 to £3,434."			

Issue 5 Disutility associated with IV infusions

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 25: "However, in their scenario exploring concomitant medications for belumosudil, the company considered the utility value for the failure health state included impact of IV infusions and thus removed the disutility for belumosudil only – Section 4.2.6." The consideration here related to the failure-free	"However, in their scenario exploring concomitant medications for belumosudil, the company considered the utility value for the failure-free health state included impact of IV infusions and thus removed the disutility for belumosudil only – Section 4.2.6."	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

health state, not the failure state.			

Issue 6 Belumosudil place in therapy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 29: "This places corticosteroids at 1L, CNIs at 2L, and other relevant treatments at 3L (Figure 4, CS). The company positioned treatment with belumosudil at 3L in their treatment pathway alongside ECP, rituximab, MMF, sirolimus, and imatinib (Figure 1, below). In the company's pathway, patients are treated with corticosteroids + CNIs as their first systemic treatment for cGvHD and so would have received 1L and 2L concurrently." The CS states on p27: "The optimal place in therapy for	"This places corticosteroids at 1L, CNIs at 1L or 2L, and other relevant treatments at 2L or 3L (Figure 4, CS). The company positioned treatment with belumosudil according to the licence at 3L in their treatment pathway alongside ECP, rituximab, MMF, sirolimus, and imatinib (Figure 1, below). In the company's pathway, patients would only receive belumosudil after CNIs if these were added at least 4 weeks after the initiation of CS in order for this combination to be considered a second line therapy." In Figure 1 (p29), we propose a footnote should be added, stating: "Only after two systemic lines of therapy".	For clarity of facts and to ensure that the population matches that of the licensed indication. In our submission we do not advocate that the first two lines of treatment can occur concurrently as suggested by the text on page 29 of the EAG report. Rather 1L treatment can be either CS alone or CS in combination with another therapy such as CNIs but only if the CNI is prescribed at the time of CS initiation or within four weeks of it. This is usual clinical practice and is aligned with the clinical trial definition of lines of treatment.	Thank you for highlighting this error. Section 2.2.1 has been edited to reflect the company's intended positioning of belumosudil. A footnote has been added to Figure 1 to note that belumosudil is an alternative to the 2L treatments, after an initial 2L treatment, which could be a later addition of CNIs.

belumosudil, intended to be used as a monotherapy, in the chronic GVHD treatment pathway is after oral CS (with or without the addition of CNIs) and at least one other systemic therapy, such as sirolimus or the later addition of a CNI". In Figure 1, a footnote (linked to the asterisk) has been removed, which was intended to reflect this by stating: "Only after two systemic treatments". Belumosudil is not licensed for use as a second line treatment option. However, the subsequent addition of treatments (e.g. CNIs added to CS) are considered a new line of therapy if added 4 weeks or later after initiation of the previous treatment. This may need clarification in the EAG report.			

Issue 7 Adolescent eligibility

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 31 (Table 9): “Despite being eligible, no adolescents (12-18yo) were recruited to the trial.” Adolescents were not eligible for the KD025-208 trial.	“Therefore, no adolescents (12-18yo) were recruited to the trial”.	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 8 Intervention dose

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 31 (Table 9): “Belumosudil 200 mg BID is the recommended dose.” Typographical error: 200 mg QD is the recommended dose.	“Belumosudil 200 mg QD is the recommended dose.”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 9 Reporting of CNI use

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 31 (Table 9): “The EAG notes that ROCKstar, KD025-208 and REACH-3 trials report on the steroid sparing within the trial but do not report reduction in the use of CNIs”. The CSRs of the ROCKstar (Table 53) and KD025-208 (Table 37) studies reported CNI reduction and discontinuation.	“The EAG notes that the REACH-3 trial reports on the steroid sparing within the trial but does not report reduction in the use of CNIs”.	Factual accuracy. CNI reduction was collected during the studies. The CSRs were provided to the EAG.	Thank you for highlighting this error. Table 9 has been corrected and no longer states that reduction in CNI use is not reported. Relevant edits have also been made to Section 2.3.4. The results in Section 3.3 have been edited by adding the CNI discontinuations reported in ROCKstar to Table 20. The discontinuations reported in the KD025-208 CSR were not appropriate for this report because they were not within the ≥2 LOT subgroup.

Issue 10 Data source for best available therapy proportion

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 41 (Table 12): “Best available therapy (n=164)” Percentages in this table (adapted from Zeiser et al. 2021 Supplementary materials) were calculated from the number of patients treated with BAT, which was 158.	“Best available therapy (n=158)”.	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 11 Patient numbers in baseline characteristics

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 53 (Table 16): “200 mg once daily (n=) / 200 mg twice daily (n=) / Combined 200 mg (n=)” Total patient numbers (included in Table A1 in the Company response to clarification question A1)	“200 mg once daily (n=92) / 200 mg twice daily (n=84) / Combined 200 mg (n=176)”	Typographical error	Thank you for highlighting this omission. The EAG report has been corrected.

have been omitted from the Table.			

Issue 12 3-year data analysis

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 57 (Table 18): “2-year analysis” Page 62 (Table 24): “2-year analysis” Page 63 (Table 25): “2-year analysis” These data refer to the September 2022 data cut, which provide 3-years data. Please note this comment also applies to the text ‘2-year analysis are presented…’ and ‘assessed over two years…’ on pages 71 and page 75 respectively.	“3-year analysis”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 13 ECA typographical error

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 64: “[redacted]” [redacted]	“[redacted]”	Typographical error	Thank you for highlighting this error. The EAG report has been corrected.

Issue 14 Probabilistic results in company’s model

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 79: In Table 35, the value of incremental QALYs for the probabilistic results is shown as 1.72. 1.72 was the value in the company’s model when a severity modifier of 1.2 was applied. Since all the other values in this table were based on the results without a severity modifier (QALY weight of 1), the correct	Replace 1.72 with 1.44	Correction of results	Thank you for highlighting this error. The EAG report has been corrected throughout.

value in this cell should have been 1.44.			
Please note that this issue also affects the following places in the EAG report: • Page 134: “In the base case probabilistic analysis, an incremental quality-adjusted life-year (QALY) gain of 1.72 over best available therapy (BAT)” Table 59 (page 134): incremental QALYs for probabilistic results displayed as 1.72			

Issue 15 Utility data in belumosudil trials

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 81: In two rows of Table 37, it is stated "QALYs based on PROMIS-GH data from ROCKstar and KD025-208 mapped to EQ-5D-3L" PROMIS-GH was only available from ROCKstar, not from KD025-208.	"QALYs based on PROMIS-GH data from ROCKstar mapped to EQ-5D-3L"	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 16 Model structure figure

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 83: Figure 3 is inaccurately reproduced. "Failure – off cGVHD treatment" is missing, and several arrows are missing, e.g., arrows that "loop" to the same state, arrows between on and off	Replace with correct model Figure.	Factual accuracy	Thank you for highlighting this error. The EAG copied the figure directly from the company submission and is thus unsure how several items from the figure were lost. Figure 3 has been reproduced again from Figure 14 of

treatment within the Failure-free state, etc.			the company submission and we have ensured the copy is not missing any information this time.

Issue 17 Response definition for BAT

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 85: In Table 38, it is stated that the definition of response used for BAT was “best response at week 24”. In order to avoid potential misinterpretations and confusion with ‘overall response at week 24’ (which was the primary endpoint in REACH-3 but was not used in our model), we would recommend rewording this to “best response up to week 24” (as the data used in our model were indeed for ‘best overall response up to week 24’).	Reword to “best response up to week 24”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected throughout.
Please note that this issue also affects the following places in the EAG report:			

- Page 92: “However, for the BAT arm, in REACH-3 response was assessed as best response at Week 24.”
- Pages 100-101 (multiple instances): “However, for the BAT arm, in REACH-3 response was assessed as best response at Week 24. The EAG considers that if the company wanted to include response in the model, then definition of response used in the model should match that of REACH-3. As such, during the clarification stage the EAG requested, and the company provided best response at Week 24 for the belumosudil arms of the model (clarification question A11).

  Table 44 presents the best response at Week 24 data for all arms of the model. Compared to the company’s base case response data for belumosudil, best ORR at Week 24 is slightly lower (72.8% and 73.8% versus 69.6% and 70.2% for belumosudil 200 mg QD and BID, respectively). The company provided a scenario based on best response at Week 24 for all treatment arms, which had minimal impact on the ICER, and results are presented in Section 5.2.2.”
- Page 101: caption for Table 44 (“Table 44. Best response at Week 24 data included – pooled analysis of ROCKstar and KD025-208 for the ≥2 LOT subgroup and REACH-3”)

Page 140 (Table 60; Scenario B10): “Best response for belumosudil measured at 24 weeks”

Issue 18 Typographical error (CYP3A inducers)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 87: “…5% of patients would be on CYP3A inhibitors or PPIs.” The SmPC dosing difference for belumosudil relates to strong CYP3A inducers, not inhibitors.	“…5% of patients would be on CYP3A inducers or PPIs.”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 19 Health-state transitions

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 87: “Table 39 summarises the clinical data used to estimate health-state transition probabilities included in the model...” This terminology is potentially confusing and was not used in the CS. This should be updated to align with the caption in Table 39.	“Table 39 summarises the clinical data used to inform health-state transitions included in the model.”	Factual accuracy	This has been amended as per the company’s suggestion for clarity.

Issue 20 Standard parametric survival distributions

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 89: “Extrapolations of the KM data were then explored using standard parametric survival distributions (exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma).”	“Extrapolations of the KM data were then explored using standard parametric survival distributions (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma and generalised gamma).”	Factual accuracy	Thank you for highlighting the omission. The EAG report has been corrected.

The gamma distribution was also explored.			

Issue 21 Statistical fits of failure-free survival distributions

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 90: "Based on statistical fit (not provided in the company's clarification response) and clinical plausibility, the company selected generalised gamma distribution for BAT, belumosudil 200 mg QD and belumosudil 200 mg BID." For FFS, AIC and BIC were provided in our first response to clarification questions A1 (Table 1).	"Based on statistical fit (provided in the company's clarification response) and clinical plausibility, the company selected generalised gamma distribution for BAT, belumosudil 200 mg QD and belumosudil 200 mg BID."	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 22 Life years in failure-free health state

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 90: "Based on the extrapolations, mean FFS for belumosudil and BAT is	"Based on the extrapolations, mean discounted FFS for belumosudil and BAT is [redacted] years and [redacted] years, respectively."	Factual accuracy	Thank you for highlighting this issue. The EAG has considered that presenting undiscounted

[redacted] years and [redacted] years, respectively." The figures presented are the discounted values. This should be explained for clarity and factual accuracy, or the undiscounted values should be presented.			values of mean FFS is more appropriate. The EAG report has been updated to state "mean undiscounted FFS for belumosudil and BAT is [redacted] years and [redacted] years, respectively".

Issue 23 Modelled FFS

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 90-91, Table 40: Values in the modelled FFS column are calculated without applying the cap by OS. This is inconsistent with Figure 5 (which presents the FINAL FFS curves used in the model, i.e., with the cap by OS) and with the text above the figure, which for instance mentions numbers of life-years (which are also based on the final FFS curves used in the model).	Modelled FFS should be updated to include figures including the OS cap.	Clarity required	Thank you for highlighting the inconsistency between Table 40 and Figure 5. The values for 20 and 30 years in Table 40 have been amended to take into account the OS cap.

Issue 24 REACH-3 reference

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 92: "From REACH-3, a KM curve for BAT DOR was available but only median TTR was published.[10]" Median FFS was not published in Zeiser *et al*. 2021 as currently cited, but Le *et al*, 2022.	Update reference to Le *et al*, 2022.	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 25 Median TTR for BAT

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 92: "For the BAT arm of the model, the company estimated a curve using an exponential distribution based on the median TTR	"For the BAT arm of the model, the company estimated a curve using an exponential distribution based on the median TTR (4 weeks) for BAT from REACH-3."	Factual accuracy	In the company submission, median TTR from REACH-3 was cited as 4 months. However, thank you for highlighting the correction, this has

(4 months) for BAT from REACH-3." Median TTR for BAT was 4 weeks, not months.			been amended in the EAG report.

Issue 26 Figure 6. Time to response

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 93, Figure 6: "Modelled time to response for belumosudil and BAT." The belumosudil curves on this graph are not exactly the ones used in the model: even when choosing to use the 'KM' approach to model TTR for belumosudil, we do not use the raw KM data (which are represented by the curves on this graph) but we determine the KM values at the exact time points corresponding to the model cycles. The curves based on the KM values at the start of each model cycle are the	Amend the figure title to: "Time to response for belumosudil and BAT."	Factual accuracy	The EAG does not consider this to be a factual inaccuracy as the KM data presented in the graphs have been used to inform the model engines. No amendment required.

final ones used by the model.			

Issue 27 Mean 'in response' time

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 94: "Based on the extrapolations, mean 'in response' time for belumosudil and BAT is [redacted] years and [redacted] years, respectively." The figures presented are the discounted values. This should be explained for clarity and factual accuracy, or the undiscounted values should be presented.	"Based on the extrapolations, mean 'in response' time (discounted) for belumosudil and BAT is [redacted] years and [redacted] years, respectively."	Factual accuracy	Thank you for highlighting this issue. The EAG has considered that presenting undiscounted values of mean 'in response' time is more appropriate. The EAG report has been updated to state "mean undiscounted 'in response' time for belumosudil and BAT is [redacted] years and [redacted] years, respectively".

Issue 28 Mean OS

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 96: “Based on the extrapolations, mean OS for belumosudil and BAT is [redacted] years and [redacted] years, respectively.” The figures presented are the discounted values. This should be explained for clarity and factual accuracy, or the undiscounted values should be presented.	“Based on the extrapolations, mean OS (discounted) for belumosudil and BAT is [redacted] years and [redacted] years, respectively.”	Factual accuracy	Thank you for highlighting this issue. The EAG has considered that presenting undiscounted values of mean OS is more appropriate. The EAG report has been updated to state “mean undiscounted OS for belumosudil and BAT is [redacted] years and [redacted] years, respectively”.

Issue 29 Modelled OS

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 96, Table 43: Values in the modelled OS column are calculated without applying the cap of general population mortality. This is inconsistent with Figure 9 (which presents the final OS	Modelled FFS should be updated to include figures including the OS cap.	Clarity required	Thank you for highlighting the inconsistency between Table 43 and Figure 9. The values in Table 43 have been amended to

curves used in the model, i.e., with the general mortality cap) and with the text above the figure, which for instance mentions numbers of life-years (which are also based on the final OS curves used in the model).			take into account the OS cap.

Issue 30 BAT observed OS reference

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 96, Table 43: "*Estimates are approximate based on published KM curves in Zeiser *et al.*" This footnote only applies to the value at 2 years in the fourth column (BAT). The value at 1 year (83.8%) was reported in Zeiser *et al*.	Amend footnote and/or asterisk position.	Factual accuracy	Thank you for highlighting this issue. The EAG report has been amended.

Issue 31 Best response up to Week 24

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 101: In Table 44, the number of patients with lack of response for belumosudil 200 mg BID is displayed as '22'. The correct value is 25 (as in Table 9 in the company's response to clarification question A11).	Replace '22' with '25'.	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 32 Typographical error

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 101: "Additionally, for BAT patients who did have a complete or partial response, they could only ruxolitinib if they had disease progression, mixed response, or unacceptable side effects from the control therapy." There is a typographical error here which could cause	"Additionally, for BAT patients who did have a complete or partial response, they could only receive ruxolitinib if they had disease progression, mixed response, or unacceptable side effects from the control therapy."	Typographical error	Thank you for highlighting this error. The EAG report has been corrected.

confusion or misinterpretation.			

Issue 33 Figure 14 reference

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 103, Figure 14: "(reproduced from Figure 9 of the company clarification response)" The correct Figure was Figure 8 in the clarification response.	"(reproduced from Figure 8 of the company clarification response)"	Typographical error	Thank you for highlighting this error. The EAG report has been corrected.

Issue 34 Non-relapse mortality

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 105: "Additionally, if no failure events were observed during a period or all failure events were due to non-relapse malignancy…" The correct term is non-relapse mortality.	"Additionally, if no failure events were observed during a period or all failure events were due to non-relapse mortality…"	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 35 Utility measurements in ROCKstar: failure

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 109: "The company explained that only 23 patients with failure had utility measurements recorded in 52 visits." These figures are for the August 2021 data cut. Other figures quoted in the same section of the EAG report were for the September 2022 data cut, therefore these figures should be updated to use those from the September 2022 data cut, which were also provided in the company's response to clarification question B18.	"The company explained that only 25 patients with failure had utility measurements recorded in 74 visits."	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 36 Standard error for utility value for recurrent malignancy

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 110: In Table 47, the standard error for the utility of patients with acute myelogenous leukaemia (AML) is displayed as 0.32.	Correct value to 0.032.	Factual accuracy We would like to acknowledge that this was mistakenly displayed as 0.32 in the company submission	Thank you for clarifying the error in Appendix N. This has been updated in the EAG report.

The correct value for this standard error is 0.032.		(Appendix N, Table 17) and therefore in the EAG's report.	

Issue 37 Adelphi DSP study (1)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 113: "However, the Adelphi DSP study only had two self-reported UK patient responses.[49]" The correct reference for this statement is the DSP slides (48).	Amend citation.	Referencing error	Thank you for highlighting this error. The EAG report has been corrected.

Issue 38 Adelphi DSP study (2)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 114: "The utility value from the Adelphi DSP study was based on sample of 10 GvHD patients from Europe, but only two were from the UK and there was no distinction for type of cGvHD (acute or chronic)."	"The utility value from the Adelphi DSP study was based on sample of 10 GvHD patients from Europe, but only two were from the UK."	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

As described in the Adelphi DSP study slides referenced, the subgroup analysed only included patients with chronic GVHD.			

Issue 39 Utility measurements in ROCKstar: complete response

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 113: "In ROCKstar, only three patients had a complete response." Six patients had a complete response in ROCKstar. Regarding the utility analysis, six patients with complete response had utility measurements recorded in 32 visits (see Table 33 in the company's response to clarification question B17).	"In ROCKstar, only six patients had a complete response."	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 40 Belumosudil drug cost per cycle

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 117: "The discounted pack price of belumosudil is [redacted], resulting in a cost per cycle (28 days) for the QD regimen of [redacted] and for the BID regimen of [redacted]." The cost per cycle for the BID regimen is [redacted].	"The discounted pack price of belumosudil is [redacted], resulting in a cost per cycle (28 days) for the QD regimen of [redacted] and for the BID regimen of [redacted]."	Correction of typographical error	Thank you for highlighting this error. The EAG report has been corrected.

Issue 41 BAT drug acquisition costs

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 118: In Table 51 (BAT drug acquisition costs), all the values are slightly inaccurate as those were based on the original company model, which used the data cut from August 2021 for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup), including for baseline	Replace the values in Table 51 with:	For accuracy of facts	Thank you for highlighting this error. The EAG report has been corrected.

characteristics (one of which [BSA] affects drug costs for BAT). This should be corrected to use the values from the company's updated (post clarification) model.	• 1st cycle (weeks 1-4): £4,285.38 • 2nd to 3rd cycle (weeks 5-12): £4,265.47 • 4th to 6th cycle (weeks 13-24): £3,415.04 • 7th cycle onwards (weeks 25+) – maximum treatment duration of 5 years: £3,202.32		

Issue 42 IV drug administration costs

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 119: “As a simplification for the base case, the company used the cheaper of the two costs (SB13X, £426.80).”	“As a simplification for the base case, the company used the cheaper of the two costs (SB13Z, £426.80).”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 43 BAT administration costs table

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 119, Table 52: These numbers include accommodation costs. This should made clear, e.g. in the footnote.	Footnote: “These numbers include accommodation costs.”	Factual accuracy	A footnote has been added to Table 52 highlighting the costs included accommodation costs for ECP.

Issue 44 Patients with AML in ROCKstar

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 122: “Acute myeloid leukaemia was the most common underlying malignancy in ROCKstar (43.9%).” 43.9% is the proportion in the QD arm only (based on the Aug 2021 data cut). The value across all belumosudil patients (QD and BID) for the same data cut was 40.9%.	“Acute myeloid leukaemia was the most common underlying malignancy in ROCKstar (40.9%).”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.
Please note that this issue also affects the following place in the EAG report: Page 124: “A one-time cost for post-progression treatment of AML was sourced from TA642, based on AML being the most common malignancy in ROCKstar (43.9%).”			

Issue 45 Positioning of subsequent treatments

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 123, 4.2.7.6 Cost of subsequent treatments: This section discusses subsequent treatment as "fourth line". Technically this should refer to "fourth line plus", since belumosudil can be used in treatment lines later than the third line. This more accurately reflects the licensed position and the clinical trial population. Please note this also applies to Page 131: "In the model, subsequent treatments for patients whose failure event was initiation of new systemic cGvHD treatment was assumed to be the same as BAT, regardless of the treatment patients received at third-line."	Update "fourth line" to "fourth line plus" or "fourth line or later".	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 46 Subsequent treatment drug acquisition costs

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 124: "The per cycle of drug acquisition and administration cost for subsequent treatments was estimated to be £810.27 and £81.99, respectively."	"The per cycle of drug acquisition and administration cost for subsequent treatments	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

The value for drug acquisition costs is slightly inaccurate as this was based on the original company model, which used the data cut from August 2021 for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup), including for baseline characteristics (one of which [BSA] affects drug costs for subsequent treatments). This should be corrected to use the values from the company's updated (post clarification) model.	was estimated to be £808.81 and £81.99, respectively."		

Issue 47 AE cost by treatment arm

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 125: In Table 55 (one-off AE cost by treatment arm), all the values are slightly inaccurate as those were based on the original company model, which used the data cut from August 2021 for the pooled analysis of ROCKstar and KD025-208 (≥2 LOT subgroup), including for incidence of adverse events. This should be corrected to use the values from the company's updated (post clarification) model.	Replace the values in Table 55 with: • Belumosudil 200 mg QD: £186.55 • Belumosudil 200 mg BID: £144.36 • BAT: £987.35	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected.

Issue 48 Table 57. Concomitant medicines usage (1)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 127: In Table 56, there is an issue in the 'Dosing regimen' column, as the text for ECP ("3.2 sessions per 28-day cycle") was spread across two rows by mistake. This resulted in an offset in all the following rows. Page 129: likewise, there is a similar issue in the 'Dosing regimen' column of Table 57	Correct the following values in the 'Dosing regimen' column of Table 56 as follows: • ECP: 3.2 sessions per 28-day cycle • Sirolimus: 2 mg QD • MMF: 1000 mg BID • Budesonide: 3 mg TDS • Montelukast: 10 mg QD • Azithromycin: 250 mg TIW Correct the values in the 'Dosing regimen' column of Table 57 as follows: • ECP: 3.2 sessions per 28-day cycle	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected

	• Sirolimus: 2 mg QD • MMF: 1000 mg BID		

Issue 49 Table 57. Concomitant medicines usage (2)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 129, Table 57: Dosing regimens included in the second column were not derived from the ROCKstar trial. They reflect UK treatment patterns and were taken from a range of sources described in the CS.	Footnote to clarify that dosing regimens may not reflect the concomitant treatment usage in the ROCKstar trial.	Factual accuracy	For clarity, the EAG has made the proposed amendment in the EAG report.

Issue 50 Table 57. Concomitant medicines usage (3)

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 129, Table 57: Footnote stating "*Weighted average of belumosudil 200 QD and BID based on proportions on each	Remove footnote	Typographical error	Thank you for highlighting this error. The EAG report has been corrected

regiment (base case values of 95% for QD regimen and 5% for BID regimen)" is not relevant to the table and should be removed.			

Issue 51 Probabilistic results for INMB and INHB

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 134: "Using the £20,000 and £30,000 threshold, the incremental net monetary benefit (INMB) is [redacted] and [redacted] and the incremental net health benefit (INHB) is [redacted] and [redacted]." These values correspond to the deterministic results, while the rest of this paragraph relates to probabilistic results.	Replace with probabilistic results for INMB and INHB: "Using the £20,000 and £30,000 threshold, the incremental net monetary benefit (INMB) is [redacted] and [redacted] and the incremental net health benefit (INHB) is [redacted] and [redacted]."	For accuracy of results	Thank you for highlighting the inconsistency in the results presented. The suggested amendment has been made in the EAG report.

Issue 52 Results for Scenario B21 in Table 60

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 140: In Table 60, the results displayed for Scenario B21 are mistakenly the same as those for Scenario B20. This should be updated with the correct results for Scenario B21.	Replace QALYs and ICER results for Scenario B21 with: • Belumosudil Total QALYs: [redacted] • BAT Total QALYs: [redacted] • Incremental QALYs: [redacted] • ICER: £3,609	Correction of results	Thank you for highlighting this error. The EAG report has been corrected

Issue 53 Results for Scenario 3 in Table 61

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 144: In Table 61, the results displayed for Scenario 3 are mistakenly the same as those for Scenario 2. This should be updated with the correct results for Scenario 3. We believe that those are the results that are (mistakenly) presented	Replace results for Scenario 3 with: • Belumosudil Total Costs: [redacted] • BAT Total Costs: £248,736 • Belumosudil Total QALYs: [redacted] • BAT Total QALYs: [redacted] • Incremental Costs: [redacted]	Correction of results	Thank you for highlighting this error. The EAG report has been corrected in line with the error corrected for Issue 3.

under Scenario 4 in this same table.	• Incremental QALYs: [redacted] • ICER: £17,085		
Please note that the proposed amendment above is intended to be consistent with other results provided by the EAG and therefore does not include the correction that would be needed in relation to the issue with the removal of costs of background therapies that we described under Issue 3 in this form.			

Issue 54 Results for Scenario 4 in Table 61

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**EAG response**
Page 144: In Table 61, the results displayed for Scenario 4 are mistakenly those corresponding to Scenario 3 (see Issue 53 above). This should be updated with the correct results for Scenario 4, i.e., for a combination of Scenarios 1 and 3.	We believe the correct results for Scenario 4 would be: • Belumosudil Total Costs: [redacted] • BAT Total Costs: £248,736 • Belumosudil Total QALYs: [redacted] • BAT Total QALYs: [redacted] • Incremental Costs: [redacted] • Incremental QALYs: [redacted] • ICER: Dominant	Correction of results	Thank you for highlighting this error. The EAG report has been corrected in line with the error corrected for Issue 3.

Please note that the proposed amendment above is intended to be consistent with other results provided by the EAG and therefore does not include the correction that would be needed in relation to the issue with the removal of costs of background therapies that we described under Issue 3 in this form.

Issue 55 Table 66 typographical error

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 150, Table 66: Sex distribution- male has a value of "58.%". To avoid confusion, this should be reported as 58.0% or 58%.	"58.0%"	Typographical error	Thank you for highlighting this error. The EAG report has been corrected.

Issue 56 Proportional QALY shortfall

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 152: "However, the severity modifier of 1.2 does not apply to the EAG's preferred cost effectiveness results, as the absolute QALY shortfall is less than 12 and the	"However, the severity modifier of 1.2 does not apply to the EAG's preferred cost effectiveness results, as the absolute QALY shortfall is less than 12 and the proportional QALY shortfall is less than 0.85".	Typographical error	Thank you for highlighting this error. The EAG report has been corrected

proportional QALY shortfall is less than 0.80”. The correct value is 0.85.			

Issue 57 Caregiver disutility in EAG base case

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Page 154: In the last row of Table 73, the header “Caregiver disutility - failure” should be edited in order to clarify that this change, made as part of the EAG base case, only concerns failure due to initiation of a new systemic cGvHD therapy. The caregiver-related disutility applied for failure due to recurrent malignancy remains the same as in the company’s base case.	Replace header with: “Caregiver disutility – failure – initiation of a new systemic cGvHD therapy”	Factual accuracy	Thank you for highlighting this error. The EAG report has been corrected

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Thursday 20 July**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Sanofi
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	**None**
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Evidence for adolescents not available from Rockstars and KD025-208	No	Given the unmet need in treatment options for cGVHD across all age groups, and the biological plausibility of using belumosudil in patients aged 12 to 18 years old, we consider it reasonable and appropriate to align the eligible population with that of the MHRA licence.
Inclusion of concomitant medication costs for belumosudil, such that the intervention for the cost-effectiveness analysis is belumosudil in addition to best available therapy	Yes	The licence for belumosudil does not state the requirement for any concomitant treatment when used to treat cGVHD. However, we acknowledge that in the clinical trials a proportion of patients did receive concomitant cGVHD medication with belumosudil, and this may be reflective of NHS clinical practice. Therefore, it is not unreasonable to include these costs in the economic model. In clinical practice we also consider it likely, based on insights from clinical experts, that concomitant treatments would be used alongside BAT to a similar extent to belumosudil. The treatment costs in the BAT arm of the model currently characterise BAT as the equivalent of one therapy but this is modelled as a basket to reflect the options available. This is based on the source trial, REACH-3, which limited co-medication to glucocorticoids and calcineurin

(BAT) (belumosudil+BAT)

inhibitors within the study protocol. However, this study protocol is not reflective of clinical practice and the absence of concomitant treatments in the BAT arm of the model means that their associated costs are likely to be underestimated.

The scenario analyses presented during clarification questions included the costs of concomitant therapies according to the pooled analysis of the ROCKstar and KD025-208 studies for the ≥2 LOT subgroup (2021 data cut). As requested by the EAG in their report, we subsequently obtained the updated 2022 data cut for this subgroup and re-ran the scenario analysis, employing the EAG's preferred assumptions (see Table A and B below). This reduced the incremental costs from [redacted] to [redacted], primarily due to the lower proportion of concomitant ECP usage in trial participants. This results in a marginal change in the magnitude of the ICER and has no impact on decision risk.

Table A. Model inputs for concomitant treatment (EAG base case and updated data cut)

Concomitant treatment	**Pooled analysis, 2021 data cut (n=156)**		**Pooled analysis, 2022 data cut (n=176)**	
	QD (n=81)	BID (n=75)	QD (n=92)	BID (n=84)
ECP n (%)	20 (24.7%)	26 (34.7%)	22 (23.9%)	28 (33.3%)
Sirolimus n (%)	17 (21.0%)	18 (24.0%)	21 (22.8%)	20 (23.8%)
MMF n (%)	11 (13.6%)	2 (2.7%)	11 (12.0%)	3 (3.6%)
Total	48 (59.26%)	46 (61.33%)	54 (58.70%)	51 (60.71%)

Table B. Cost-effectiveness results from deterministic analysis

	EAG base case (2021 data cut)	**EAG base case using 2022 data cut**
Incremental costs	[redacted]	[redacted]
Incremental QALYs	[redacted]	[redacted]

		Cumulative ICER (£/QALY) | Dominant | Dominant
Naïve comparison of belumosudil versus BAT	No	We acknowledge the limitations of using the BAT arm from the REACH-3 study, which was conducted in patients receiving an earlier line of therapy, in a naïve indirect comparison with belumosudil. As described in the company submission, we explored multiple different approaches to address this uncertainty. We agree with the EAG and clinical experts that the naïve comparison is currently the only feasible option to compare clinical outcomes for belumosudil (+BAT) with BAT. There are currently no Phase III RCTs of belumosudil versus BAT in third line cGVHD ongoing either in the UK or internationally. Expert clinicians considered the results of this naïve comparison to be clinically plausible and suitable for decision making in the absence of direct head-to-head data. Therefore, despite the uncertainty around the expected impact of this naïve comparison, the consistently low ICERs presented in the company base case, EAG preferred scenario, and extensive scenario analyses, provide confidence that belumosudil (+BAT) is cost-effective compared to BAT and there is low decision risk.
Removal of response outcomes from the economic model	No	We agree with the EAG that failure-free survival is the most clinically relevant outcome for patients and consider it to be the most suitable endpoint for cost-effectiveness modelling. This was consistent with the clinical expert opinion we received during the development of the model and company submission. We also agree with the EAG that excluding response from the model removes a source of unresolvable uncertainty whilst having a minimal impact on the resulting ICERs.

Removal of overall survival benefit for belumosudil+BAT	No	As stated in our company submission, we have no direct data which demonstrate a relative overall survival benefit for belumosudil vs. standard care. Considering the uncertainty around relative overall survival between the two treatments, we do not consider it unreasonable for the EAG to remove the overall survival benefit originally included in the company submission model. Whilst the impact of this change on the ICER is not insignificant, it is worth reflecting that if the original assumption of overall survival benefit for belumosudil is maintained but all the other EAG adjustments to the company model are implemented the ICER remains under £30,000/QALY. If the OS benefit is removed from the company base case belumosudil is the dominant strategy. Whilst the OS assumption is undoubtedly the biggest driver of the ICER, taken together these ICER estimates indicate that the inclusion or exclusion of the OS benefit does not introduce a risk of decision error in judging belumosudil to be a cost-effective option.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response

Additional issue 1: Subsequent treatment of cGVHD duration may exceed 5 years in some patients	Section 4.2.7.6 – pages 122-123 Section 4.2.7.9 – pages 130-132	No	The duration of subsequent treatments for patients for whom third-line treatment fails and the proportion of these patients on treatment was identified as a secondary issue for the EAG and a primary driver of cost-effectiveness in the model. The EAG preferred base case assumes a 5-year subsequent treatment duration with only 60% of patients remaining on treatment post-failure. We consider this to be an underestimate. Assuming that 40% of patients do not receive any subsequent treatment and that the remaining 60% stop receiving any treatment for cGVHD by 5 years does not seem clinically plausible considering the chronic nature of the disease and its likely severity at this stage in the treatment pathway. Feedback we received from clinicians was that cGVHD patients were unlikely to remain on a *single third line therapy* for more than 5 years. However, after third line treatment failure, they could subsequently cycle through multiple lines of treatment over the remainder of their lifetime. This is reflected in the basket of subsequent treatments modelled in the company base case, which assumes that patients remain on some form of treatment for 60% of their remaining life years after third line (or later) treatment failure. The 60% proportion is intended to reflect the overall proportion of time spent on treatment, rather than the proportion of patients going on to receive a subsequent treatment, thereby accounting for adherence issues, treatment gaps, and or therapy discontinuations. Despite the limited available evidence to accurately model post-failure treatment costs, the range of assumptions employed by the company and EAG consistently result in ICERs below the cost-effectiveness threshold.

Additional issue 2: Choice of utility value for treatment failure	Section 4.2.6 – pages 105-109, 112-114	No	The midpoint utility value of 0.608 selected by the EAG for patients in the "failure-new systemic therapy" health state is associated with considerable uncertainty. It was derived in part from a published cGVHD economic analysis by Crespo *et al*, which reported a utility value for a 'Progression' health state of 0.696. This figure, quoted in Crespo *et al*, could not be traced back to a value in the source study cited by Crespo *et al*, and despite contacting the authors we were unable to establish how they derived this estimate. The other data point informing the midpoint value (0.52) was collected from European patients with cGVHD. Although the sample size was very small (n=10), it is at least a verifiable source and is likely to reflect the utility of patients more accurately at this point in the treatment pathway. Indeed, it may be conservative as it is above the value that we validated for this health state (0.479) in an advisory board and subsequent individual consultancies with clinicians. It should be remembered that this utility is carried forward for the remaining time in the model and so should reflect the worsening stages of the disease. In the general population vignette-based utility elicitation we conducted, the utility value for this health state was substantially lower than any of the alternative sources. Whilst seemingly extreme, this study was conducted following Decision Support Unit's best practice recommendations in line with the NICE process and methods guide and reflects the UK societal valuation of this health state. Therefore, it should not be completely discounted in this appraisal. This is an especially important parameter because the applicability of the severity modifier rests upon the severity of disease defined by quality of life decrement in the later parts of the patient journey. Patient and clinician testimony that we have received indicates that cGVHD is an extremely burdensome disease affecting all areas of a patient's life, some of which, like social functioning will not be fully captured in the QALY. Therefore, we disagree with the EAG analysis which removes the severity modifier.

			We consider this an important point for deliberation and would encourage the patient perspective to be incorporated during Committee discussions. Acknowledging the uncertainty of this parameter, we consider the utility value in the EAG base case to be an upper estimate of the plausible value, resulting in an underestimation of the cost-effectiveness of belumosudil. Nonetheless, even with this upper estimate the ICER remains below the commonly accepted threshold and so decision risk is low.

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR 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 incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

Key areas for consideration

1. Demographics
2. Comparator and definitions of second- and third-line therapy
3. Disutilities
4. Modelling around overall survival.

Of these points 2,3,4 are likely to have the greatest potential impact on the economic impact.

Demographics

B 1.3.1.2

Based on the 2022 BSBMT report to commissioners (13th Edition), reporting results from 2016-19, the number of allogeneic transplants performed per year in the UK is around 1300 (adult) and 330 (paediatric). The number of procedures performed in 2020 and 2021 fell due to the COVID pandemic and as such 2019 data (1571 procedures total) are more representative; assuming 90.8% in England this gives a total of 1426 procedures. Not all of these are eligible for the technology as a significant proportion of the paediatric patients will be under 12 years.

Approximately 30% develop cGVHD of which around 45% is extensive. Recently published US data suggest that up to 45% will require third line therapy[1] although this may not translate precisely to a UK population due to variable clinic practice. According to the 2022 BSBMT report, cGvHD is much more common in adult patients (33%) compared to paediatric patients (16%)

Prevalence data for cGvHD are sparse however appear to have been underestimated. Based on US data, the prevalence was around 2:50000 although this is likely to be an underestimate due to the source data for the estimates.[1] The median survival of cGvHD is better than suggested. This depends on a number of factors including stage of disease and patient age however median OS may be as long as 4-5 years.[2] On this basis an estimate of prevalence could be up to 5:50000.

B 1.3.1.3

GvHD is a significant cause of death in allogeneic transplant recipient however the figure given relates to all GvHD rather than cGvHD. Extensive cGvHD is however a major cause of morbidity as reported.

B 1.4

I am not aware of any published data (which might inform expert opinion) to support the hypothesis that there is more cGvHD in minority ethnic transplant recipients.

Comparator and definitions of second- and third-line therapy

B 1.3.2.3

Belumosudil is presented being a third line therapy on the basis that steroids are first line therapy and CNI are second line therapy (additional arguments in support made in 1.3.2.1). As such the direct comparators would be ECP, rituximab, MMF, imatinib, sirolimus. This is however not consistent with UK clinical practice where CNI are unlikely to be used as a second line treatment but

[1] Bachier CT et al Transplantation and Cellular Therapy 2021;504:e1
[2] Pidala J et al Haematologica 2011;96:1678

more likely to be used alongside steroids as part of first line therapy. As a third line therapy, belumosudil should be positioned for patients who have failed recognised second line treatment (ECP, rituximab, imatinib, pentostatin); or 2 lines of second line therapy if the 2017 NHS commissioning guidance is to be followed. Of these treatments, ECP is most widely used in the UK and it would be reasonable to assume that the majority of patients reaching third line therapy would have failed prior ECP and as such this is not a plausible comparator in the third line setting. Interpretation of the trial data is not straightforward on the basis that

1. cGVHD is clinically heterogenous and treatments may show differential efficacy in different patients dependent on the extent of their disease. For example, data presented in figures 8 and 11 indicate that the efficacy of belumosudil is greatest in joints/fascia, gut and mouth whereas published data for ECP (for example) have reported greatest efficacy in skin, mouth and eyes.[3]
2. The belumosudil trial data report that the majority of patients included had received treatments beyond steroids and CNI and as such the evidence does not really support use in 3rd line therapy (as defined by the company).[4,5]
3. In the control arm of REACH-3, 42% of patients received CNI in addition to steroids supporting the case that CNI are not generally considered to be a line of treatment.[6]

B 2.9.2

It is regrettable that external control data could not be used to provide a more direct comparison. Outcome data regarding outcomes of UK patients post ECP would have been informative had they been available. I don't have access to appendix M to comment further.

B 3.2.3

Agree that identification of a comparator is challenging and using REACH 3 control arm was a pragmatic compromise accepting that the trial populations are not directly superimposable. In REACH-3 less than 10% of patients had received treatment for cGVHD beyond steroids and CNI compared to more than 50% of those in the belumosudil trials. Better responses might be anticipated for patients at earlier stages of disease and pragmatically if there is a benefit of belumosudil at third line over BAT at second line this may be real. This does not however imply that belumosudil should necessarily be placed as second line therapy as suggested.

Disutility

B 1.3.1.4

I struggle to understand the use of utility values obtained from members of the public. As referenced, published data suggest that these vary compared to those obtained from patients. Whilst I acknowledge that general public can appreciate the impact of GvHD on wellbeing however observational data in patients should be applied in this analysis.

[3] Flowers MED et al Blood 2008;112:2667
[4] Cutler C et al Blood 2021;138:2278
[5] Jagasia M et al J Clin Oncol 2021;39:1888
[6] Zeiser R et al N Engl J Med 2021;385:228

The impact of therapy section needs to be more objective. Whilst accepting that treatments can be associated with significant toxicity this should be quantified in terms of frequency and severity.

B 3.4.3

The assumptions over utility of 4th line therapy (which is equated to recurrent malignancy) are very weak and not really justified. The disutility for haematological cancers appears reasonable however is likely to over-estimate the disutility associated with cGvHD

B 3.4.6

I am unclear why MS has been chosen to inform assumptions over disutility in caregivers looking after patients with cGvHD. Choosing other chronic illness may give different results and this also depends on other patient factors (eg age)

B 3.4.8

There is a significant proposed difference in the utilities for failure-free: LR and failure: new therapy required. In clinical practice, it is unlikely that a patient who showed a lack of response would be considered failure-free and as such not clear on how this differs from those with failure needing a new therapy. The utilities associated with both states would be expected to be very similar. Access to therapy may be an issue however no data are presented apart from expert opinion. In my centre we have no issues with access to ECP or capacity although this is due to having this available locally. The majority of the service is provided by NHSBT and consultation with them would give a better view on any national capacity/logistic constraints than what is presented.

Overall survival

B 3.3.2.2

The case made for OS advantage of belumosudil over BAT is not robust. From a historical perspective, there has been very little progress made in improving survival for patients with cGvHD in recent years. Assessment of OS requires data maturity and cannot be inferred from response endpoints. In the REACH-3 trial OS was the same in both arms despite a significant benefit a measured by response rate although this is potentially confounded by heterogeneity in the BAT control arm and crossover. With regard to this technology, REACH-3 and the belumosudil trials is short and the populations are not directly comparable. An OS advantage is not entirely implausible however based on the published data this is less likely and certainly cannot be assumed.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Clinical expert statement and technical engagement response form

Thank you for agreeing to comment on the external assessment report (EAR) for this evaluation, and for providing your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. The EAR and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR section 1.1. You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

A clinical perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

Please note, part 1 can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm** on **Thursday 20 July**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating chronic graft versus host disease and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Professor Adrian Bloor
2. Name of organisation	Christie NHS Foundation Trust
3. Job title or position	Consultant Haematologist
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 chronic graft versus host disease? ☒ A specialist in the clinical evidence base for chronic graft versus host disease or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Nil
8. What is the main aim of treatment for chronic graft versus host disease?	Resolution of disease is primary goal. Secondary objectives are control of symptoms, time to next therapy and improvement in quality of life

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	Complete remission is desirable. Otherwise improved time to next treatment
10. In your view, is there an unmet need for patients and healthcare professionals in chronic graft versus host disease?	Yes. Treatment is suboptimal particularly for patients resistant to first line therapy (steroids)
11. How is chronic graft versus host disease currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	This is managed per 2017 NHS commissioning guidance (https://www.england.nhs.uk/wp-content/uploads/2017/03/gvhd-heamatopoietic-stem-cell.pdf). Treatment is anticipated to be broadly uniform however subject to some local variation (eg access to ECP, local funding agreements or access to clinical trials). The technology could have a significant impact on treatment pathways offering an orally active alternative to currently available therapies. As listed below, there is however a concern as to where this would be placed in the current treatment pathway due to an inconsistent approach taken in the submission.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	The technology is not currently available in the UK and is not used. This is an orally active therapy which would be used exclusively in secondary care setting within allogeneic stem cell transplant units (<30 nationally). As an oral therapy, this would not require investment in infrastructure and if superior to currently available treatments then could lead to reduced infrastrure requirement (eg due to less iv therapy or use of ECP which is delivered in the hospital setting)

13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	If effective then the technology should impact positively on health related QoL as chronic GvHD is a major cause of morbidity post transplant. It is however far from certain if this would lead to an overall survival advantage which has not been demonstrated for other therapies.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	There is no reason to expect this. With regard to section B1.4, I am not aware of any published data (which might inform expert opinion) to support the hypothesis that there is more cGvHD in minority ethnic transplant recipients.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	As an oral therapy with a favourable toxicity profile this would be expected to be easier to use for patients and HCP compared to other treatments
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Response is assessed clinically. Treatment would be expected to continue until resolution of symptoms, failure of treatment or unacceptable side effects.
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen	There is a potential health related benefit however the data presented are hampered by the following considerations. 1. Lack of a comparator 2. Reliance on data obtained from US treatment pathways which differ from those used in the UK 3. Inclusion of an overall survival advantage

may be more easily administered (such as an oral tablet or home treatment) than current standard of care	4. Positioning of technology as a third line treatment based on flawed assumptions over sequencing of cGvHD therapies in UK (ie considering calcineurin inhibitors as second line therapy)
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Any effective treatment for steroid refractory cGvHD offers a significant benefit for stem cell transplant patients in the UK. The major question is however how (cost) effective this is compared to currently available technologies which is informed to a significant extent by the comparator(s) used and where this is placed in treatment
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The technology appears to be generally well tolerated and as such is unlikely to adversely affect patients' quality of life
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The results of the trials are broadly applicable to UK patients however the lack of comparator in the published data hamper this. The REACH-3 standard of care arm is not really directly comparable as this was used at an earlier stage of disease and includes therapies not available in the UK. There is no defined optimal endpoint for cGvHD studies; overall survival would be desirable however may not be feasible within the context of a trial. With regard to surrogate endpoints, use of ORR as the primary outcome measure is not entirely unreasonable however is of limited importance to patients. Use of endpoints related to survival, duration of response (eg ongoing response at 12 months – Martin PJ et al Blood 2017;130:360) or time to next therapy are more relevant
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No

23. How do data on real-world experience compare with the trial data?	Very limited high quality data are available to comment. The company have attempted to collect real world data unsuccessfully. It is regrettable that external control data (B 2.9.2) could not be used to provide a more direct comparison. Outcome data regarding outcomes of UK patients post ECP would have been informative had they been available. I don't have access to appendix M to comment further.
24. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why.	I have no concerns or comments

More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Evidence for adolescents not available from Rockstars and KD025-208 *Would you expect the efficacy, as well as the side effect profile of belumosudil to be similar in adolescents as in adults?*	The therapeutic index for many cancer drugs is better in younger patients. I have no data to directly support this related to this technology however if anything would anticipate a better outcome in younger patients and see no reason that they would not be included.
Inclusion of concomitant medication costs for belumosudil, such	Belumosudil is presented being a third line therapy on the basis that steroids are first line therapy and CNI are second line therapy (additional arguments in support made in 1.3.2.1). As such the direct comparators would be ECP, rituximab, MMF, imatinib, sirolimus. This is however not consistent with UK

that the intervention for the cost-effectiveness analysis is belumosudil in addition to best available therapy (BAT) (belumosudil+BAT) *The company included belumosudil in the model as a monotherapy and costs used reflect this assumption.* *The EAG suggest that the intervention in the model should reflect belumosudil+BAT.* *Do you agree that the intervention in the model should reflect belumosudil+BAT? Would this reflect UK clinical practice?*	clinical practice where CNI are unlikely to be used as a second line treatment but more likely to be used alongside steroids as part of first line therapy. As a third line therapy, belumosudil should be positioned for patients who have failed recognised second line treatment (ECP, rituximab, imatinib, pentostatin); or 2 lines of second line therapy if the 2017 NHS commissioning guidance is to be followed. Of these treatments, ECP is most widely used in the UK and it would be reasonable to assume that the majority of patients reaching third line therapy would have failed prior ECP and as such this is not a plausible comparator in the third line setting. Using the REACH 3 control arm was a pragmatic compromise accepting that the trial populations are not directly superimposable. In REACH-3 less than 10% of patients had received treatment for cGVHD beyond steroids and CNI compared to more than 50% of those in the belumosudil trials. Better responses might be anticipated for patients at earlier stages of disease and pragmatically if there is a benefit of belumosudil at third line over BAT at second line this may be real. This does not however imply that belumosudil should necessarily be placed as second line therapy as suggested. The technology has been incorrectly placed as an alternative to ECP/Rituximab/MMF/sirolimus etc which does not correctly reflect third line therapy according to UK practice. Interpretation of the trial data is not straightforward on the basis that 1. cGVHD is clinically heterogenous and treatments may show differential efficacy in different patients dependent on the extent of their disease. For example, data presented in figures 8 and 11 indicate that the efficacy of belumosudil is greatest in joints/fascia, gut and mouth whereas published data for ECP (for example) have reported greatest efficacy in skin, mouth and eyes (Flowers MED et al Blood 2008;112:2667) 2. The belumosudil trial data report that the majority of patients included had received treatments beyond steroids and CNI and as such the evidence does not really support use in 3rd line therapy (as defined by the company) (Cutler C et al Blood 2021;138:2278, Jagasia M et al J Clin Oncol 2021;39:1888)

	3. In the control arm of REACH-3, 42% of patients received CNI in addition to steroids supporting the case that CNI are not generally considered to be a line of treatment (Zeiser R et al N Engl J Med 2021;385:228)
Naïve comparison of belumosudil versus BAT	See above
Removal of response outcomes from the economic model *The EAG considers that inclusion of response in the model is potentially adding unnecessary complexity to the analysis.* *Should response be included in the model? Is failure-free survival a more clinically relevant outcome for patients?*	As above, ORR is not particularly clinically relevant and survival endpoints or time to next therapy would be more meaningful for patients. ORR may of course be a surrogate for later efficacy however I agree this may not be important for the analysis
Removal of overall survival (OS) benefit for belumosudil+BAT *What is the clinical plausibility of an OS*	The case made for OS advantage of belumosudil over BAT is not robust. From a historical perspective, there has been very little progress made in improving survival for patients with cGvHD in recent years. Assessment of OS requires data maturity and cannot be inferred from response endpoints. In the REACH-3 trial OS was the same in both arms despite a significant benefit a measured by response rate although this is potentially confounded by heterogeneity in the BAT control arm and crossover. With regard to this technology, REACH-3 and the belumosudil trials is short and the populations are not

benefit for belumosudil+BAT?	directly comparable. An OS advantage is not entirely implausible however based on the published data this is less likely and certainly cannot be assumed.
Are there any important issues that have been missed in EAR?	See comments below regarding disutility assumptions. B 1.3.1.4 I struggle to understand the use of utility values obtained from members of the public. As referenced, published data suggest that these vary compared to those obtained from patients. Whilst I acknowledge that general public can appreciate the impact of GvHD on wellbeing however observational data in patients should be applied in this analysis. The impact of therapy section needs to be more objective. Whilst accepting that treatments can be associated with significant toxicity this should be quantified in terms of frequency and severity. B 3.4.3 The assumptions over utility of 4th line therapy (which is equated to recurrent malignancy) are very weak and not really justified. The disutility for haematological cancers appears reasonable however is likely to over-estimate the disutility associated with cGvHD B 3.4.6 I am unclear why MS has been chosen to inform assumptions over disutility in caregivers looking after patients with cGvHD. Choosing other chronic illness may give different results and this also depends on other patient factors (eg age) B 3.4.8 There is a significant proposed difference in the utilities for failure-free: LR and failure: new therapy required. In clinical practice, it is unlikely that a patient who showed a lack of response would be considered failure-free and as such not clear on how this differs from those with failure needing a new therapy. The utilities associated with both states would be expected to be very similar. Access to therapy may be an issue however no data are presented apart from expert opinion. In my centre we have no

	issues with access to ECP or capacity although this is due to having this available locally. The majority of the service is provided by NHSBT and consultation with them would give a better view on any national capacity/logistic constraints than what is presented.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

Information on completing this form

In part 1 we are asking you about living with chronic graft versus host disease or caring for a patient with chronic graft versus host disease. The text boxes will expand as you type.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Patient expert statement

NICE National Institute for Health and Care Excellence

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

The deadline for your response is **5pm** on **7 September 2023**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Part 1: Living with this condition or caring for a patient with chronic graft versus host disease

Table 1 About you, chronic graft versus host disease, current treatments and equality

1. Your name	
2. Are you (please tick all that apply)	☑ A patient with chronic graft versus host disease? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with chronic graft versus host disease? ☑ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Anthony Nolan
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☑ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☑ I agree with it and **will be** completing ☐
5. How did you gather the information included in your statement? (please tick all that apply)	☑ I am drawing from personal experience ☑ I have other relevant knowledge or experience (for example, I ☐am drawing on others' experiences). Please specify what other experience: Drawing on experiences of others as shared on Anthony Nolan'sPatient & Families forum.

Patient expert statement

	☐ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement
6. What is your experience of living with chronic graft versus host disease? **If you are a carer (for someone with chronic graft versus host disease) please share your experience of caring for them**	I had my Stem Cell Transplant in June 2020 and had no acute GvHD. Once I came off Steroids and Ciclosporin, signs began to appear of Chronic GvHD in my eyes, skin, mouth and gut. Since then I have been diagnosed with GvHD in my eyes, mouth, gut and skin, with suspected lung GvHD as well. I will take each of these individually and then present my overall thoughts at the end. 1. My gut GvHD manifests itself in two different ways: significant acid reflux & diarrhoea. The acid reflux is being managed with Omeprazole (20mg twice daily) and flares up regularly - it also has a quality of life impact, since it restricts what I can and cannot eat and drink. During flare ups it can be debilitating, with significant pain that neither painkillers nor regular heartburn medicine can help. The Diarrhoea is being managed with Budesonide (3mg thrice daily) and at its worst can be very frequent, painful and can include blood in my stool. The Budesonide does a good job managing it, but itself causes Acid Reflux and it does not prevent flare ups. 2. The GvHD in my mouth presents as Severe Dry Mouth and is constant. It is easily manageable when quiescent but does require constant sips of water as breathing/talking dries my mouth out quickly. This can be challenging to achieve when on the move, but with planning is achievable. This also makes eating a chore, since every bite of a sandwich for example must be followed immediately by a mouthful of water, and my mouth is VERY sensitive to strong flavours, acidic foods, any hint of spice etc.

	This also has a significant impact on my ability to exercise, as anything vaguely challenging will require me to stop constantly for sips of water and then the necessary toilet breaks for someone that drinks 4-5L of water per day. Outside of the initial symptoms, the more significant effects are to oral hygeine. Due to my lack of saliva, I am far more susceptible to tooth decay due to plaque build up - indeed I have had to have 15 fillings in the last 36 months, with only one filling previously in my life. I also get regular ulcers throughout my mouth, which can be very painful. Flare ups can occur, but are rare these days, with the ones that do happen being managed with a Nystatin paste and a steroid mouthwash. 3. My Skin GvHD has been diagnosed recently and we are going through an evaluation process currently with the Dermatology Team - currently I have very dry skin all over, with several patches of pigmentation change and some patches that are clearly distressed. On top of this my skin heals very slowly or not at all, which presents its own challenges. 4. The GvHD in my eyes has been the most limiting to date, as it has caused me to be unable to continue working. It presents as dry eyes and I am under the care of Moorfields Eye Hospital for treatment and management. We have tried a lot of different management methods, and I am currently taking: ● Ciclosporin eye drops (twice daily) ● Prednisolone Eye Drips (once daily) ● Warm Eye Compresses & wipes (thrice daily) ● Nightly eye ointment ● Serum Eye Drops (as needed - typically every 15-30 minutes) ● Doxycycline Oral antibiotics

	The Serum Eye Drops have been prescribed recently and I have only been using them for a few weeks - initial impressions are that they are better than the commercially available lubricating eye drops, but that I still need to use them far more often than recommended, with amounts varying wildly day by day. They are also incredibly hard to manage unless you are at home all day every day since they need to be kept in the freezer and defrosted on the day of use. Given that the amount that I need to take varies day by day, I am using more than one vial per day roughly 50% of the time. So a challenge here is that if I go out for the day (to visit family out of london for example) do I take one or two? I get a fixed amount per 90 days, so by taking two when I might not need them I get through my vials far more quickly than I would like and I am left without any for the last month, or by only taking one I take the significant risk that I run out and then have nothing with me. In this situation it would leave me in significant pain and unable to see/open my eyes. In particular, this issue is exacerbated by looking at computer screens and by being on Video Calls especially. This caused me to have to go back off sick from work recently, which eventually led to me being let go by my company, effectively ending my previously very successful corporate career. Other aggravating factors are being outside when sunny or windy, being in air conditioned environments or being in cold environments. During flare ups I can have both Blepharitis and Conjunctivitis, which can be debilitating and very difficult to deal with. Overall this causes the most issues on a day to day basis and requires both the most planning and the most time to manage. It has also caused me to be unable to work and has caused the end of my career in technology sales.

	5. Lung GvHD - This is currently suspected, following several lung function tests and scans - one of the lung functions showed a significant reduction in lung capacity (45%) and I have a preliminary diagnosis of Chronic Bronchiolitis, but this is under investigation. This also makes me very vulnerable to colds, flu & other respiratory infections, which causes significant anxiety. Overall, because of the combination of GvHD that I currently have, my quality of life is significantly diminished. It has cost me my ability to play sports, exercise, work and makes day to day life challenging and complex with the treatments and planning that needs to go into keeping on top of things. It has also led to significant weight loss (15+kg) since my body is using a significant number of calories per day. Even with dietician support and specialist supplements topping me up to 3500 calories per day, it is a challenge to keep weight on, especially when the symptoms themselves can make eating difficult. I also suffer badly from fatigue, meaning that everyday tasks can be difficult and the fatigue factor has to be taken into account when planning what to do on a day or how to approach a new job. Outside of the physical symptoms, the impact can be felt both mentally and socially. Mentally it has been very hard for me to accept that I have long term chronic conditions at the age of 30 that are side effects from the treatment that saved my life, and that will limit both my quality of life and my lifespan moving forwards. I have had clinical psychological support for this and it is something that we are continually working on. My wife has also been heavily impacted by this and doesn't have the same specialist psychological support available to her on the NHS.

	Due to the physical manifestation of GvHD, it is exceptionally hard not to be isolated socially, since symptoms can flare up at any time and even without flare ups, something that seems simple (ie, going for a catch up with a friend at a pub) can be incredibly challenging events. Managing all the different symptoms whilst not in a home setting is very hard and has lead to me declining invitations that otherwise I would have loved to accept. There have also been cases where the anxiety around being in public (due to the possibility of picking up an infection that leads to serious damage to my lungs) has caused physical symptoms to flare up. This then leads to pulling out of an event last minute and can have significant impacts on friendships & other relationships. This also leads to people stopping inviting me to things as they know it is hard for me to say no, so they wait for me to instigate things, leading to more isolation. Social isolation and the mental toll of suffering from Chronic GvHD are some of the most significant impacts that I think are not taken into consideration enough by health professionals or discussed in the wider public at the moment.

7a. What do you think of the current treatments and care available for chronic graft versus host disease on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	a. I think the bar to access systemic treatments is very high, and the withdrawal of Ruxolitinib has had a significant negative impact on the available treatments overall. My GvHD has to this point been managed with topical treatments rather than systemic and it is exceptionally difficult to manage. The medicines I am on cause conflicting side effects that can mask the GvHD symptoms, and need to be taken at very specific times of the day due to their inability to be taken together. b. One of my volunteer roles at Anthony Nolan is as an online community champion, so I help to moderate the patients and families forum. The general view on there is similar to mine, with patients exasperated that Ruxolitinib has been withdrawn when people seem to be getting positive results from it. A lot of patients are having to deal with serious GvHD with topical treatments that are not having the right positive impact, but instead still come with sometimes significant side effects. Others are struggling with the invasiveness of the systemic treatments that are on the market today and the mental, social and physical impacts of those.
8. If there are disadvantages for patients of current NHS treatments for chronic graft versus host disease (for example, how they are given or taken, side effects of treatment, and any others) please describe these	I will focus on 2 specific systemic treatments in my answer - ECP and Immunosuppressants (combined with steroids) ECP - Extremely invasive. The need for a central line and for significant hospital time makes this a big deal for patients. Especially since it tends to come at a time where most patients will have had their central line removed, so another one will need to be inserted. The mental impact of something like this can be significant as it can feel like a step backwards in condition rather than forwards.

	Immunosuppressants (for example Ciclosporin) - has obvious disadvantages here with a significantly weaker immune system that will require significant observation and management. Add in high-dose steroids and you have a complex variety of side effects that can hide the improvement in the GvHD - for example and improvement in gut GvHD being offset by the side effects of high dose steroids. As well as this, it comes with a significant quality of life and mental impact due to needing to go back on treatment the patients previously believed they were finished with and the need to go to hospital for any fever as they become a significant Sepsis risk - again.
9a. If there are advantages of belumosudil over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does belumosudil help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	Being able to manage GvHD with a single at home drug would have a significant positive impact. Quality of life would improve instantly from simply not having to manage 8-10 different medicines, as well as driving the cost down for patients. Compared to the existing Systemic treatments, the obvious advantages are in the ease of access to treatment (being able to take it from home, without substantial hospital time) and the different levels of side effects (not deliberately weakening your immune system for an already vulnerable patient.)
10. If there are disadvantages of belumosudil over current treatments on the NHS please describe these. For example, are there any risks with belumosudil? If you are concerned about any potential side effects you have heard about, please describe them and explain why	

11. Are there any groups of patients who might benefit more from belumosudil or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	I can see specific positive impacts for the following groups: 1. Patients with other complex health conditions - not needing to add another regular hospital appointment with the toll that can take would be significant. The more simple and less invasive the treatments are, the easier they will be to deal with. 2. Patients with mobility issues - for some patients, getting to hospital is a real challenge, so the more that can be done to help with at home treatments rather than in-patient or out-patient clinic treatments the better. 3. Patients with cognitive impairments - managing multiple instances of GvHD across your body with different treatments for each one can be a challenge. Particularly when the treatments cause side effects that can seem similar to GvHD symptoms - for people with cognitive impairments, staying on top of all of that would be very hard, they would almost certainly need external help to stay safe. One drug that they can take at home is a significant improvement for this group.
12. Are there any potential equality issues that should be taken into account when considering chronic graft versus host disease and belumosudil? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme	

Find more general information about the Equality Act and equalities issues here.	
13. Are there any other issues that you would like the committee to consider?	The impact on people from a more disadvantaged background, for whom paying for travel to and from hospital, as well as getting time off from work (which they may not get paid for.) A treatment like this would go a long way to helping close the health gap between different socioeconomic groups.

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Overall, because of the combination of GvHD that I currently have, my quality of life is significantly diminished; It has cost me my ability to play sports, exercise, work and makes day to day life challenging and complex with the treatments and planning that needs to go into keeping on top of things
- Social isolation and the mental toll of suffering from Chronic GvHD are some of the most significant impacts that I think are not taken into consideration enough by health professionals or discussed in the wider public at the moment.
- A treatment like this would go a long way to helping close the health gap between different socioeconomic groups. Existing treatments are very invasive, very intensive and require significant side effects, which makes them inaccessible for some patients.
- The ability to treat GvHD with a single treatment that is available to be administered at home would be a very positive step change in the right direction
- My Chronic GvHD is the single hardest thing I have had to deal with mentally in my life, and I believe the impact on patients is severely underestimated. It is the biggest topic of discussion on the Patients & Families Forum that I help moderate and effects both quality of life and lifespan of cancer survivors.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Thursday 20 July**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

NICE National Institute for Health and Care Excellence

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)	Novartis Pharmaceuticals UK Limited
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry	[redacted]

	[redacted]

Table 1 About you

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Evidence for adolescents not available from Rockstars and KD025-208	Yes/No	No comments
Inclusion of concomitant medication costs for belumosudil, such that the intervention for the cost-effectiveness analysis is belumosudil in addition to best available therapy (BAT) (belumosudil+BAT)	Yes/No	No comments
Naïve comparison of belumosudil versus BAT	Yes/No	No comments
Removal of response outcomes from the economic model	Yes/No	No comments
Removal of overall survival benefit for belumosudil+BAT	Yes/No	No comments

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR 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 incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Belumosudil for treating chronic graft versus host disease after 2 or more lines of systemic therapy [ID4021]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is **5pm** on **Thursday 20 July**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Sanofi
Disclosure Please disclose any funding received from the company bringing the treatment to NICE for evaluation or from any of the comparator treatment companies in the last 12 months [Relevant companies are listed in the appraisal stakeholder list.] Please state: • the name of the company • the amount • the purpose of funding including whether it related to a product mentioned in the stakeholder list • whether it is ongoing or has ceased.	**None**
Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response	EAG response
Evidence for adolescents not available from Rockstars and KD025-208	No	Given the unmet need in treatment options for cGVHD across all age groups, and the biological plausibility of using belumosudil in patients aged 12 to 18 years old, we consider it reasonable and appropriate to align the eligible population with that of the MHRA licence.	There are no efficacy and safety data for belumosudil in adolescents, therefore the EAG cannot confirm if the clinical outcomes for adults would be seen in adolescents. However, as mentioned in the EAG report, the EAG's clinical experts agreed, that from a biological perspective, there is no reason why belumosudil would not work as effectively as in adults. Many of the drugs used for cGvHD in adolescents do not have marketing authorisation due to a lack of research.
Inclusion of concomitant medication costs for belumosudil,	Yes	The licence for belumosudil does not state the requirement for any concomitant treatment when used to treat cGVHD. However, we acknowledge that in the clinical trials a proportion of patients did receive concomitant cGVHD medication with	As discussed in the EAG report, the company's scenario which includes concomitant medications for the BAT arm of the model is not considered to

such that the intervention for the cost-effectiveness analysis is belumosudil in addition to best available therapy (BAT) (belumosudil+BAT)		belumosudil, and this may be reflective of NHS clinical practice. Therefore, it is not unreasonable to include these costs in the economic model. In clinical practice we also consider it likely, based on insights from clinical experts, that concomitant treatments would be used alongside BAT to a similar extent to belumosudil. The treatment costs in the BAT arm of the model currently characterise BAT as the equivalent of one therapy but this is modelled as a basket to reflect the options available. This is based on the source trial, REACH-3, which limited co-medication to glucocorticoids and calcineurin inhibitors within the study protocol. However, this study protocol is not reflective of clinical practice and the absence of concomitant treatments in the BAT arm of the model means that their associated costs are likely to be underestimated. The scenario analyses presented during clarification questions included the costs of concomitant therapies according to the pooled analysis of the ROCKstar and KD025-208 studies for the ≥2 LOT subgroup (2021 data cut). As requested by the EAG in their report, we subsequently obtained the updated 2022 data cut for this subgroup and re-ran the scenario analysis, employing the EAG's preferred assumptions (see Table A and B below). This reduced the incremental costs from [redacted] to [redacted], primarily due to the lower proportion of concomitant ECP usage in trial participants. This results in a marginal change in the magnitude of the ICER and has no impact on decision risk.	be clinically valid by the EAG. By definition, BAT is a composition of treatments that reflects established clinical management and thus concomitant medications are implicitly part of the basket. Additionally, the usage of concomitant medications for BAT is not based on evidence as it was not permitted in REACH-3. Nonetheless, inclusion of concomitant medications for BAT is not part of the company or EAG base case. The EAG thanks the company for providing the updated concomitant medication data from ELIPSE. The EAG has produced a separate document with an updated EAG base case which includes the latest data cut from ELIPSE for concomitant medications.

Table A. Model inputs for concomitant treatment (EAG base case and updated data cut)

Concomitant treatment	Pooled analysis, 2021 data cut (n=156)		Pooled analysis, 2022 data cut (n=176)	
	QD (n=81)	BID (n=75)	QD (n=92)	BID (n=84)
ECP n (%)	20 (24.7%)	26 (34.7%)	22 (23.9%)	28 (33.3%)
Sirolimus n (%)	17 (21.0%)	18 (24.0%)	21 (22.8%)	20 (23.8%)
MMF n (%)	11 (13.6%)	2 (2.7%)	11 (12.0%)	3 (3.6%)
Total	48 (59.26%)	46 (61.33%)	54 (58.70%)	51 (60.71%)

Table B. Cost-effectiveness results from deterministic analysis

	EAG base case (2021 data cut)	EAG base case using 2022 data cut
Incremental costs	[redacted]	[redacted]
Incremental QALYs	[redacted]	[redacted]
Cumulative ICER (£/QALY)	Dominant	Dominant

Naïve comparison of belumosudil versus BAT	No	We acknowledge the limitations of using the BAT arm from the REACH-3 study, which was conducted in patients receiving an earlier line of therapy, in a naïve indirect comparison with belumosudil. As described in the company submission, we explored multiple different approaches to address this uncertainty. We agree with the EAG and clinical experts that the naïve comparison is currently the only feasible option to compare clinical outcomes for belumosudil (+BAT) with BAT. There are currently no Phase III RCTs of belumosudil versus BAT in third line cGVHD ongoing either in the UK or internationally. Expert clinicians considered the results of this naïve comparison to be clinically plausible and suitable for decision making in the absence of direct head-to-head data. Therefore, despite the uncertainty around the expected impact of this naïve comparison, the consistently low ICERs presented in the company base case, EAG preferred scenario, and extensive scenario analyses, provide confidence that belumosudil (+BAT) is cost-effective compared to BAT and there is low decision risk.	As mentioned in the EAG report, the only feasible option to compare clinical outcomes for belumosudil+BAT with BAT is via a naïve comparison. However, the EAG emphasises the uncertainty associated with naïve comparisons of clinical outcomes from different trials.
Removal of response outcomes from the economic model	No	We agree with the EAG that failure-free survival is the most clinically relevant outcome for patients and consider it to be the most suitable endpoint for cost-effectiveness modelling. This was consistent with the clinical expert opinion we received during the development of the model and company submission.	The company and the EAG are aligned on this issue. However, the EAG notes that the company has not revised their base case to remove response outcomes from the analysis.

		We also agree with the EAG that excluding response from the model removes a source of unresolvable uncertainty whilst having a minimal impact on the resulting ICERs.	
Removal of overall survival benefit for belumosudil+BAT	No	As stated in our company submission, we have no direct data which demonstrate a relative overall survival benefit for belumosudil vs. standard care. Considering the uncertainty around relative overall survival between the two treatments, we do not consider it unreasonable for the EAG to remove the overall survival benefit originally included in the company submission model. Whilst the impact of this change on the ICER is not insignificant, it is worth reflecting that if the original assumption of overall survival benefit for belumosudil is maintained but all the other EAG adjustments to the company model are implemented the ICER remains under £30,000/QALY. If the OS benefit is removed from the company base case belumosudil is the dominant strategy. Whilst the OS assumption is undoubtedly the biggest driver of the ICER, taken together these ICER estimates indicate that the inclusion or exclusion of the OS benefit does not introduce a risk of decision error in judging belumosudil to be a cost-effective option.	The company and the EAG are aligned on this issue. However, the EAG notes that the company has not revised their base case to exclude an overall survival benefit for belumosudil+BAT from the analysis.

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response	EAG response

Additional issue 1: Subsequent treatment of cGVHD duration may exceed 5 years in some patients	Section 4.2.7.6 – pages 122-123 Section 4.2.7.9 – pages 130-132	No	The duration of subsequent treatments for patients for whom third-line treatment fails and the proportion of these patients on treatment was identified as a secondary issue for the EAG and a primary driver of cost-effectiveness in the model. The EAG preferred base case assumes a 5-year subsequent treatment duration with only 60% of patients remaining on treatment post-failure. We consider this to be an underestimate. Assuming that 40% of patients do not receive any subsequent treatment and that the remaining 60% stop receiving any treatment for cGVHD by 5 years does not seem clinically plausible considering the chronic nature of the disease and its likely severity at this stage in the treatment pathway. Feedback we received from clinicians was that cGVHD patients were unlikely to remain on a *single third line therapy* for more than 5 years. However, after third line treatment failure, they could subsequently cycle through multiple lines of treatment over the remainder of their lifetime. This is reflected in the basket of subsequent treatments modelled in the company base case, which assumes that patients remain on some form of treatment for 60% of their remaining life years after third line (or later) treatment failure. The 60% proportion is intended to reflect the	No new evidence presented. The EAG notes that in its base case, subsequent treatments are modelled as a basket of treatments (with the proportions of treatments unchanged from the company's base case) and not a single third line therapy as implied in the company response. Additionally, the EAG maintained the company's assumption that 60% of patients on subsequent treatment, to account for various changes in treatment (such as treatment weaning, treatment pauses and treatment discontinuations). Thus, the only difference between the EAG and company base case is the duration of subsequent treatment (five years versus lifetime). As such, the EAG maintains its position that assuming patients will be on subsequent treatments for the remainder of their lifetime is a clinically implausible assumption. However, as implied by the company in their TE response, both the company and EAG base case results are below the £20-30,000 cost-effectiveness threshold which use different assumptions for modelling duration of subsequent therapies.

			overall proportion of time spent on treatment, rather than the proportion of patients going on to receive a subsequent treatment, thereby accounting for adherence issues, treatment gaps, and or therapy discontinuations. Despite the limited available evidence to accurately model post-failure treatment costs, the range of assumptions employed by the company and EAG consistently result in ICERs below the cost-effectiveness threshold.	

Additional issue 2: Choice of utility value for treatment failure	Section 4.2.6 – pages 105-109, 112-114	No	The midpoint utility value of 0.608 selected by the EAG for patients in the "failure-new systemic therapy" health state is associated with considerable uncertainty. It was derived in part from a published cGVHD economic analysis by Crespo *et al*, which reported a utility value for a 'Progression' health state of 0.696. This figure, quoted in Crespo *et al*, could not be traced back to a value in the source study cited by Crespo *et al*, and despite contacting the authors we were unable to establish how they derived this estimate. The other data point informing the midpoint value (0.52) was collected from European patients with cGVHD. Although the sample size was very small (n=10), it is at least a verifiable source and is likely to reflect the utility of patients more accurately at this point in the treatment pathway. Indeed, it may be conservative as it is above the value that we validated for this health state (0.479) in an advisory board and subsequent individual consultancies with clinicians. It should be remembered that this utility is carried forward for the remaining time in the model and so should reflect the worsening stages of the disease. In the general population vignette-based utility elicitation we conducted, the utility value for this health state was substantially lower than any of the alternative sources. Whilst seemingly extreme, this study was conducted following Decision Support Unit's best practice	During technical engagement, clinical expert Adrian Bloor provided a response in which they advised that "*assumptions over utility of 4th line therapy (which is equated to recurrent malignancy) are very weak and not really justified*". As such, the EAG considers that this additional clinical expert response reinforces the EAG's clinical advice that failure due to recurrent malignancy is a far worse outcome for patients than failure due to a change in cGvHD treatment. Thus, the assumption that the utility value for recurrently malignancy and new cGvHD systemic therapy being the same is unreasonable. As such, the EAG considers that its base case assumption of using a utility value of 0.608 for failure - new cGvHD systemic therapy represents a more plausible assumption.

			recommendations in line with the NICE process and methods guide and reflects the UK societal valuation of this health state. Therefore, it should not be completely discounted in this appraisal. This is an especially important parameter because the applicability of the severity modifier rests upon the severity of disease defined by quality of life decrement in the later parts of the patient journey. Patient and clinician testimony that we have received indicates that cGVHD is an extremely burdensome disease affecting all areas of a patient's life, some of which, like social functioning will not be fully captured in the QALY. Therefore, we disagree with the EAG analysis which removes the severity modifier. We consider this an important point for deliberation and would encourage the patient perspective to be incorporated during Committee discussions. Acknowledging the uncertainty of this parameter, we consider the utility value in the EAG base case to be an upper estimate of the plausible value, resulting in an underestimation of the cost-effectiveness of belumosudil. Nonetheless, even with this upper estimate the ICER remains below the commonly accepted threshold and so decision risk is low.	

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR 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 incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy [ID4021]

Updated EAG base case post technical engagement

August 2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135825.

Table 1. EAG's preferred model assumptions (deterministic) – belumosudil+BAT versus BAT

Preferred assumption	Section in EAG report	Incremental costs	Incremental QALYs	ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case – post clarification	-	[redacted]	[redacted]	3,571	-
Removal of response outcomes – company scenario	4.2.4.5	[redacted]	[redacted]	3,434	£3,434
Removal of OS benefit	4.2.4.5	[redacted]	[redacted]	Dominant	Dominant
Concomitant medication costs for belumosudil only	4.2.7.9	[redacted]	[redacted]	16,745	Dominant
Removal of cost of background therapies*	4.2.7.9	[redacted]	[redacted]	Dominant	Dominant
KM TTD data for belumosudil	4.2.7.9	[redacted]	[redacted]	2,047	Dominant
Exponential distribution for BAT TTD	4.2.7.9	[redacted]	[redacted]	5,174	Dominant
Removal of accommodation costs for patients in ECP	4.2.7.9	[redacted]	[redacted]	4,080	Dominant
Maximum subsequent treatment duration of five years (except for rituximab)	4.2.7.9	[redacted]	[redacted]	7,638	Dominant
Midpoint utility value of 0.608 for failure new cGvHD systemic therapy utility value	4.2.6.5	[redacted]	[redacted]	4,213	Dominant
Caregiver disutility for failure – new cGvHD systemic therapy equal to failure-free (PR/LR)	4.2.6.5	[redacted]	[redacted]	4,065	Dominant
Disutility and duration for central line-related infection based on disutility for infections and infestations from TA689[36]	4.2.6.5	[redacted]	[redacted]	3,568	Dominant
Removal of IV disutility for BAT	4.2.6.5	[redacted]	[redacted]	3,613	Dominant

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; EAG, External Assessment Group; ECP, extracorporeal photopheresis; ICER, incremental cost-effectiveness ratio; IV, intravenous; KM, Kaplan-Meier; LR, lack of response; OS, overall survival; PR, partial response; QALY, quality adjusted life year; TTD, time to treatment discontinuation

*Scenario combines the assumptions of removal of OS benefit, concomitant medication costs only for belumosudil and removal of cost of background therapies.

Table 63. EAG's base case results – post technical engagement

Interventions	Total Costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
BAT	[redacted]	[redacted]	[redacted]	-	-	-	-
Belumosudil+ BAT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Probabilistic results							

BAT	[redacted]	[redacted]	[redacted]	-	-	-	-
Belumosudil+ BAT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Abbreviations: BAT, best available therapy; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year.

1 Introduction

2 References

1. Centre for Reviews and Dissemination (CRD). CRD's guidance for undertaking reviews in healthcare. Available at https://wwwyorkacuk/media/crd/Systematic_Reviewspdf (Accessed 21 March 2016) 2011.

Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy [ID4021]

Committee requested scenarios

September 2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135825.

1 Introduction

Prior to the first appraisal committee meeting (ACM1) for belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy, issues were raised with by the committee lead team via NICE and several additional scenarios around the External Assessment Group (EAG) base case were requested.

A summary of the issues raised by the lead team were as follows:

- The company and EAG preferred utility value for failure – new cGvHD systemic therapy, was considered to be too low considering that the reason for failure is a change in treatment. As such, the lead team requested a scenario that assumes the utility value for failure – new cGvHD systemic therapy, is set equal to the utility value for failure-free – lack of response (i.e. [redacted]).
- The assumption of a constant disease management cost for failure – new cGvHD systemic therapy, was considered be a pessimistic assumption. Instead, the lead team requested to see a scenario where the disease management cost for failure – new cGvHD systemic therapy, had a linear decline over five years equal to the year five failure-free partial/lack of response (PR/LR) disease management cost. However, the EAG highlights that the year 5 failure-free PR/LR disease management cost is assumed to be the same as the disease management for complete responders (CR). As such, there is an inherent assumption in the lead team request that over five years, patients in the failure – new cGvHD systemic therapy health state will accrue the same health state costs as a patient who is failure-free and experience a complete response to treatment.

The lead team raised two other issues which they considered warranted further modelling of the main treatment outcomes. However, due to a paucity of time, the EAG could not implement these changes but discusses these issues in Section 2, and encourages the committee to seek advice from the clinical experts on these issues at ACM1.

- Extrapolation of REACH-3 failure-free survival (FFS) for the best available therapy (BAT) arm. In the EAG report it was highlighted that in REACH-3, 38% of BAT patients crossed over to ruxolitinib at week 24. The EAG considered that crossover would have an impact on overall survival (OS) outcomes. However, the lead team considered that due to the open-label

nature of REACH-3, there could have been investigator bias when to change treatment for patients in the BAT arm of the trial, resulting in biased FFS results.

- The suitability of FFS as an outcome of interest for chronic graft versus host disease (cGvHD).

Each of the issues is discussed in Section 2 and results of the EAG's scenarios around its preferred base case are presented in Section 3. A confidential discount is available for rituximab, as such the EAG has produced a confidential appendix to this document.

2 EAG discussion of committee lead team issues

2.1 Utility value for failure – new cGvHD systemic therapy

In the External Assessment Group (EAG) report, the EAG considered that the utility value for failure – new cGvHD systemic therapy, is key driver of quality-adjusted life-years (QALYs) in the model, as patients in the best available therapy (BAT) arm spend the majority of their time in this health state. In the company's base case, the utility value for failure – new cGvHD systemic therapy, is assumed to be equal to the utility value for failure – recurrent malignancy (0.479).

During the clarification stage, the EAG requested, and the company provided, utility data for new cGvHD systemic therapy and recurrent malignancy (Table 40 of the company clarification response). The resulting utility value of [redacted], based on 69 observations from 22 patients, for failure – new cGvHD systemic therapy was higher than the utility value for failure-free ([redacted]), which was based on 1,197 observations from 140 patients.

The EAG considers that there is a high degree of uncertainty around the utility value for failure – new cGvHD systemic therapy derived from ROCKStar due to the limited number of observations. Additionally, as advised by its clinical experts, the EAG considers that patients who require a change in treatment for their cGvHD represent a population with a poorer health-related quality of life (HRQoL) and a poorer prognosis than those that remain failure free, as the failure event indicates more advanced cGvHD. As such, the EAG preferred to use a utility value of 0.608, which was a midpoint value derived from Crespo *et al.*,[1] and the Adelphi disease specific programme (DSP study).

The lead team considered that because the utility analysis from ROCKStar demonstrated that the utility value for failure – new cGvHD systemic therapy was similar to the failure-free utility, it is likely that the impact on HRQoL was likely to be minimal and potentially similar to patients who are failure-free and have a lack of response. As such, the lead team requested a scenario that assumes the utility value for failure – new cGvHD systemic therapy is set equal to the utility value for failure-free – lack of response ([redacted]). Results of the scenario are presented in Section 3.

Based on advice from its clinical experts and acknowledged by the lead team, a change in cGvHD systemic therapy does indicate that a patient's disease has progressed and could potentially indicate being refractory to treatment. The EAG's clinical experts advised that failure events in clinical practice, whether because of new treatment or recurrent malignancy have a negative impact on

patients, with recurrent malignancy being the more severe outcome for patients. Furthermore, the treatment pathway after third line is limited and consists of best available therapy.

The EAG does not consider that the company's base case assumption that quality of life for all failure patients is the same (0.479), nor does it agree with the lead team that quality of life for patients who have had a change in treatment would be the same as a patient who is failure-free but has a lack of response to treatment ([redacted]). The EAG considers its base case assumption of a utility value of 0.608 sits between the company's base case assumption and the lead team's assumption and urges the committee to seek further advice on the impact of failure events on HRQoL from clinical and patient experts during the first appraisal committee meeting (ACM1).

2.2 Disease management costs for failure – new cGvHD systemic therapy

In the company base case, disease management state costs in the model were derived from Hospital Episode Statistics (HES) data. In the EAG report, the EAG acknowledged that disease management costs are a primary driver of cost-effectiveness in the model but considers the company's HES study to be thorough, with data reflecting the UK population. Additionally, the EAG's clinical experts were satisfied with the underlying assumptions used to estimate costs from the HES data. As a reminder, the company's assumptions for the disease management costs included in the model are as follows:

- **Patients in the failure-free health state with CR**: assumed to be the mean cost incurred by haematopoietic stem cell transplant (HSCT) patients without cGvHD in the HES study throughout the time horizon of the model.
- **Patients in the failure-free health state with PR and LR**: assumed to be the mean cost incurred by all HSCT patients with cGvHD in the HES study in the first year, with a linear decrease in each year to reach the disease management cost of patients with CR in the fifth year. The model assumes that patients remaining failure-free incur the same costs regardless of response status after the fifth year.
- **Patients in the failure state with a new systemic therapy**: assumed to incur the mean cost of HSCT patients with two or more records of high-cost therapy in the HES study.
- **Patients in the failure state with recurrent malignancy**: These were not available from the HES study and so were sourced from TA642[2] that included the total costs incurred by patients with AML-related inpatient admissions, ICU, emergency department, outpatient visits, diagnostic procedures, lab tests, and blood transfusions. Acute myeloid leukaemia was the most common underlying malignancy in ROCKstar (40.9%).

The assumption of a constant disease management cost for failure – new cGvHD systemic therapy was considered by the lead team to be a pessimistic assumption. Instead, the lead team requested to see a scenario where the disease management cost for failure – new cGvHD systemic therapy had a linear decline over five years to equal the year five failure-free partial/lack of response (PR/LR) disease management cost (see Table 1).

Table 1. Disease management costs per cycle per year

Health states	Mean cost per cycle per year					Source
	1st year	2nd year	3rd year	4th year	≥5th year	
Failure-free						
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Partial response and Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Failure						
New cGvHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database[59]
Recurrent malignancy	£2,719.46	£2,719.46	£2,719.46	£2,719.46	£2,719.46	NICE TA642[41]
Committee requested scenario - New cGvHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**First year cost is the same as the company base case with a linear decline down to the ≥5th year cost for failure-free (CR/PR/LR)**

Abbreviations: cGvHD, chronic graft versus host disease; CS, company submission.

The EAG highlights that the year 5 failure-free PR/LR disease management cost in the lead team's requested scenario is assumed to be the same as the disease management for complete responders (CR). As such, there is an inherent assumption in the lead team's scenario that over five years, patients in the failure – new cGvHD systemic therapy health state will accrue the same health state costs as a patient who is failure-free and experience a complete response to treatment. Furthermore, the lead team acknowledged that it is unknown what percentage of failure patients who start a new treatment will experience a resolution in their cGvHD or how long it will take before they experience another failure event. However, in the EAG's base case, life years spent in the failure-free health state for BAT patients is approximately [redacted] and is unlikely to be longer than that when patients progress to their next therapy.

The EAG considers that the requested scenario may not be clinically plausible and is potentially biased in favour of BAT. The EAG also reiterates its clinical experts view that cGvHD is aggressive and

difficult to treat disease as it affects multiple organs, which becomes harder to treat the if a patient has had number of prior therapies. Thus, the EAG considers that it is clinically plausible costs will increase as a result of patients starting a new therapy.

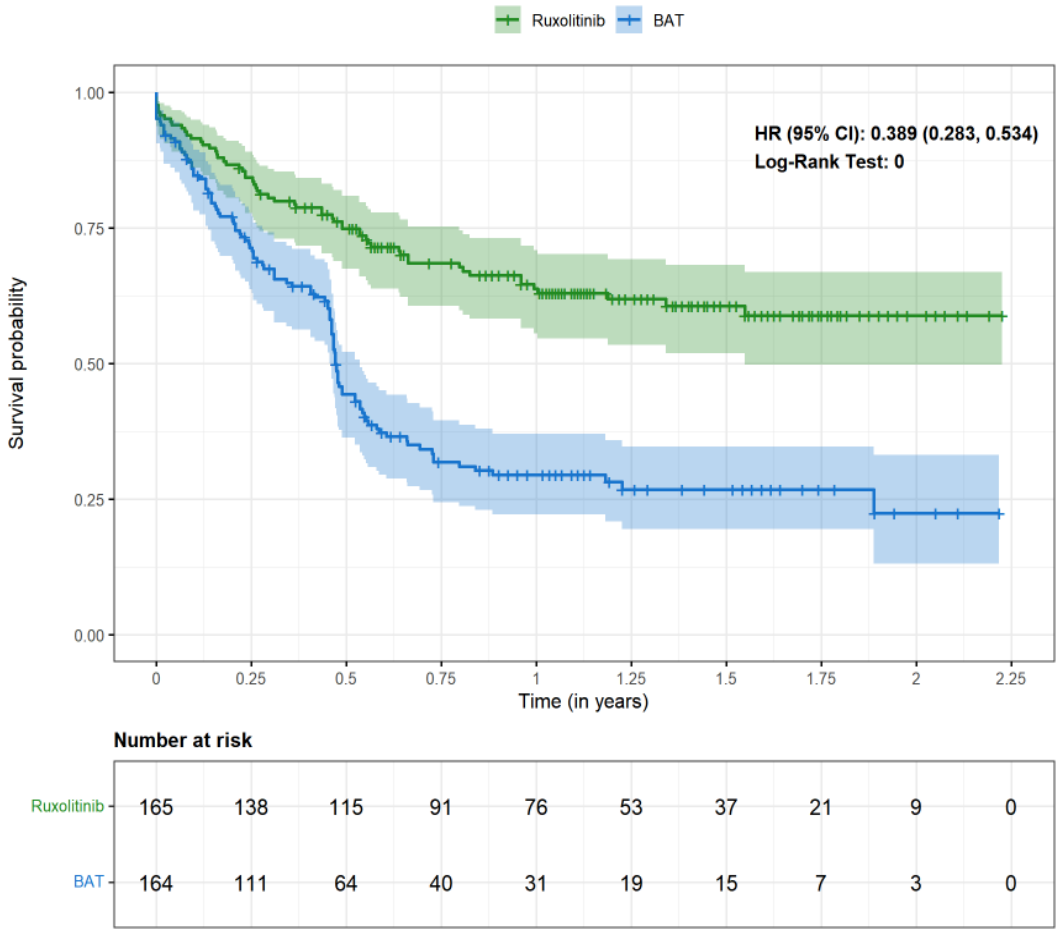
Nonetheless, the EAG ran the requested scenario, assuming a linear decline in the disease management cost for failure – new cGvHD systemic therapy over five years to equal the year five failure-free partial/lack of response (PR/LR) disease management cost (see Table 1). However, the scenario assumed 100% of failure – new cGvHD systemic therapy patients incur a reduction in costs to manage their disease, which may not be clinically plausible. As such, the EAG explored scenarios ranging from 25%-75% of failure – new cGvHD systemic therapy patients experiencing a reduction in disease management costs. Furthermore, the EAG highlights that in its base case, the duration of subsequent treatments was assumed to be five years. Thus, in the long-term disease management costs are likely to increase for those patients who still have cGvHD.

2.3 Extrapolation of REACH-3 failure-free survival for the best available therapy arm

In REACH-3, for patients who did not have or maintain a complete or partial response, had unacceptable side effects from a control therapy, or had a flare of chronic GVHD, crossover from control therapy to ruxolitinib could occur on or after week 24.[3] Additionally, patients in the control group who had a complete or partial response at week 24 could not cross over to ruxolitinib unless they had disease progression, mixed response, or unacceptable side effects from the control therapy.[3] Overall, 38% of BAT patients crossed over to ruxolitinib on or after week 24. The EAG notes that response in REACH-3 was defined as best overall response (complete or partial) up to week 24 and that failure-free survival (FFS) was defined as relapse or recurrence of underlying disease or death due to underlying disease, non-relapse mortality, or addition or initiation of another systemic therapy for cGvHD, whichever came first.

Figure 1 presents FFS from REACH-3. REACH-3 was an open-label study, and the lead team were concerned that the drop in the Kaplan-Meier (KM) curve at approximately six months could be due to investigator bias to change treatment to ruxolitinib for patients in the BAT arm of the trial, resulting in biased FFS results.

Figure 1. Kaplan-Meier curve for FFS in the BAT arm in REACH-3[3] (reproduced from Figure 21 of the company submission)*



*In the company submission, the company note that Figure 21 was created based on reconstructed individual patient-level data (RIPD) using the Guyot *et al.* algorithm.

The EAG considered that crossover would have an impact on overall survival (OS) outcomes. Additionally, in the CADTH review of ruxolitinib, it was noted that crossover may confound OS but the relationship between crossover and FFS was not highlighted as a cause for concern.[4]

The lead team requested a scenario where FFS KM data for BAT from REACH-3 was truncated at week 24 and then extrapolated for use in the economic model. However, the EAG were unable to supply the scenario due to a paucity of time. The EAG considers that it is plausible, albeit unknowable, that there could be a "halo" effect around 24 weeks where BAT patients in REACH-3 who were failing prior to this time did not initiate new treatment and were "switched late" so that they could receive ruxolitinib and some failing but not yet failed patients could have been "switched early" when ruxolitinib became available at 24 weeks.

The EAG considers that if these assumptions are true, then truncating the KM curve at 24 weeks is very likely to give a clinically implausible overestimation of the treatment effect for BAT. Thus, the base case estimation of FFS may not reflect an "unbiased" FFS curve but it is likely to be less biased than an extrapolation of the BAT FFS KM data truncated at 24 weeks. The EAG considers that the assumptions of delaying or early initiation of a change in cGvHD treatment depends on the affected organs of the patient and whether treatment choices are time sensitive. The EAG notes that its clinical experts consider that mean FFS for BAT patients is likely to be one year, [redacted]. As such, the EAG considers clinical expert input during ACM1 is critical in understanding disease progression and clinical decision making around cGvHD treatments for patients.

Furthermore, the EAG considers that a scenario where the FFS curve for BAT is truncated would need careful consideration by the committee to interpret the resulting incremental cost-effectiveness ratio (ICER). The EAG recommends that the committee seeks additional advice from the clinical experts at ACM1, whether or not they consider that the extrapolated FFS for BAT based on truncated data from REACH-3 is clinically plausible or implausible. If the clinical experts consider it to be clinically implausible, which the EAG considers is quite likely, the scenario wouldn't produce an ICER suitable for decision making.

2.4 The suitability of failure-free survival as an outcome of interest for chronic graft versus host disease

In the NICE final scope, FFS was listed as an outcome of interest. However, the lead team were concerned that in the US National Institutes of Health (NIH) consensus guidelines, it was stated that new treatment decisions are not always driven by lack of efficacy and are subject to bias, making FFS as a primary endpoint inadequate for regulatory purposes.[5] Additionally, the lead team considered the limitations of FFS highlight in studies by Inamoto *et al.* which aimed to investigate failure for GvHD and whether or not FFS was clinically meaningful outcome.[6, 7] The authors considered that FFS is a potentially useful, efficient, and robust basis for interpreting results of initial treatment of chronic GVHD and that it can be used as a robust benchmark for designing and interpreting future phase 2 trials of initial treatment of chronic GVHD.[6, 7] However, the authors outlined that FFS does not give any direct information about changes in GVHD-related symptoms, activity, damage, or disability and results with this end point require careful interpretation in nonblinded trials as decisions to change treatment may be subjective.[6, 7]

The EAG's clinical experts considered that FFS to be a clinically relevant outcome and was deemed suitable for use in the economic model. Additionally, in the CADTH report for ruxolitinib, clinical experts considered FFS was a clinically meaningful endpoint.[4] It should be noted that in the Scottish Medicines Consortium (SMC) appraisal for belumosudil, no concerns were raised about the use of FFS in the economic model. Furthermore, as ruxolitinib was a non-submission for NICE, there is no published opinion from NICE on what would be appropriate within a NICE appraisal.

In conclusion, the EAG does not consider that it is definitive that: a) FFS is always an inappropriate outcome for cGvHD (and it was listed as an outcome of interest in the NICE final scope); and b) a model that includes FFS is always inappropriate. Additionally, the EAG recommends that committee seeks further advice from the clinical experts at ACM1 to determine the suitability of FFS as an outcome of interest for use in the economic model for this appraisal.

3 Committee lead team requested scenarios

As discussed in Section 2, several scenarios around the EAG base case were requested by the lead team and results are presented in Table 2.

Table 2. Results of committee lead team requested scenarios

	Results per patient	Belumosudil	BAT	Incremental value
0	**EAG base case – post technical engagement**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
1	**Utility value for failure – new cGvHD systemic therapy equal to failure-free lack of response ([redacted])**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
2	**Linear decline of Failure new treatment disease management costs to Failure-free CR disease management cost**			
	Total costs (£)	[redacted]	172,780	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	80,173
3	**75% of Failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	188,514	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	52,949
4	**50% of Failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	204,248	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	25,725
5	**25% of Failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	219,982	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
6	**Scenario 1+2**			
	Total costs (£)	[redacted]	172,780	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	310,826

8	**Scenario 1+3**			
	Total costs (£)	[redacted]	188,514	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	205,280
7	**Scenario 1+4**			
	Total costs (£)	[redacted]	204,248	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	99,734
9	**Scenario 1+5**			
	Total costs (£)	[redacted]	219,982	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; CR, complete response; EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year.

4 References

1. Crespo C, Pérez-Simón JA, Rodríguez JM, Sierra J, Brosa M. Development of a population-based cost-effectiveness model of chronic graft-versus-host disease in Spain. *Clin Ther* 2012; **34**: 1774-87.
2. National Institute for Health and Care Excellence. Gilteritinib for treating relapsed or refractory acute myeloid leukaemia (TA642), 2020. Available from: https://www.nice.org.uk/guidance/ta642. Date accessed: April 2023.
3. Zeiser R, Polverelli N, Ram R, Hashmi SK, Chakraverty R, Middeke JM, et al. Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. *The New England journal of medicine* 2021; **385**: 228-38.
4. CADTH. CADTH reimbursement review: Ruxolitinib (Jakavi). *Canadian Journal of Health Technologies* 2022; **2022**: 1-220.
5. Martin PJ, Lee SJ, Przepiorka D, Horowitz MM, Koreth J, Vogelsang GB, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. The 2014 Clinical Trial Design Working Group Report. *Biol Blood Marrow Transplant* 2015; **21**: 1343-59.
6. Inamoto Y, Flowers ME, Sandmaier BM, Aki SZ, Carpenter PA, Lee SJ, et al. Failure-free survival after initial systemic treatment of chronic graft-versus-host disease. *Blood* 2014; **124**: 1363-71.
7. Inamoto Y, Storer BE, Lee SJ, Carpenter PA, Sandmaier BM, Flowers ME, et al. Failure-free survival after second-line systemic treatment of chronic graft-versus-host disease. *Blood* 2013; **121**: 2340-6.

Belumosudil for treating chronic graft versus host disease after two or more lines of systemic therapy [ID4021]

Committee requested scenarios

September 2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135825.

1 Additional committee lead team scenarios

At the request of the committee lead team, the External Assessment Group (EAG) has run a scenario which uses the hospital episodes statistics (HES) estimate for cGVHD patients with first high-cost therapy, presented in Table 4 of the company submission, for the failure-free year 1 health state cost. The assumption of linear decline in the failure-free health state cost to the year 5 failure-free complete response health state cost is maintained. In addition, the committee lead team requested the additional scenario to be combined with previously requested scenarios, summarised as follows:

- Utility value for failure – new cGvHD systemic therapy is set equal to the utility value for failure-free – lack of response ([redacted]).
- Disease management cost for failure – new cGvHD systemic therapy had a linear decline over five years to equal the year five failure-free partial/lack of response (PR/LR) disease management cost (see Table 1).

Table 1 presents the costs for the committee lead team requested scenario. The EAG notes, that in its preferred base case, response is excluded from the model and the failure-free health state cost is based on the failure-free partial/lack of response cost presented in Table 1.

The EAG considers that the lead team requested scenarios aims to explore the impact of increasing the cost of failure-free survival such that the cost difference of failure relating to a change in treatment is reduced. However, there are a few considerations for when using the HES cost associated with first high-cost therapy and the scenarios combining a linear decline in failure – new cGvHD systemic therapy:

- The marketing authorisation restricts the use of belumosudil to patients who have received at least two prior systemic therapies. Therefore, assuming the cost of first high-cost therapy (which is assumed by the company to be ECP, pentostatin, rituximab, ruxolitinib, imatinib) might not be appropriate, as at third-line, patients would have already received a high-cost treatment.
- In the EAG base case, the intervention is belumosudil+BAT, as opposed to belumosudil monotherapy assumed in the company's base case, thus restricting the cost to patients with one high-cost treatment may not be reflective of the resource use these patients incur.

- By year 5, patients in either the failure-free health state and failure-new cGvHD systemic therapy health state incur the same health state costs as a person who is failure-free with a complete response (i.e. their cGvHD is resolved). The EAG considers that this assumption may not be clinically plausible as based in advice from the EAG's clinical experts, patients who fail treatment are in a clinically worse position than patients who are failure-free. However, the EAG acknowledges there will be a proportion of patients in their next line of treatment who maybe failure-free for that line of treatment but notes that next line of treatment is best available therapy (BAT) and so failure-free survival may be shorter than at third-line.

Table 1. Disease management costs per cycle per year

Health states	Mean cost per cycle per year					Source
	1st year	2nd year	3rd year	4th year	≥5th year	
Failure-free						
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database
Partial response and Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database
Committee requested scenario - Partial response and Lack of response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database. Chronic GVHD patients with first high-cost therapy (Table 4 of the CS)
Failure						
New cGvHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	HES database
Recurrent malignancy	£2,719.46	£2,719.46	£2,719.46	£2,719.46	£2,719.46	NICE TA642
Committee requested scenario - New cGvHD systemic therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**First year cost is the same as the company base case with a linear decline down to the ≥5th year cost for failure-free (CR/PR/LR)**

Abbreviations: cGvHD, chronic graft versus host disease; CS, company submission.

The results of **all** the committee lead team requested scenarios are presented in Table 2.

Table 2. Results of committee lead team requested scenarios

	Results per patient	Belumosudil	BAT	Incremental value
0	**EAG base case – post technical engagement**			

	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
1	**Utility value for failure – new cGvHD systemic therapy equal to failure-free lack of response ([redacted])**			
	Total costs (£)	[redacted]	235,716	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
2	**Linear decline of failure new treatment disease management costs to failure-free CR disease management cost**			
	Total costs (£)	[redacted]	172,780	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	80,173
3	**75% of failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	188,514	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	52,949
4	**50% of failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	204,248	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	25,725
5	**25% of failure new treatment patients experiencing a linear reduction in disease management cost to failure-free disease management cost**			
	Total costs (£)	[redacted]	219,982	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
6	**Scenario 1+2**			
	Total costs (£)	[redacted]	172,780	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	310,826
7	**Scenario 1+3**			
	Total costs (£)	[redacted]	188,514	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	205,280
8	**Scenario 1+4**			
	Total costs (£)	[redacted]	204,248	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	99,734

9	**Scenario 1+5**			
	Total costs (£)	[redacted]	219,982	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
10	**Failure-free health state cost assumed to be cost of first high-cost therapy from HES + linear reduction in cost.**			
	Total costs (£)	[redacted]	239,672	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	Dominant
11	**Scenario 10+6**			
	Total costs (£)	[redacted]	176,737	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	336,556
12	**Scenario 10+7**			
	Total costs (£)	[redacted]	192,471	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	231,010
13	**Scenario 10+8**			
	Total costs (£)	[redacted]	208,205	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	125,464
13	**Scenario 10+9**			
	Total costs (£)	[redacted]	223,939	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	19,918

Abbreviations: BAT, best available therapy; cGvHD, chronic graft versus host disease; CR, complete response; EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year.