Single Technology Appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

Committee Papers

© National Institute for Health and Care Excellence 2023. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

Contents:

The following documents are made available to stakeholders:

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

Pre-technical engagement documents

1. **Company submission** from Roche

2. **Clarification questions and company responses**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. Lymphoma Action

4. **External Assessment Report** prepared by Warwick Evidence

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

Post-technical engagement documents

6. **Technical engagement response from company**
   a. Technical engagement response form
   b. Updated analysis

   N.B. Mosunetuzumab vs Flatiron report submitted by the company but all contents are academic in confidence, so it is not included in these papers.

7. **Technical engagement responses and statements from experts:**
   a. Dr Kim Linton – clinical expert, nominated by Roche
   b. Dr Mark Bishton – clinical expert, nominated by NCRI-ACP-RCP-RCR
   c. Zoe Drymoussi – patient expert, nominated by the Follicular Lymphoma Foundation

© National Institute for Health and Care Excellence 2023. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

8. **External Assessment Report critique of company response to technical engagement** prepared by Warwick Evidence
    a. Critique of company response to technical engagement
    b. EAG revised base case tables
    c. Additional EAG plots
    d. Updated Table 14 in EAG critique of company response to technical engagement

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 2023. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

ID3931: Mosunetuzumab for treating relapsed or refractory follicular lymphoma

Document B

Company evidence submission

May 2022

File name	Version	Contains confidential information	Date
ID3931_Mosun RR FL_Doc B_ACIC REDACTED_v2.0	2.0	No	23 June 2022

© Roche (2022). All rights reserved

Contents

© Roche (2022). All rights reserved

© Roche (2022). All rights reserved

© Roche (2022). All rights reserved

Tables

© Roche (2022). All rights reserved

Figures

© Roche (2022). All rights reserved

© Roche (2022). All rights reserved

© Roche (2022). All rights reserved

Appendices

The following appendices are provided as separate files accompanying this company submission:

1. Appendix C: Draft Summary of Product Characteristics and UK public assessment report
2. Appendix D: Report from the systematic literature review of published clinical evidence, described in Section B.2.1
3. Appendix E: Report from the indirect treatment comparison between mosunetuzumab and relevant comparators, described in Section B.2.9
4. Appendix F: Report from the systematic literature review of published cost-effectiveness evidence
5. Appendix G: Report from the systematic literature review of health-related quality of life studies
6. Appendix H: Cost and healthcare resource identification, measurement and valuation
7. Appendix I: Clinical outcomes and disaggregated results from the model
8. Appendix J: Price details of treatments included in the submission
9. Appendix K: Checklist of confidential information

Note that no appendices detailing subgroup analyses or adverse events are provided with this submission, as all relevant information is included within the main body of this document.

© Roche (2022). All rights reserved

Abbreviations

ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse events of special interest
alloSCT	Allogeneic stem cell transplant
ALT	Alanine transaminase
ASCT	Autologous stem cell transplantation
aSMD	Absolute standardised mean difference
ASTCT	American Society for Transplantation and Cellular Therapy
ATE	Average treatment effect
BHS	British Society for Haematology
BMI	Body mass index
BNLI	British National Lymphoma Investigation
BSC	Best supportive care
CAR-T	Chimeric antigen receptor T-cell
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CNS	Central nervous system
CR	Complete response
CRS	Cytokine release syndrome
CT	Computed tomography
DLBCL	Diffuse large B-cell lymphoma
DOR	Duration of response
ECG	Electrocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Score
EMA	European Medicines Agency
EORTC QLQ-C30	European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire
EPAR	European public assessment report
ESS	Effective sample size
FACT-Lym	Functional Assessment of Cancer Therapy – Lymphoma
ESMO	European Society for Medical Oncology
Fab	Fragment antigen-binding
FDG-PET	Fluorodeoxyglucose-positron emission tomography
FL	Follicular lymphoma
FLIPI	Follicular Lymphoma International Prognostic Index
G-CSF	Granulocyte colony stimulating factor
GELF	Groupe d'Etude des Lymphomes Folliculaires
Hb	Haemoglobin
HMRN	Haematological Malignancy Research Network
HR	Hazard ratio
HRQoL	Health-related quality of life

© Roche (2022). All rights reserved

Ig	Immunoglobulin
IPD	Individual patient data
IPTW	Inverse probability of treatment weighting
IRF	Independent review facility
ISRT	Involved-site radiotherapy
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenously
LDH	Lactate dehydrogenase
LLN	Lower limit of normal
MAIC	Matching-adjusted indirect comparisons
MCL	Mantle cell lymphoma
MHRA	Medicines and Healthcare Products Regulatory Agency
NE	Not evaluable
NHL	Non-Hodgkin lymphoma
ORR	Objective response rate
OR	Odds ratio
OS	Overall survival
PD	Progressive disease
PET-CT	Positron emission tomography - computed tomography
PFS	Progression-free survival
PI3K	Phosphoinositide 3-kinase
PICOS	Population, Intervention, Comparator, Outcomes, and Study Design
POD24	Progression of disease within 24 months from front-line therapy
PR	Partial response
PRO	Patient reported outcome
Q3W	Every 3 weeks
R^2	Rituximab with lenalidomide
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
RCT	Randomised controlled trials
R/R	Relapsed / refractory
RWD	Real-world data
SAE	Serious adverse event
SD	Stable disease
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
TA	Technology appraisal
tFL	Transformed follicular lymphoma
VAS	Visual analogue scale

© Roche (2022). All rights reserved

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

B.1.1 Decision problem

Mosunetuzumab does not currently have a marketing authorisation in the UK, so the precise wording of the marketing authorisation is not yet available. This submission covers the technology's full anticipated marketing authorisation for this indication, i.e., [redacted]

© Roche (2022). All rights reserved

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with relapsed or refractory follicular lymphoma	[redacted]	Clinical data for mosunetuzumab is currently only available in FL patients who had received at least 2 prior systemic therapies. As such, this population is included in the submission, in line with the anticipated marketing authorisation.
Intervention	Mosunetuzumab	Mosunetuzumab	No difference
Comparator(s)	Established clinical management without mosunetuzumab. Treatment choice will depend on previous treatments, and how effective those treatments were. • Obinutuzumab with bendamustine followed by obinutuzumab maintenance • Lenalidomide with rituximab • Rituximab in combination with chemotherapy • BSC	• Lenalidomide with rituximab (R^2) • Rituximab in combination with chemotherapy, represented by rituximab + bendamustine (RB) • Obinutuzumab with bendamustine followed by obinutuzumab maintenance (OB)	BSC was not included, as it could be considered a palliative approach in those FL patients who require treatment. As the age-standardised 5-year net survival rate for FL in England is estimated at 83.0%[1], a palliative treatment approach (BSC) is therefore unlikely to be of major relevance. In the ITC (see Section B.2.9) and economic model (see Section B.3.2.5) rituximab in combination with chemotherapy was represented solely by the rituximab + bendamustine regimen. An ITC against R-CHOP was attempted but, despite availability of patient-level data from the EORTC 20981 trial[2,3], the analysis proved not to be methodologically feasible. OB is included for completeness but the company do not consider it to be a relevant comparator considering that clinical expert advice states that this regimen is not commonly used in the third-line setting and beyond, which is also corroborated by market share data which states that [redacted] of patients receive OB in the third-line setting.
Outcomes	• Overall survival	• Overall survival	No difference

© Roche (2022). All rights reserved

	• Progression free survival • Response rates • Adverse effects of treatment • Health-related Quality of Life	• Progression free survival • Response rates • Adverse effects of treatment • Health-related Quality of Life	
Subgroups to be considered	Not specified	Subgroups defined by: • Demographics (gender, age range categories, race/ethnicity, ECOG PS), • Presence of bulky disease • Number of prior systemic therapies and refractory status to those prior treatments • FLIPI prognostic score	Available data from key subgroups of clinical relevance is presented in the submission.
Abbreviations: BSC, best supportive care; ECOG PS, Eastern Cooperative Oncology Group Performance Score; FL, follicular lymphoma; FLIPI, Follicular Lymphoma International Prognostic Index; ITC, indirect treatment comparison; OB, obinutuzumab plus bendamustine; R^2, lenalidomide plus rituximab, RB, rituximab plus bendamustine; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone			

© Roche (2022). All rights reserved

B.1.2 Description of the technology being appraised

FL is a tumour arising from B-cells[4]. Late pre-B cells and mature B-cells (both healthy and malignant) express the CD20 antigen, which is a well-established therapeutic target in B-cell malignancies – both rituximab and obinutuzumab target the CD20 antigen[4].

Mosunetuzumab is a bispecific antibody that harnesses the patient's immune response by recruiting and activating cytotoxic T-cells to eliminate malignant B-cells. One fragment antigen-binding (Fab) region of mosunetuzumab targets the CD20 antigen expressed on B-cell surface, while the other targets CD3, a part of the T-cell receptor complex. Mosunetuzumab is a conditional agonist. When it is bound to both a B-cell and a cytotoxic T-cell, the T-cell is activated against the B-cell, resulting in B-cell death. This mechanism of action of mosunetuzumab is presented in Figure 1.

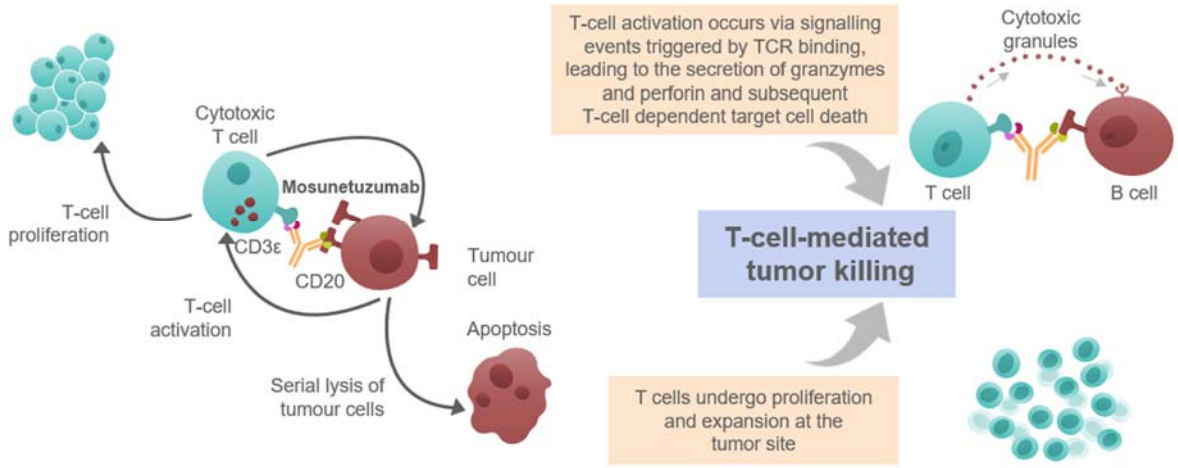


Figure 1. Mechanism of action of mosunetuzumab[5,6]

Abbreviations: TCR: T-cell receptor
Adapted from Aldoss et al. 2017[6]

At the time of this submission, marketing authorisation for mosunetuzumab in the UK has not yet been granted. In the European Union, the Committee for Medicinal Products for Human Use (CHMP) opinion was received on 21 April 2022. The expected EMA approval date is [redacted], with UK marketing authorisation (via the EU RELIANCE route) expected in [redacted]

© Roche (2022). All rights reserved

Table 2. Technology being appraised

UK approved name and brand name	Mosunetuzumab (Lunsumio®)
Mechanism of action	Mosunetuzumab is a T-cell–recruiting bispecific antibody targeting CD20-expressing B cells. One of its Fab regions is directed against the extracellular domain of the CD3ε subunit of the T-cell receptor complex, and the other Fab region is directed against the extracellular domain of CD20. Engagement of both arms of mosunetuzumab results in the formation of an immunologic synapse between a target B-cell and a cytotoxic T-cell, leading to T-cell activation and subsequent lysis of the target B-cell. Mosunetuzumab is a conditional agonist; targeted B-cell killing is observed only upon simultaneous binding to CD20 on B-cells and CD3 on T-cells.
Marketing authorisation/CE mark status	In the European Union, mosunetuzumab has held Orphan Drug Designation for the treatment of follicular lymphoma since 12 November 2021 (EMA/OD/0000058552). Mosunetuzumab does not currently have a marketing authorisation in the UK or elsewhere. On 10 September 2021, Roche submitted the initial marketing authorisation application to the EMA for mosunetuzumab as a single agent for the treatment of adult patients with relapsed or refractory FL who have previously received at least two prior systemic therapies. A positive CHMP opinion was issued on 21 April 2022 and the European Commission decision is anticipated in [redacted]. [redacted]
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication for mosunetuzumab is "[redacted]." The most recent draft of the SmPC is provided in Appendix C.
Method of administration and dosage	Mosunetuzumab should be administered for 8 cycles unless a patient experiences unacceptable toxicity or disease progression. For patients who achieve a complete response, no further treatment beyond 8 cycles is required. For patients who achieve a partial response or have stable disease in response to treatment with mosunetuzumab after 8 cycles, an additional 9 cycles of treatment (17 cycles total) should be administered, unless a patient experiences unacceptable toxicity or disease progression. The dosing of mosunetuzumab is as follows: • Cycle 1, Day 1: 1 mg • Cycle 1, Day 8: 2 mg • Cycle 1, Day 15: 60 mg

© Roche (2022). All rights reserved

	• Cycle 2, Day 1: 60 mg • Cycle 3 and subsequent cycles, Day 1: 30 mg Cycle duration is 3 weeks. See draft SmPC (Appendix C) for details of dose adjustments in case of toxicity. In the first cycle, mosunetuzumab should be administered over a minimum of 4 hours as intravenous infusion. If the infusions are well-tolerated in Cycle 1, the subsequent cycles may be administered over a 2-hour infusion.
Additional tests or investigations	Patients should receive premedication with IV corticosteroids at least 1 hour prior to mosunetuzumab infusion, as well as an antihistamine and an anti-pyretic at least 30 minutes prior to mosunetuzumab infusion. Premedication should be administered to all patients for mosunetuzumab Cycles 1 and 2. During subsequent cycles, premedication should be administered to patients who experienced any grade CRS with previous doses.
List price and average cost of a course of treatment	£220 per 1 mg vial (£6,600 per 30 mg vial)
Patient access scheme (if applicable)	[redacted] (simple discount)
Abbreviations: CHMP, Committee for Medicinal Products for Human Use; CRS, cytokine release syndrome; EMA, European Medicines Agency; Fab, fragment antigen-binding; Ig, immunoglobulin; IV, intravenously; MHRA, Medicines and Healthcare Products Regulatory Agency; SmPC, Summary of Product Characteristics	

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

B.1.3.1 FL overview and clinical course

FL is the most common type of low-grade non-Hodgkin lymphoma (NHL)[7]. As mentioned above, it is a B-cell lymphoma. FL is usually characterised by a chronic course with disease relapses[7]. Age-standardised net survival from FL in England is estimated to be 94.8% (95% CI: 94.3%, 95.3%) at 1 year from diagnosis and 83.0% (95% CI: 81.2%, 84.8%) at 5 years from diagnosis[1]. However, survival and duration of remission worsen significantly as patients progress through multiple lines of therapy[8,9].

The clinical course of the disease is heterogenous, and FL may transform into a high-grade diffuse large B-cell lymphoma (DLBCL)[10]. The risk of transformation is about 28% over 10 years, and the prognosis of affected patients is poor, with a median survival of 1.2 years from transformation[10].

© Roche (2022). All rights reserved

B.1.3.2 FL staging and grading

NHL, including FL, is currently often staged according to the Lugano classification, based on fluorodeoxyglucose-positron emission tomography (FDG-PET)[11]. The Lugano classification is a modification of the older Ann Arbor system and includes stages from I (localised disease) to IV (widespread disease outside of the lymphatic system)[12] (Table 3). Stage I–II FL is considered limited stage, while advanced stage disease comprises stages III-IV[12].

In addition to the staging described above, histological appearance of FL biopsy samples is used to grade the disease based on the number of centroblasts per high-powered field[13]. Grades 1 to 3A are considered indolent (low-grade) disease, while grade 3B should be treated as an aggressive disease[13] and is outside of the scope of the current appraisal.

Table 3. The Lugano staging system for lymphomas[12]

Stage	Involvement	Extranodal* status
Limited stage		
I	One node or a group of adjacent nodes	Single extranodal lesions without nodal involvement
II	Two or more nodal groups on the same side of the diaphragm	Stage I or II by nodal extent with limited contiguous extranodal involvement
II bulky	II as above with "bulky" disease	Not applicable
Advanced stage		
III	Nodes on both sides of the diaphragm; nodes above the diaphragm with spleen involvement	Not applicable
IV	Additional non-contiguous extralymphatic involvement	Not applicable
*Note: Tonsils, Waldeyer's ring, and spleen are considered nodal tissue.		

B.1.3.3 Symptom burden of FL

Symptoms of FL include swelling due to enlarged lymph nodes, commonly in the neck, armpit, or groin area[14]. Patients may also experience fatigue[15]. Some patients experience general symptoms, also known as B symptoms, that include night sweats, unexplained fevers, and unintended weight loss[14].

FL, particularly during disease relapses, has a substantial impact on health-related quality of life (HRQoL)[15,16]. A study of 181 FL patients from the Haematological Malignancy Research Network (HMRN) reported EQ-5D-5L crosswalk utilities ranging from 0.83 during disease remission to 0.74 in patients on treatment[16], compared with a UK general population reference utility of 0.857[17].

© Roche (2022). All rights reserved

B.1.3.4 Epidemiology of FL in England

2,476 people were newly diagnosed with FL in England in 2019, with men affected slightly more frequently: age-standardised incidence rates per 100,000 persons were 4.9 (95% CI: 4.6, 5.2) in men and 4.4 (95% CI: 4.2, 4.7) in women[18]. In terms of age, incidence was very low in people under 45 years of age and then rose gradually, with highest rates (up to 20 per 100,000 people) observed in those aged over 60 years[18].

With regards to prevalence, the HMRN estimated 3-year prevalence of FL in the UK at 6,700 people and 10-year prevalence at 16,220 people (latest year of data was 2016)[19].

Age-standardised mortality rates from FL per 100,000 population were similar between men (0.5, 95% CI: 0.4, 0.5) and women (0.4, 95% CI: 0.3, 0.5)[20].

B.1.3.4.1 Estimate of the target population size for mosunetuzumab

There were 2,470 reported cases of adults with follicular lymphoma in the 2019 edition of the NHS Digital Cancer Registration Statistics[18], resulting in an incidence of 0.0053% for FL when estimated as a proportion of the overall adult population of England in 2019 (45,470,282) as per ONS[21]. Applying this incidence to the overall adult population of England results in an estimated 2,517 cases of adults with FL in 2020. The adult population of England is assumed to be aged 15+ in line with the incidence banding reported in the NHS Digital Cancer Registration Statistics[18].

Based on the 2022 HMRN report, approximately 74.3% of patients (n=1,871 patients) per year can be expected to receive first-line chemotherapy or radiotherapy, including 1,461 patients who receive first-line therapy directly after diagnosis and a further 419 patients who go on to receive first-line therapy following a “watch and wait” period of no initial treatment[9] (see Section B.1.3.6 for an overview of current FL management). Of the 1,871 patients receiving first-line treatment, 27.3% (n=511 patients) are expected to relapse and receive second-line treatment and, of those, 33.7% (n=172 patients) are expected to relapse following second-line therapy and move onto third line[9]. The vast majority of these patients are likely to have grade 1 to 3A FL[9], so approximately 172 patients per year are expected to be eligible for treatment with mosunetuzumab.

B.1.3.5 Prognostic factors in FL

The Follicular Lymphoma-specific International Prognostic Index (FLIPI) and its revised version (FLIPI2) are frequently used to assess the risk of adverse outcomes among FL patients, stratifying patients into three groups (low, intermediate, and high risk)[11]. The prognostic factors contributing to the FLIPI include: age ≥60 years, haemoglobin concentration

© Roche (2022). All rights reserved

<120 g/l, elevated lactate dehydrogenase (LDH), stage III–IV disease, and ≥5 involved nodal areas[11]. FLIPI2 uses the same age and haemoglobin criteria as the FLIPI, but the remaining three risk factors are different from FLIPI and include elevated β2 microglobulin (as opposed to LDH), bone marrow involvement (as opposed to advanced stage disease), and longest diameter lymph node >6 cm (as opposed to the number of affected nodal areas)[11]. The latter risk factor in FLIPI2 pertains to bulky disease, i.e., the presence of a large, single nodal mass as opposed to multiple smaller lymph nodes[12].

Another important adverse prognostic factor in FL is progression of disease within 24 months from front-line therapy (POD24)[11,22]. A US-based study of 588 FL patients reported that early (within 24 months) disease progression after first-line therapy was associated with a 5-year overall survival (OS) of 50% compared to 90% in the group without relapse within 24 months[22].

While strong candidates for molecular prognostic factors in FL exist[23], these molecular factors do not currently seem to affect clinical decision making in routine practice and are therefore not described in detail in this submission.

B.1.3.6 Current management of FL in England

B.1.3.6.1 Key recent clinical guidelines in FL

The British Society for Haematology (BSH) Guideline on the investigation and management of FL[11] comprehensively described the recommended treatment of FL in the UK. The European Society for Medical Oncology (ESMO) practice guidelines on FL were also recently updated[13]. In addition, several technology appraisals (TAs) have been conducted by NICE in the area of R/R FL and are highly relevant to the current submission (TA627, TA629 and TA137[24-26], as well as TA604 for idelalisib which, however, was not recommended[27]).

B.1.3.6.2 First-line management of limited-stage disease

Approximately 10–15% of FL patients present initially with limited stage disease and can be treated with a curative intent[13]. Involved-site radiotherapy (ISRT) is the international standard and should be offered to patients with limited stage FL whose tumour can be encompassed within a radiotherapy field[11]. The recommended dose is 24 Gy in 12 daily fractions[11].

B.1.3.6.3 First-line management of advanced-stage disease

For the majority of FL patients who present with advanced stage disease, no curative therapy has been established[13]. A “watch and wait” approach involving observation with no therapy should be considered in patients with *asymptomatic*, advanced stage FL[11]. When determining the need for therapy, the UK guidelines refer to the Groupe d’Etude des Lymphomes

© Roche (2022). All rights reserved

Folliculaires (GELF) criteria[11,28], while the ESMO guidelines propose the use of tumour burden criteria adapted from the GELF and the British National Lymphoma Investigation (BNLI)[13]. The latter guidelines recommend that therapy should be initiated only upon the development of symptoms or complications of the disease, such as hematopoietic impairment, bulky disease, vital organ compression, ascites, pleural effusion or rapid lymphoma progression[13]. As an alternative to "watch and wait", rituximab monotherapy may also be considered[11,29], constituting a cost-effective treatment option[30]. However, rituximab does not currently have a UK marketing authorisation for this indication and is also not commissioned by NHS England in this setting[11].

Patients with advanced-stage, *symptomatic* disease should be treated with an anti-CD20 antibody combined with chemotherapy, the choice of which depends on patient characteristics and preferences of the treating clinician[11]. NICE TA234 recommends rituximab in combination with several chemotherapy regimens as an option for the treatment of symptomatic, advanced stage FL[31]. In patients with a FLIPI score ≥ 2 (corresponding to intermediate or high risk[32]), obinutuzumab in combination with chemotherapy is another first-line treatment option (recommended within TA513)[33].

Patients who respond to the initial induction treatment should receive maintenance therapy with rituximab or obinutuzumab (whichever antibody was used during induction) for up to 2 years[11,33,34].

B.1.3.6.4 Management of R/R FL

Therapy selection for relapsed disease depends on several factors:[11]

- The indication for therapy. There is no evidence that treatment can improve outcomes for patients with asymptomatic, relapsed FL. Observation alone should be considered for patients with asymptomatic, relapsed disease.
- Patient's fitness for therapy. Importantly, patients with relapsed FL who are fit enough should be considered for consolidation with high-dose therapy and autologous stem cell transplantation (ASCT) after achieving a second or subsequent remission.
- Previous treatments received and the duration of response to them (see below).
- Patient preference.

At present, two treatments are broadly recommended by NICE for the treatment of R/R FL and are considered relevant comparators for this submission:

- **Rituximab in combination with chemotherapy**, followed by maintenance treatment with rituximab (TA137[26]). BSH guidelines recommend that patients who had a relatively

© Roche (2022). All rights reserved

long duration of remission should be considered for repeated therapy with rituximab and the previously administered chemotherapy regimen. For those with a shorter remission duration, rituximab in combination with an alternative chemotherapy regimen from that administered previously should be considered[11].

- **Rituximab with lenalidomide** (R^2, TA627[24]), recommended as an option for previously treated follicular lymphoma (grade 1 to 3A) in adults[24].

Another comparator for this submission is **obinutuzumab in combination with bendamustine**, recommended for the treatment of patients who are refractory to rituximab, i.e., did not respond or progressed up to 6 months after treatment with rituximab or a rituximab-containing regimen (TA629)[25]. However, clinical experts consulted by the company noted that the OB is infrequently used in patients who had received 2 or more prior therapies, with some clinical experts only using this regimen in the second-line setting. This advice corroborates market share data obtained by the company which highlights that OB is used in just [redacted] of patients. As such, OB is included as a comparator in the current appraisal to align with the NICE scope but the company does not feel that it should be considered a relevant comparator.

Best supportive care (BSC), although specified in the NICE scope, was not included as a comparator in this submission. While interventions considered as BSC were not defined in the NICE scope for this appraisal, the BHS guidelines recommend low-dose radiotherapy for symptom control in patients who are not appropriate candidates for systemic therapy[11]. However, patients who are unsuitable for therapy with currently available systemic options are also unlikely to be candidates for treatment with mosunetuzumab, limiting the relevance of the BSC comparator. Clinical experts consulted by the company also noted that the number of patients receiving BSC only at third line would be very small.

B.1.3.6.5 Unmet medical need in R/R FL

Although initially an indolent and chemo-sensitive disease, the natural history of FL is relapsing-remitting. While the addition of rituximab to chemotherapy improved the outcomes of patients with FL[11], outcomes remain poorer in patients with POD24 compared with patients who do not progress within 24 months of first-line treatment[22].

A patient with FL usually receives several courses of treatment over several years, each associated with progressively shortening progression-free survival (PFS)[8,9,35]. In the UK, the treatment landscape at third and subsequent treatment lines is highly fragmented[9] with no established standard of care, and clinical advice to the company indicates that the therapeutic choices are heavily dependent on prior treatments received. As such, there is clearly a need for more effective, chemotherapy-free treatment options for patients with R/R FL who have

© Roche (2022). All rights reserved

received at least 2 prior therapies, particularly for patients whose disease is resistant to multiple agents or who do not tolerate chemotherapy, and who therefore have limited treatment options.

B.1.3.6.6 Proposed positioning of mosunetuzumab

Mosunetuzumab is proposed for use within NHS England as an alternative to any third- or later-line therapy option and irrespective of transplantation status (i.e., as a bridge to ASCT, in patients relapsing post-ASCT, and in those unsuitable for ASCT). Due to its unique mechanism of action, mosunetuzumab is also expected to provide a therapeutic option regardless of the patient's disease being refractory to various chemotherapy options, or rituximab. Therefore, within current NHS clinical practice, mosunetuzumab may be used instead of rituximab combined with various chemotherapy regimens, R^2, or obinutuzumab with bendamustine (Figure 2). Since clinical expert opinion indicates that the treatment pathway for FL is heterogenous across the UK, this versatility of mosunetuzumab is likely to be an important asset in real-world clinical practice.

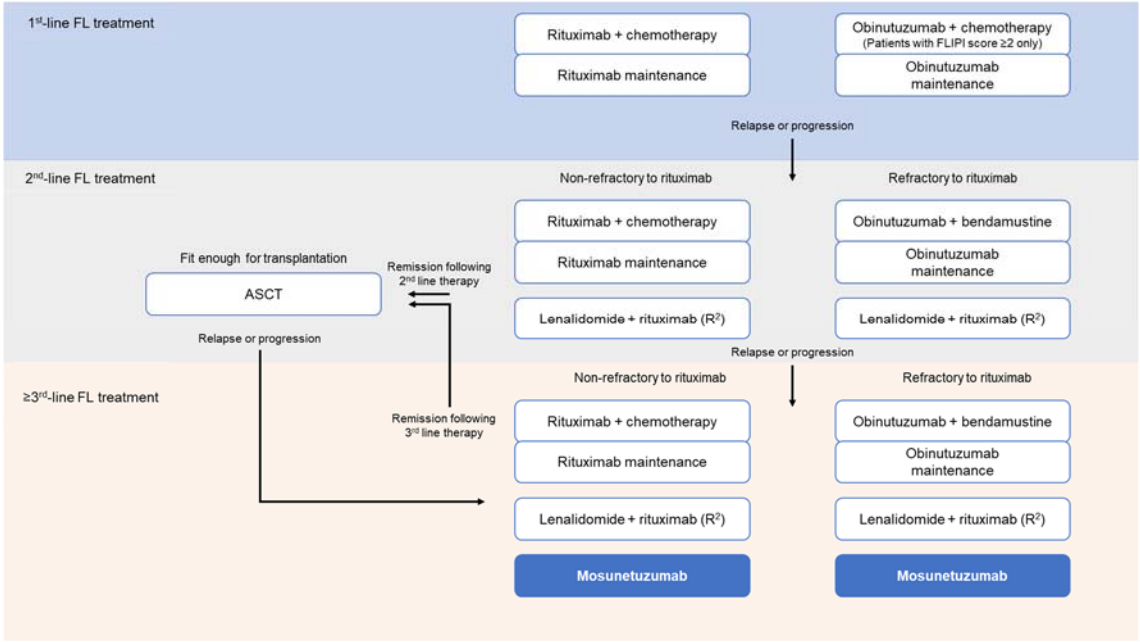


Figure 2. Proposed positioning of mosunetuzumab within the FL treatment pathway.

Abbreviations: ASCT, autologous stem cell transplantation; FL, follicular lymphoma; FLIPI, follicular lymphoma prognostic index.
Refractory to rituximab defined as disease progression on or within 6 months of rituximab treatment as per TA629[25]
Adapted from TA604[27], TA627[24], and the British Society for Haematology guidelines on investigation and management of FL[11]

B.1.4 Equality considerations

Treatment with mosunetuzumab within the NHS is not expected to raise any equality issues.

© Roche (2022). All rights reserved

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

A systematic literature review (SLR) was performed to identify studies of treatments for patients with R/R FL, including clinical evidence from randomised controlled trials (RCTs), non-randomised clinical studies, and real-world observational studies. The results of the clinical SLR were used for a feasibility assessment, exploring the feasibility of conducting matching-adjusted indirect comparisons (MAIC) between mosunetuzumab and comparator treatments in patients with R/RFL who have received ≥2 prior therapy lines (see Section B.2.9 for details of the MAICs and other indirect treatment comparisons [ITCs] performed).

Searches were performed in Medline®, Embase, and EBM Reviews on the 17th of April 2021 initially and were subsequently updated on the 24th of December 2021. The database searches were supplemented with searching conference proceedings from the past 3 years, reference lists of included publications, websites of HTA bodies, clinical trial registries, and European public assessment report (EPAR) documents.

The following Population, Intervention, Comparator, Outcomes, and Study Design (PICOS) were applied:

- The eligible patient population was adult patients with R/R FL, i.e., patients receiving second or later line of therapy. This broad population was selected to ensure that all relevant studies were identified, including those with mixed line populations which could include a high proportion of patients treated at third and subsequent line, or which reported subgroup results for this population.
- There were no restrictions on treatments (intervention or comparator) or outcomes.
- All prospective and retrospective study designs were eligible but economic evaluations, case report, pharmacokinetic and animal or *in vitro* studies were excluded.

Data extraction was performed by a single reviewer and quality checked for all data elements by a second reviewer, with disputes referred to a third-party advisor.

A total of 172 studies reported in 214 publications were identified for inclusion into the SLR. There were 76 studies conducted exclusively in patients with FL across various lines of therapy, including 18 studies focusing exclusively on patients receiving third or subsequent therapy line and another 5 studies reporting subgroup results for this population.

© Roche (2022). All rights reserved

Forty-two studies evaluated treatments of interest and were potentially relevant for the MAIC feasibility assessment. Upon further review of these studies, the following criteria were applied which resulted in 19 studies being considered for the feasibility assessment.:

- 3L+ FL study (n=6)
- Mixed line FL study if median ≥ 2 prior lines (≥ 50% 3L) or with results for 3L+ (n=5)
- 3L+ mixed lymphoma study if > 80% FL (or if ≤ 80% with results for FL patients) (n=2)
- Mixed line and mixed lymphoma study if > 80% FL (or if ≤ 80% with results for FL patients) AND median ≥ 2 prior lines (≥ 50% 3L) (n=6)

Two further studies, AUGMENT and EORTC 20981, were added to this list as they assessed treatments of high relevance to this appraisal (R^2 and R-CHOP, respectively), thus expanding the list of studies considered in the feasibility assessment to a total of 21. Of these, 5 were sponsored by Roche and individual patient data (IPD) was available; therefore, propensity score analyses applying the methods recommended by NICE DSU TSD 17[36] were performed. The feasibility of conducting a MAIC was assessed for the other 16 studies. It should, however, be noted that the list of comparators considered in the feasibility assessment was broader than the range of comparators considered in this appraisal. For details of ITC performed against comparators specified in the scope of this appraisal, please see Section B.2.9. The report from the SLR and feasibility assessment is provided as Appendix D.

B.2.2 List of relevant clinical effectiveness evidence

Mosunetuzumab was studied in the Phase I/II, multicentre, open-label, dose-escalation and dose-expansion GO29781 study (Table 4). This study enrolled patients with R/R haematologic malignancies expected to express CD20 (including FL) and evaluated the use of mosunetuzumab alone and in combination with atezolizumab.

The focus of this submission is the pivotal cohort of FL patients who had relapsed after or failed to respond to at least two prior lines of systemic therapy AND received single-agent mosunetuzumab at the recommended phase II dose (which is also the planned label dose) during the dose expansion phase of the trial.

Table 4. Clinical effectiveness evidence

Study	**GO29781 pivotal cohort**
Study design	Phase I/II, multicentre, open-label, dose-expansion study.
Population	90 patients with FL, who had relapsed after or failed to respond to at least two prior lines of systemic therapy and received single-agent mosunetuzumab at the recommended phase II dose.
Intervention(s)	Mosunetuzumab IV at a dose of 1/2/60/30 mg
Comparator(s)	None

© Roche (2022). All rights reserved

Study	GO29781 pivotal cohort				
Indicate if trial supports application for marketing authorisation	Yes	✓	Indicate if trial used in the economic model	Yes	✓
	No			No	
Rationale for use/non-use in the model	This pivotal study provided key clinical efficacy and safety data supporting the modelling.				
Reported outcomes specified in the decision problem	• OS • PFS • Response rates • AEs • HRQoL				
All other reported outcomes	• DOR • Pharmacokinetics • Pharmacodynamics				
Abbreviations: AE, adverse event; DOR, duration of response; FL, follicular lymphoma; HRQoL, health-related quality of life; IV, intravenously; OS, overall survival; PFS, progression-free survival; Q3W, every 3 weeks; R/R, relapsed or refractory					

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

Methodological aspects of the GO29781 study design that are relevant to the pivotal cohort, and therefore this submission, are summarised in Table 5.

Please note, that baseline characteristics of study participants are provided in Section B.2.6.

Table 5. Summary of GO29781 pivotal cohort design

Trial design	Phase I/II, international, multicentre, open-label, dose-expansion study
Key eligibility criteria for participants	• Adult patients with grade 1–3a FL that had relapsed after or failed to respond to at least two prior lines of systemic therapy • ECOG PS ≤1 • Prior treatment with an anti–CD20-directed therapy and an alkylating agent • Patients who had received alloSCT or had CNS lymphoma were excluded
Settings and locations	Multiple countries including the UK: United States (n=40), Australia (n=17), Canada (n=13), Spain (n=8), Germany (n=6), South Korea (n=4), United Kingdom (n=2)
Summary of trial drugs	• Mosunetuzumab monotherapy administered by IV infusion at a step-up dosing regimen of 1/2/60/30 mg (1.0 mg on Day 1 of Cycle 1, 2.0 mg on Day 8 of Cycle 1, 60.0 mg on Day 15 of Cycle 1 and Day 1 of Cycle 2, and 30 mg on Day 1 of subsequent cycles, with a 3-week cycle duration) • Mosunetuzumab was administered for 8 cycles unless the patient experienced unacceptable toxicity or disease progression. For patients achieving CR, no additional treatment beyond the 8 cycles was required. Patients achieving PR or SD received up to 9 additional cycles (17 cycles in total), unless the patient experienced unacceptable toxicity or disease progression

© Roche (2022). All rights reserved

	• No comparator therapy was included. Concomitant use of any treatment for lymphoma or leukaemia resulted in the patient being taken off study
Primary outcome	Proportion of patients whose best overall response was a CR based on IRF assessment. Response was assessed by CT and PET-CT using the International Working Group evaluation criteria described by Cheson et al. 2007[37].
Other outcomes used in the economic model/specified in the scope	*Efficacy:* • CR rate per investigator assessment • ORR per IRF and investigator assessment • PFS per IRF and investigator assessment, from the time of first mosunetuzumab treatment • OS, from the time of first mosunetuzumab treatment Response assessment was based on International Working Group criteria[37]. *HRQoL:* • EORTC QLQ-C30 • FACT-Lym subscale • EQ-5D-5L *Safety:* • Treatment-emergent AEs, including SAEs
Pre-planned subgroups	Comprehensive analyses including stratifications by demographics, baseline disease characteristics, prior treatment history (number of prior systemic therapies and refractory status to those prior treatments), and prognostic factors. Please see Section B.2.7 for a complete list.
Abbreviations: AE, adverse event; alloSCT, allogeneic stem cell transplant; CNS, central nervous system; CR, complete response; CT, computed tomography; ECOG PS, Eastern Cooperative Oncology Group Performance Score; EORTC QLQ-C30: European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; FL, follicular lymphoma; HRQoL, health-related quality of life; IRF, independent review facility; IV, intravenously; ORR, objective response rate; OS, overall survival; PET-CT, positron emission tomography - computed tomography; PFS, progression-free survival; PR, partial response; R/R, relapsed or refractory; SAE, serious adverse event; SD, stable disease	

B.2.3.1 Overview of overall GO29781 study design (including the pivotal and other cohorts)

GO29781 was a phase I/II, international, multicentre, open-label, dose-escalation and dose-expansion study of mosunetuzumab administered as a single agent and in combination with atezolizumab in patients with R/R haematologic malignancies expected to express CD20 (including FL).

The ***dose-escalation stage*** of the study was designed to assess the safety, tolerability, and pharmacokinetics of mosunetuzumab administered by IV infusion or subcutaneous (SC) injection. Up to five dose escalation groups could be enrolled (group C was removed from the protocol before any patients were enrolled).

© Roche (2022). All rights reserved

- Group A: Single dose Cycle 1 non-fractionated single-agent mosunetuzumab escalation, IV infusion. Enrolment into dose-escalation Group A was stopped to prioritize assessment of other dosing schedules and route of mosunetuzumab administration
- Group B: Cycle 1 step-up single-agent mosunetuzumab escalation, IV infusion
- Group D: Cycle 1 non-fractionated single-agent mosunetuzumab escalation, SC injection
- Group E: Cycle 1 step-up single-agent mosunetuzumab escalation with concurrent administration of atezolizumab starting in Cycle 2, IV infusion
- Group F: Cycle 1 step-up single-agent mosunetuzumab escalation, SC injection.

The ***dose-expansion stage*** provided additional safety, tolerability, pharmacokinetics, and clinical activity data with mosunetuzumab doses up to the maximum tolerated dose / maximum assessed dose. Patients with R/R FL, R/R diffuse large B-cell lymphoma (DLBCL) and transformed FL (tFL), R/R mantle cell lymphoma (MCL), and R/R Richter's transformation were enrolled into separate expansion cohorts specific to each NHL type. Key study design aspects relevant to the pivotal cohort (described as the R/R FL cohort B11 in the clinical study report), which is of key interest to this appraisal, are presented in Figure 3. Please note, that only the methodology aspects relevant to this cohort are described in the remainder of Section B.2.3.

© Roche (2022). All rights reserved

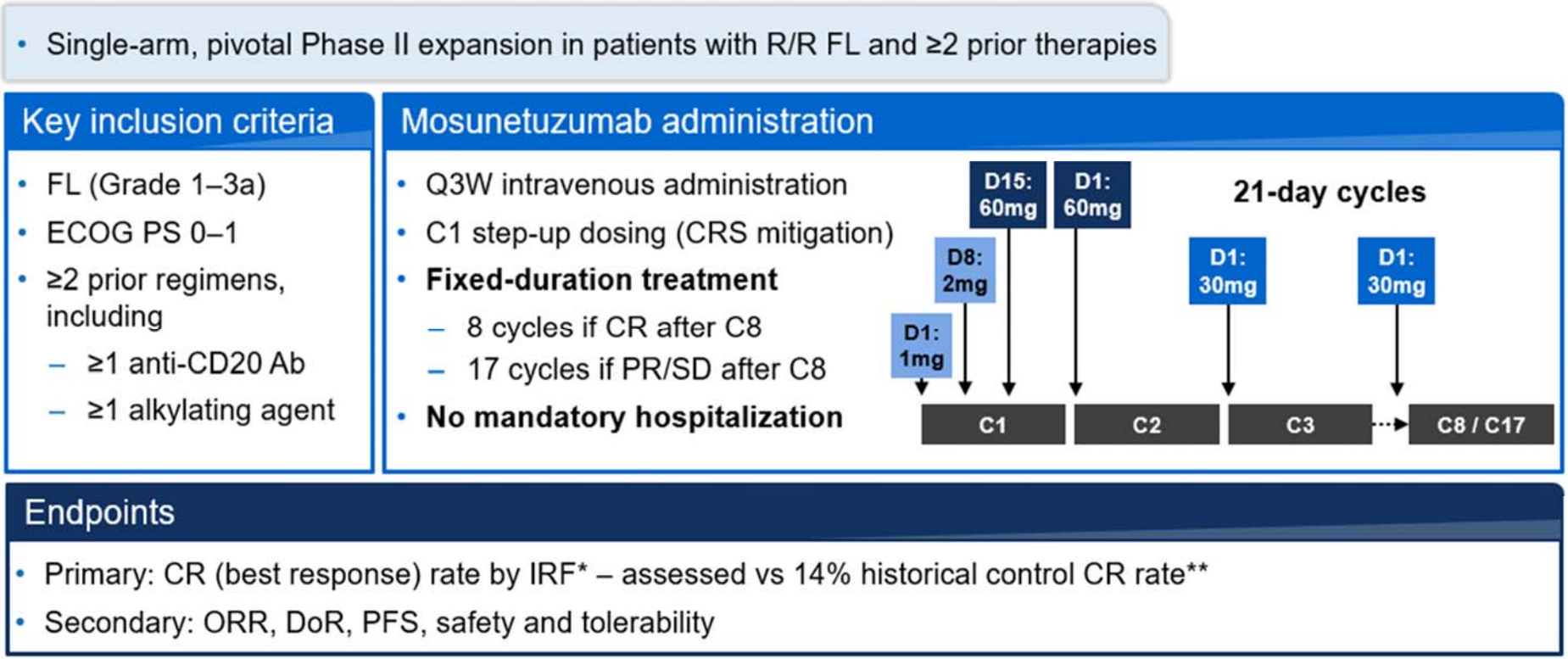


Figure 3. Overview of GO29781 pivotal cohort study design.

*Assessed by CT and PET-CT using Cheson 2007 criteria[37]

**Historical control CR rate was derived from Dreyling et al. 2017[38]

Abbreviations: Ab, antibody; CR, complete response; CT, computed tomography; D, Day; DOR, duration of response; IRF, independent review facility; ORR, objective response rate; PET, positron emission tomography; PFS, progression-free survival; PR, partial response; Q3W, once every 3 weeks; SD, stable disease.

© Roche (2022). All rights reserved

B.2.3.2 Eligibility criteria

Key inclusion and exclusion criteria for the pivotal GO29781 cohort are listed below. Please see the trial protocol for a complete list of eligibility criteria.

Key inclusion criteria:

- Provided signed informed consent
- Age ≥18 years
- Eastern Cooperative Oncology Group Performance Score (ECOG PS) ≤1
- Diagnosis of grade 1–3a FL that had relapsed after or failed to respond to at least two prior lines of systemic therapy
- Prior treatment with an anti–CD20-directed therapy and an alkylating agent
- At least one bi-dimensionally measurable lesion (>1.5 cm in its largest dimension for nodal lesions, or >1.0 cm in its largest dimension for extranodal lesions by computerized tomography [CT] scan or magnetic resonance imaging [MRI])

Key exclusion criteria:

Patients who met any of the following criteria were to be excluded from study entry:

- Prior use of any monoclonal antibody, radioimmunoconjugate, or antibody-drug conjugate within 4 weeks before first mosunetuzumab administration
- Prior treatment with systemic immunotherapeutic agents for which the mechanism of action involves T cells (including but not limited to cytokine therapy and anti CTLA-4, anti–PD-1 and anti–PD-L1 therapeutic antibodies) within 12 weeks or five half-lives of the drug, whichever was shorter, before the first mosunetuzumab administration
- Prior treatment with CAR-T therapy within 30 days before first mosunetuzumab administration
- Treatment with any chemotherapeutic agent, or treatment with any other anti-cancer agent (including investigational agents) within 4 weeks or five half-lives of the drug, whichever was shorter, prior to the first mosunetuzumab administration
- Treatment with radiotherapy within 2 weeks prior to the first mosunetuzumab administration. If patients had received radiotherapy within 4 weeks prior to the first mosunetuzumab administration, patients must have had at least one measurable lesion outside of the radiation field. Patients who had only one measurable lesion that was previously irradiated but subsequently progressed were eligible
- Autologous stem cell transplant (ACST) within 100 days prior to first mosunetuzumab administration

© Roche (2022). All rights reserved

- Prior allogeneic stem cell transplant
- Current or past CNS lymphoma

B.2.3.3 Study procedures and assessments

Written informed consent was obtained prior to performing any study-specific screening tests or evaluations. At the screening visit, eligibility for participation was reviewed and medical history and demographic characteristics were collected. Tumour response was assessed using CT or PET-CT performed at screening (baseline assessment) and subsequently at 3 months (± 1 week) and at 6 months (± 2 weeks) after first mosunetuzumab infusion, followed by every 3 months (± 2 weeks) thereafter. An additional early assessment at 6 weeks (± 1 week) was optional at the investigator's discretion. The tumour assessment at 6 months (± 2 weeks) was used to determine the duration of mosunetuzumab treatment (see Section B.2.3.4.1.2). In addition to radiological assessment, a bone marrow biopsy was performed at screening. In patients whose screening biopsy was positive for tumour, a subsequent biopsy was required to confirm CR. Patient-reported outcomes (PROs, see Section B.2.3.5.1) were collected prior to mosunetuzumab infusion in cycle 1 (baseline assessment), cycle 2, and every other cycle thereafter (i.e., at even-numbered cycles). Collection of adverse events (AEs) included all AEs and serious adverse events (SAEs) occurring after the initiation of treatment with mosunetuzumab and until 90 days following the last treatment administration or until study completion or patient discontinuation, whichever was later. After this period, AE reporting included only SAEs that the investigators believed to be related to prior treatment with mosunetuzumab.

Please see the study protocol for a complete schedule of assessments.

B.2.3.4 Study treatments and concomitant medications

B.2.3.4.1 Study treatment

B.2.3.4.1.1 Dose

Mosunetuzumab monotherapy was administered by IV infusion in a flat (weight-independent), step up dosing regimen of 1.0 mg on Day 1 of Cycle 1, 2.0 mg on Day 8 of Cycle 1, 60.0 mg on Day 15 of Cycle 1 and Day 1 of Cycle 2, and 30 mg on Day 1 of subsequent cycles, with a 3-week cycle duration. This dosing is hereafter referred to as 1/2/60/30 mg.

B.2.3.4.1.2 Number of treatment cycles

Mosunetuzumab was administered for 8 cycles unless progressive disease (PD) (see Section B.2.3.4.1.3) or unacceptable toxicity was observed prior to completion of the 8 cycles. Tumour

© Roche (2022). All rights reserved

assessment at 6 months (± 2 weeks) was scheduled during Cycle 8 and was used to determine the duration of mosunetuzumab treatment, as follows:

- Patients who achieved a complete response (CR) after receiving 8 cycles of treatment did not receive any additional cycles of mosunetuzumab but continued to be monitored. Patients who subsequently developed PD were eligible for retreatment with single-agent mosunetuzumab for at least 8 additional cycles. Data on retreatment will not be presented in this submission.
- Patients who achieved a partial response (PR) or maintained stable disease (SD) after receiving 8 cycles of mosunetuzumab were to continue treatment for up to a total of 17 cycles unless PD or unacceptable toxicity was observed. Patients who achieved CR, PR, or SD after 17 cycles of treatment were monitored. As for patients achieving CR after 8 cycles, patients achieving CR after 17 cycles of mosunetuzumab were also eligible for retreatment if they subsequently experienced PD. Data on retreatment will not be presented in this submission.

Multiple rounds of re-treatment were allowed per the protocol.

B.2.3.4.1.3 Treatment after radiographic progression

Treatment beyond radiographic progression was permitted only if pseudoprogression was suspected. Patients continuing mosunetuzumab treatment despite apparent radiographic progression were strongly encouraged to undergo a biopsy to assess whether increases in tumour volume were due to immune cell infiltration or neoplastic proliferation, provided that such a biopsy could be performed safely on a non-target lesion. Patients also had to provide written consent to acknowledge discussion with the treating investigator about the risks and benefits of continuing mosunetuzumab treatment beyond radiographic progression.

If true progression was suspected based on the investigator's judgment, clinical factors, or biopsy findings that were consistent with neoplastic proliferation, or if radiographic disease progression was confirmed at a subsequent tumour assessment, the patient was ineligible to receive further treatment with mosunetuzumab.

B.2.3.4.2 Concomitant medications

Concomitant therapy was defined as any medication (e.g., prescription or over-the-counter drugs, supplements, herbal remedies, etc.) used by a patient from 7 days prior to screening to the study completion/discontinuation visit. Key permitted and prohibited medications are listed below. Please see trial protocol for a complete list.

Permitted medications:

© Roche (2022). All rights reserved

- Oral contraceptives, hormone-replacement therapy, or other maintenance therapy
- Granulocyte colony stimulating factor (G-CSF) for both prophylactic and therapeutic purposes
- Anti-infective prophylaxis for viral, fungal, bacterial, or pneumocystis infections according to institutional practice
- Symptomatic treatment of mosunetuzumab infusion-related symptoms, including analgesics, anti-pyretics, and antihistamines as indicated.
- Treatment of severe cytokine release syndrome (CRS) according to published recommendations and/or institutional practice, including supportive care, tocilizumab, and corticosteroids. For patients refractory to tocilizumab, siltuximab, anakinra, dasatinib and emapalumab could be administered at the discretion of the investigator.

Prohibited medications:

- Cytotoxic chemotherapy intended for treatment of lymphoma or leukaemia
- Radiotherapy
- Immunotherapy
- Hormone therapy for the treatment of cancer, whether approved by local regulatory authorities or investigational. However, adjuvant endocrine therapy for non-metastatic, hormone receptor positive breast cancer was permitted.
- Biologic agents other than haematopoietic growth factors.
- Any therapies intended for the treatment of lymphoma or leukaemia, whether approved by local regulatory authorities or investigational

Patients who required the use of any aforementioned prohibited therapy were to be discontinued from treatment with mosunetuzumab; however, they continued to be followed for safety outcomes for 90 days after their last mosunetuzumab dose or start of another anti-cancer therapy, whichever occurred first.

- Vaccination with live vaccines was not permitted for at least 4 weeks before initiation of treatment with mosunetuzumab, at any time during treatment, and until B-cell levels recovered to normal range after the last dose of mosunetuzumab. Patients who required vaccination with a live virus vaccine were to be discontinued from treatment with mosunetuzumab.

© Roche (2022). All rights reserved

B.2.3.5 Endpoints

Primary and secondary efficacy and patient-reported outcomes (PRO) endpoints are summarised in Table 6. Tumour response was assessed by CT and PET-CT using the International Working Group response evaluation criteria described by Cheson et al. 2007[37].

Safety assessments included AEs, laboratory assessments (haematology and chemistry), electrocardiograms (ECGs) and vital signs, and anti-drug antibodies (ADA). Only treatment-emergent AEs (defined as AEs with onset on the day of or after first administration of mosunetuzumab) are presented in this submission. The causal relationship of AEs to treatment with mosunetuzumab was assessed by the investigators.

Table 6. Endpoints and their definitions

Endpoint name	Definition
Primary efficacy endpoint	
CR rate per IRF assessment	Proportion of patients whose best overall response was a CR based on IRF assessment.
Secondary efficacy endpoints	
CR rate per investigator assessment	Proportion of patients whose best overall response was a CR based on investigator assessment.
ORR, per IRF and investigator assessment	Proportion of patients whose best overall response was a PR or CR.
Duration of CR, per IRF and investigator assessment	Time from the initial occurrence of a documented CR until documented disease progression or death due to any cause, whichever occurred first.
DOR, per IRF and investigator assessment	Time from the initial occurrence of a documented PR or CR until documented disease progression or death due to any cause, whichever occurred first.
PFS, per IRF and investigator assessment	Time from the first mosunetuzumab dose to the first occurrence of disease progression or death from any cause, whichever occurred first.
OS	Time from the first mosunetuzumab dose to the date of death from any cause.
Secondary patient-reported outcome measures	
EORTC QLQ-C30	Summary statistics and change from baseline in HRQoL based on EORTC QLQ C30
FACT-Lym subscale	Summary statistics and change from baseline in disease-related symptoms based on the FACT-Lym subscale
EQ-5D-5L	Descriptive results of the EQ-5D-5L during patients' participation in the study
Abbreviations: CR, complete response; DOR, duration of response; EORTC QLQ-C30: European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire; FACT-Lym, Functional Assessment of Cancer Therapy – Lymphoma; IRF, independent review facility; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PR, partial response	

B.2.3.5.1 PRO measures

Three PRO scales were used in the GO29781 study: the EORTC-QLQ-C30, the FACT-Lym subscale, and the EQ-5D-5L.

© Roche (2022). All rights reserved

- EORTC QLQ-C30 is a widely used questionnaire that assesses HRQoL of cancer patients[39,40]. The questionnaire includes 30 items measuring five functioning scales (physical, social, role, cognitive, and emotional functioning), eight symptom scales (fatigue, nausea/vomiting, pain, dyspnoea, sleep disturbances, appetite loss, constipation, and diarrhoea), financial impact, and overall HRQoL[41].
- FACT-Lym subscale (version 4) enables assessment of the changes from baseline with respect to B-symptoms and impact on HRQoL brought about by symptom worsening or alleviation and treatment toxicity. The subscale range is 0–60, with a higher score reflecting better HRQoL, and a recall period of the previous week. The validity and reliability of the FACT-Lym subscale for NHL patients has been established[42].
- EQ-5D-5L evaluates health status. It was introduced in 2009 to improve sensitivity and reduce ceiling effect compared with EQ-5D-3L[43]. The questionnaire comprises a descriptive system including 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) with five levels each (no problems, slight problems, moderate problems, severe problems and extreme problems), and a visual analogue scale (VAS) indicating self-rated health[43].

B.2.3.6 Planned subgroup analyses

Subgroup analyses were conducted on the primary efficacy endpoint of CR rate, with results presented as forest plots displaying ORR and CR rates (per independent review facility [IRF] assessment) with 95% CIs for each subgroup. Subgroup analyses were not adjusted for multiplicity and all analyses should be considered exploratory. Pre-planned subgroups analysed included:

- Age (<65 vs ≥65 years)
- Sex
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not stated or Unknown)
- Race (White, Black/ African American, Asian, American Indian or Alaska Native, Multiple, Unknown)
- BMI (<median vs ≥median)
- ECOG score (0, ≥1)
- CD20 status (positive, negative)
- Prior lines of therapy (2, ≥3)
- R/R to last line of therapy (refractory, non-refractory)
- Received prior CAR-T therapy (yes, no)

© Roche (2022). All rights reserved

- R/R to prior anti-CD20 therapy (refractory, non-refractory)
- Time since last anti-CD20 therapy (≤3 months >3 months)
- Double Refractory to Prior Anti-CD20 Therapy and Prior Alkylator Therapy (Yes, No)
- R/R to prior alkylator therapy (refractory, non-refractory)
- R/R to prior PI3K inhibitor (refractory, non-refractory)
- PD within 24 months of start of 1L therapy (yes, no)
- FLIPI Score (low [0-1], intermediate [2], high [3-5])
- Bulky disease, i.e., >6cm (yes, no)
- EZH2 mutation (mutant, wild type)
- Received prior rituximab and lenalidomide (yes, no)

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

Statistical analyses in the pivotal cohort of the GO29781 study are summarised in Table 7 and described in detail below.

© Roche (2022). All rights reserved

Table 7. Summary of statistical analyses in the GO29781 study

Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
A comparison of CR rate between GO29781 pivotal cohort patients and historical controls was conducted using an exact binomial test with two-sided α level of 5% based on the following statistical hypothesis: H_0: CR rate=14% vs H_a: CR rate≠14%.	The primary endpoint, IRF-assessed CR rate, was estimated along with the Clopper Pearson exact 95% CI and compared against historical controls. For secondary tumour response endpoints, the exact 95% CIs using the Clopper Pearson method were provided. Kaplan-Meier analysis was performed for time-to-event secondary endpoints. The Brookmeyer-Crowley method was used to construct the 95% CIs around median time to event estimates and the Greenwood's formula was used to estimate 95% CIs around timepoint estimates for each endpoint (e.g., 12-month PFS).	With observed CR rates of 24% and 28%, a sample size of 80 patients would result in 95% CIs of (15%, 35%) and (18%, 39%), respectively, i.e., a true CR rate below 14% could be ruled out. Additionally, 80 patients would provide an 83% power to detect a 14% increase in CR rate from 14% to 28%, at the 5% two-sided significance level.	• Patients who discontinued from the study for any reason prior to achieving a CR were classified as non–complete responders. • Patients who started a new treatment for lymphoma prior to achieving CR were classified as non–complete respondents.
Abbreviations: CI, confidence interval; CR, complete response; FL, follicular lymphoma; IRF, independent review facility; R/R relapsed or refractory			

© Roche (2022). All rights reserved

B.2.4.1 Statistical hypothesis

The primary endpoint, IRF-assessed CR rate, was estimated along with the Clopper Pearson exact 95% confidence interval (CI). Comparisons of CR rate between the GO28781 pivotal cohort population and historical controls was tested. The historical control rate of 14% was based on the single-arm, multicentre study (CHRONOS-1) in 104 patients with R/R FL who had received ≥2 prior treatment lines[44].

Hypothesis testing on the primary endpoint of IRF-assessed CR rate was conducted at the time of the primary efficacy analysis based on the data cut of 15 March 2021 and occurred when the following conditions were met:

- At least approximately 6 months after the last patient was enrolled in the pivotal cohort or the Group B R/R DLBCL/tFL expansion cohort (whichever was later)
- Efficacy-evaluable population included approximately 80 patients for each histology

The following hypotheses were tested at the time of the primary analysis, using an exact binomial test with a two-sided significance level of 0.05: H_o: CR rate=14% vs H_a: CR rate≠14%.

Please note that the results presented in this submission are based on the more recent 27 August 2021 data cut-off, rather than the primary analysis cut-off of 15 March 2021.

B.2.4.2 Analysis framework for the primary endpoint

The data for the primary endpoint were presented as the proportion of patients whose best overall response was a CR based on IRF assessment out of the number of patients in the pivotal cohort. As described above, comparisons of CR rate between the pivotal cohort population and historical controls were conducted using an exact binomial test with two-sided alpha level of 5%. Primary analysis rules concerning intercurrent events are presented in Table 8.

© Roche (2022). All rights reserved

Table 8. Primary analysis rules

Intercurrent event	Analysis rule
Discontinuation from study (including for COVID-related reasons) prior to achieving CR	Patient included as a non–complete responder
Start of new anti-lymphoma therapy prior to achieving CR	Patient included as a non–complete responder
Discontinuation of study drug (including for COVID-related reasons) prior to achieving CR	Available assessments after the discontinuation of study drug were used to determine CR status
Missed any scheduled response assessments (including for COVID-related reasons) prior to achieving CR	Available assessments were used to determine CR status
Abbreviations: COVID, coronavirus disease; CR, complete response	

B.2.4.3 Analysis of secondary endpoints

- For tumour response endpoints, the exact 95% CIs using the Clopper Pearson method were provided.
- For time to event endpoints, the Kaplan-Meier estimate was provided, and the Brookmeyer-Crowley method was used to construct the 95% CIs for the median duration of CR, median DOR, median PFS, and median OS. Kaplan-Meier method was used to estimate 6-month and 12-month duration of CR and DOR, as well as 6 month and 12-month PFS and OS, along with the corresponding 95% CIs using Greenwood's formula. The following censoring rules applied for time to event endpoints of PFS, duration of CR, and DOR:
    - Patients starting new anti-lymphoma therapy were censored at last assessment prior to the start of new therapy.
    - Patients who neither died nor experienced disease progression prior to the data cut-off were censored at last assessment.
    - Patients who discontinued from the study and subsequently experienced disease progression or died were censored at last assessment.
- A descriptive comparison was also performed between the IRF and investigator assessments of CR rate, ORR, and DOR, to evaluate the concordance between the two.

B.2.4.4 PRO analysis

Visit summary and change from baseline analyses were performed for all PRO measures. The number and proportion of patients with a clinically meaningful improvement were also summarised for the EORTC QLQ-C30 and the FACT-Lym.

© Roche (2022). All rights reserved

PRO data was analysed up to Cycle 8 of treatment, due to the low number of evaluable patients at cycles beyond Cycle 8 (typically <25% of the patient population available at baseline).

For each questionnaire, a patient was considered as compliant if they completed at least one question.

B.2.4.4.1 EORTC QLQ-C30

The PRO analyses for the EORTC QLQ-C30 questionnaire were performed on the physical functioning domain (comprised of questions 1–5) and the fatigue symptom domain (comprising questions 10, 12, and 18). After linear transformation, scores for each domain ranged from 0 to 100. For the physical functioning domain, higher scores represented better physical functioning. However, for the fatigue symptom domain, higher scores represented worsening symptoms.

Clinically meaningful improvement at any time was defined as ≥10-point increase (for physical functioning) and decrease (for fatigue) compared to baseline[45].

B.2.4.4.2 FACT-Lym

For the FACT-Lym subscale, after linear transformation, scores ranged from 0 to 60, with a higher score representing better HRQoL. Clinically meaningful improvement was defined as a ≥3-point increase compared to baseline[46,47].

B.2.4.4.3 EQ-5D

Summary statistics for the health status according to the VAS and changes in the index utility score from baseline will be calculated for EQ-5D-5L.

B.2.4.5 Sample size and power calculations

The expansion pivotal FL cohort was designed to rule out a 14% CR rate[44] (see Section B.2.4.1) and was powered to detect a 14% increase in CR rate from 14% to 28%. With observed CR rates of 24% and 28%, a sample size of 80 patients would result in 95% CIs of (15%, 35%) and (18%, 39%), respectively, i.e., a true CR rate below 14% could be ruled out. Additionally, 80 patients would provide an 83% power to detect a 14% increase in CR rate from 14% to 28%, at the 5% two-sided significance level.

© Roche (2022). All rights reserved

B.2.4.6 Interim analyses

Continuous safety monitoring and interim analyses were performed to guide potential early stopping of enrolment in case of unacceptable toxicity or lower than expected response rate. Please see the trial protocol for additional details.

B.2.4.7 Analysis populations

Efficacy was assessed in all patients enrolled in the pivotal cohort. Therefore, the efficacy analysis followed the intention-to-treat principle.

Safety was assessed in patients who received at least one dose of mosunetuzumab. The safety data presented in this submission refers only to the pivotal cohort, unless explicitly stated otherwise.

The PRO evaluable population included all enrolled patients in the pivotal cohort who had a baseline and at least one post-baseline assessment of PRO scales. The PRO-evaluable population was used for descriptive analyses of visit summary and change from baseline.

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

Risk of bias was assessed by two reviewers independently, with disagreements resolved by discussion and/or involvement of additional referees. Quality (risk of bias) assessment of RCTs was conducted using the seven-criteria checklist provided in section 2.5 of the NICE single technology appraisal user guide, which evaluates selection, performance, attrition, and detection bias. The quality assessment of observational studies was conducted using the Downs and Black checklist for non-randomised studies[48].

Since the GO29781 pivotal cohort provided non-randomized evidence, its quality assessment using the Downs and Black checklist[48] is shown in Table 9. The quality assessment of the five comparator RCTs included in the ITC: AUGMENT[49], GADOLIN[50], CONTRALTO[51], GO29365[52], and EORTC 20981[2,3] using the checklist provided in the NICE user guide is provided in Table 10, which also includes the GO29781 pivotal cohort for cross-comparison purposes.

Table 9. GO29781 pivotal cohort assessment using the Downs and Black checklist[48]

Assessment criterion	Score
Reporting	
Is the hypothesis/aim/objective of the study clearly described?	Yes

© Roche (2022). All rights reserved

Are the main outcomes to be measured clearly described in the Introduction or Methods section?	Yes
Are the characteristics of the patients included in the study clearly described?	Yes
Are the interventions of interest clearly described?	Yes
Are the distributions of principal confounders in each group of subjects to be compared clearly described?	Not applicable (single-arm design)
Are the main findings of the study clearly described?	Yes
Does the study provide estimates of the random variability in the data for the main outcomes?	Yes
Have all important adverse events that may be a consequence of the intervention been reported?	Yes
Have the characteristics of patients lost to follow-up been described?	Yes
Have actual probability values been reported (e.g., 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Not applicable (single-arm design, so between-group statistical testing is not relevant)
External validity	
Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Yes (the CSR accounts for all screening failures)
Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	Unable to determine (written informed consent was obtained before any study-specific procedures were performed, so the characteristics of non-consenting patients were not collected)
Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?	Yes (international multicentre study)
Internal validity - bias	
Was an attempt made to blind study subjects to the intervention they have received?	No
Was an attempt made to blind those measuring the main outcomes of the intervention?	No
If any of the results of the study were based on “data dredging”, was this made clear?	Not applicable (pre-planned analyses only)
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	Yes
Were the statistical tests used to assess the main outcomes appropriate?	Yes
Was compliance with the intervention/s reliable?	Yes
Were the main outcome measures used accurate (valid and reliable)?	Yes
Internal validity - confounding (selection bias)	
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Not applicable (single-arm design)

© Roche (2022). All rights reserved

Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Not applicable (single-arm design)
Were study subjects randomised to intervention groups?	Not applicable (single-arm design)
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	Not applicable (single-arm design)
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Yes
Were losses of patients to follow-up taken into account?	Yes
Power	
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	Yes
Abbreviations: CSR, clinical study report Assessment was based on the information available in the GO29781 protocol and CSR	

© Roche (2022). All rights reserved

Table 10. Quality assessment of the five comparator trials included in the ITC

Study name	Was the method used to generate the random allocation adequate?		Was the treatment allocation adequately concealed?		Were the groups similar at baseline?		Were participants and personnel blind to treatment allocation?		Were outcome assessors blind to treatment allocation?		Is there any evidence to suggest that the authors measured more outcomes than they reported?		Did the analysis include an intention to treat analysis?	
	Yes/no/ unclear/ NA	Justification	Yes/no/ unclear /NA	Justification	Yes/no/un clear/NA	Justification	Yes/no/ unclear/ NA	Justification	Yes/no/un clear/NA	Justification	Yes/no/un clear/NA	Justification	Yes/no/ unclear /NA	Justification
GO29781 pivotal cohort	NA	Single-arm design	NA	Single-arm design	NA	Single-arm design	NA	Single-arm design	NA	Single arm design, but outcomes assessed by IRC	No	Full CSR is provided in the reference pack	Yes	ITT, all patients enrolled in the pivotal cohort were included in the analysis
GADOLIN[50]	Yes	Hierarchical dynamic randomisation	Yes	IVRS	Yes		No	Open-label trial	Yes	IRC assessment	No	Specified some outcomes will be reported separately	Yes	ITT all randomised patients were included in the analysis
AUGMENT[49]	Yes	1 to 1 randomisation using IVRS/IWRS	Yes	IVRS/IWRS	Yes		Yes	Double-blind trial	Yes	IRC assessment	No	All outcomes were reported	Yes	ITT all randomised patients were included in the analysis
EORTC 20981[2]	Yes	Stratified randomisation using minimisation	Unclear	No details of allocation concealment	Yes		No	Open-label trial	Unclear	No details of IRC assessment	No	All outcomes were reported	Yes	ITT all randomised patients were included in the analysis
CONTRALTO[51]	Yes	Randomised using stratified permuted block randomisation	Unclear	No details of allocation concealment	Yes		No	Open-label trial	Yes	IRC assessment	No	All outcomes were reported	Yes	ITT all randomised patients were included in the analysis
G029365[52]	Unclear	No details of randomisation methods	Unclear	No details of allocation concealment	Unclear	Abstract with limited baseline data	No	Open-label trial	Yes	IRC assessment	Unclear	Abstract with limited reporting of results	Yes	ITT all randomised patients were included in the analysis

Abbreviations: CSR, Clinical study report; IRC, independent review committee; ITT, intention-to-treat; IVRC, interactive voice response systems; IWRS, interactive web response systems

© Roche (2022). All rights reserved

B.2.6 Clinical effectiveness results of the relevant trials

B.2.6.1 Patient disposition

A total of 90 patients with R/R FL who had received at least 2 prior treatment lines were enrolled into the GO29781 pivotal cohort. All results presented herein are from this patient cohort and are derived from the 27 August 2021 data cut-off (unless explicitly stated otherwise), providing a median follow-up of 18.3 months (range: 2.0–27.5) from first mosunetuzumab dose to study discontinuation date, death or data cut off, whichever is the earliest. Disposition of patients, with regards to initial treatment (as opposed to retreatment) with mosunetuzumab, is summarised in Figure 4.

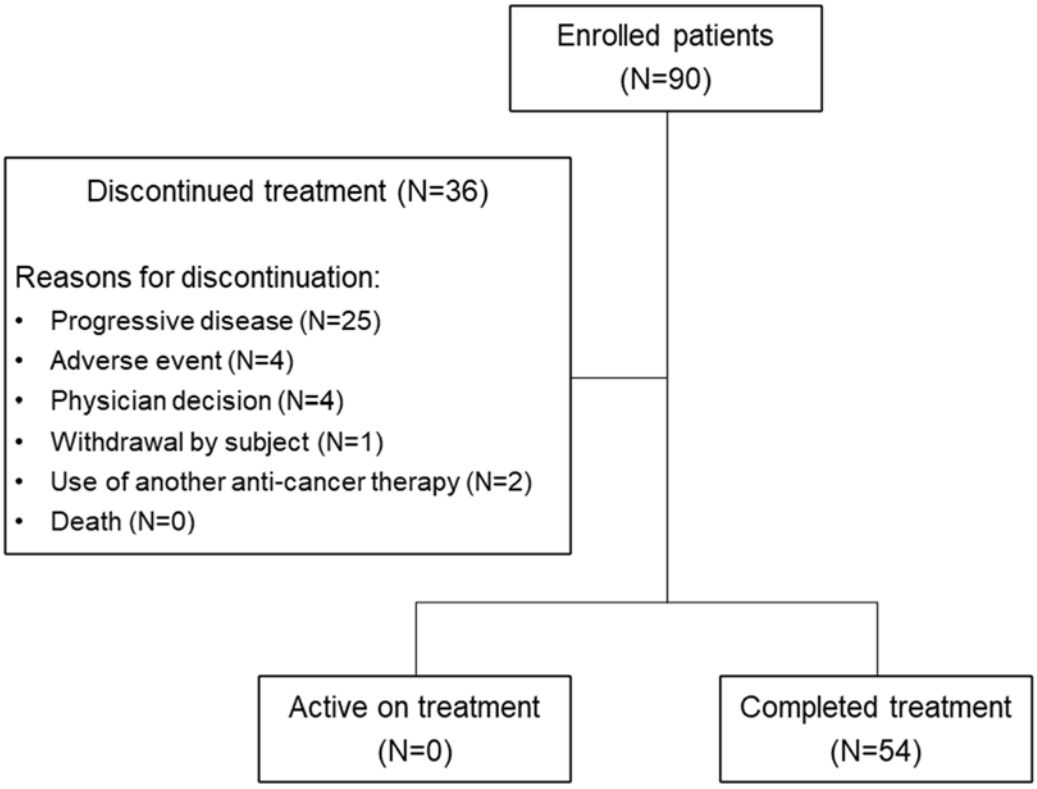


Figure 4. Patient disposition in the GO29781 pivotal cohort

B.2.6.2 Baseline characteristics

Patient characteristics are summarised in Table 11. The enrolled patients had a median age of 60.0 years, approximately a third were aged >65 years, and 61% were male. The patients entered the study at a median of 6.8 years from diagnosis and approximately three quarters had advanced disease, defined as Ann Arbor stage III-IV. Consistent with their long disease duration, the patients were heavily pretreated with a median of 3 and a maximum of 10 prior

© Roche (2022). All rights reserved

therapies. The enrolled population corresponded to a setting of high unmet need in FL. More than half of the patients were double refractory, i.e., refractory to both anti-CD20 therapy and an alkylating agent. Regarding prognostic factors, 52.2% had POD24 after their first systemic treatment, approximately a third of patients had bulky disease and over [redacted] were classified as intermediate or high risk as per FLIPI.

Table 11. Baseline demographics and clinical characteristics in GO29781 pivotal cohort (N=90)

Demographics	
Age	
Median (range), years	60.0 (29–90)
18–65 years, n (%)	[redacted]
>65 years, n (%)	[redacted]
Sex, n (%)	
Male	55 (61.1)
Female	35 (38.9)
Race, n (%)	
White	[redacted]
Asian	[redacted]
Black/African American	[redacted]
American Indian/Alaska Native	[redacted]
Multiple	[redacted]
Unknown	[redacted]
ECOG PS, n (%)	
0	53 (58.9)
1	37 (41.1)
BMI (kg/m2), median (range)	[redacted]
Disease characteristics	
Time from initial diagnosis to first mosunetuzumab dose, median (range), months	[redacted]
Ann Arbor Stage, n (%)	
I	5 (5.6)
II	16 (17.8)
III	25 (27.8)
IV	44 (48.9)
Bulky Disease (>6 cm), n (%)	[redacted]
FLIPI Risk Group, n (%)	
low (0,1)	[redacted]
intermediate (2)	[redacted]
high (3-5)	[redacted]
Prior therapy	
No. of prior lines of anti-lymphoma therapies n (%)	
Median (range)	3.0 (2–10)
2	[redacted]

© Roche (2022). All rights reserved

3	[redacted]
>3	[redacted]
Prior cancer therapy, n (%)	
anti-CD20	90 (100)
alkylating agent	90 (100)
auto-SCT	19 (21.1)
CAR-T	3 (3.3)
PI3K	17 (18.9)
Refractory to: n (%)	
last prior therapy	62 (68.9)
any prior therapy	[redacted]
any prior anti-CD20	71 (78.9)
double refractory to prior anti-CD20 and an alkylating agent	48 (53.3)
POD24 after first systemic therapy, n (%)	47 (52.2)
Abbreviations: BMI, body mass index; CAR-T, chimeric antigen receptor T-cell therapy; ECOG, Eastern Cooperative Oncology Group; FLIPI, Follicular lymphoma international prognostic index; PI3K, Phosphoinositide 3-kinase; POD24, progression of disease within 24 months	

B.2.6.3 Efficacy results

B.2.6.3.1 Response assessment, including primary efficacy endpoint of IRF-assessed CR rate

The primary efficacy endpoint of CR rate as assessed by IRF was met at the primary analysis based on the data cut of 15 March 2021. The CR rate in the pivotal GO29781 cohort was 52/90 patients corresponding to 57.8% (95% CI: 46.9, 68.1). Formal statistical comparison demonstrated that the CR rate of patients in the GO29781 pivotal cohort was significantly greater than the historical control CR rate (57.8% vs 14%; $p < 0.0001$).

As of the 27 August 2021 data cut-off, IRF-assessed CR rate was 60.0% (95% CI: 49.1, 70.2) and IRF-assessed ORR was 80.0% (95% CI: 70.3, 87.7). Investigator-assessed CR rate was 60.0% (95% CI: 49.1, 70.2) and investigator-assessed ORR was 77.8% (95% CI: 67.8, 85.9). Agreement between IRF- and investigator-assessed CR was 93.3% (83/89 patients evaluable for concordance). Further details of response assessment are provided in Table 12.

Table 12. Tumour response in the GO29781 pivotal cohort (N=90)

Best Overall Response by IRF - with or without PET Scan*	
Responders, n (% with 95% CI)	72 (80.0%, 95% CI: 70.3, 87.7)
Non-responders, n (%)	18 (20.0%)
Response classification by IRF	
CR, n (% with 95% CI), primary endpoint	***54 (60.0%, 95% CI: 49.1, 70.2)***
PR, n (% with 95% CI)	[redacted]

© Roche (2022). All rights reserved

SD, n (% with 95% CI)	[redacted]
PD**, n (% with 95% CI)	[redacted]
Not evaluable	[redacted]
Missing or not done	[redacted]
Best Overall Response by Investigator - with or without PET Scan	
Responders, n (% with 95% CI)	70 (77.8%, 95% CI: 67.8, 85.9)
Non-responders, n (%)	20 (22.2%)
Response classification by Investigator	
CR, n (% with 95% CI)	54 (60.0%, 95% CI: 49.1, 70.2)
PR, n (% with 95% CI)	[redacted]
SD, n (% with 95% CI)	[redacted]
PD**, n (% with 95% CI)	[redacted]
Not evaluable	[redacted]
Missing or not done	[redacted]
* FDG PET and CT scans were required for response assessments during study treatment and CT scans with or without PET scans could be utilised during post-treatment follow up ** PD includes missing, not evaluable and not done (ND) assessments where the patient has otherwise had a PD recorded Abbreviations: CI, confidence interval; CR, complete response; IRF, independent review facility; PD, progressive disease; PET, positron emission tomography; PR, partial response; SD, stable disease	

B.2.6.3.2 Duration of response

B.2.6.3.2.1 DOR per IRF

Of the 72 patients who achieved a response (CR or PR) as determined by the IRF, 29 patients (40.3%) subsequently had disease progression or died. After a median follow-up of 14.9 months (95% CI: 13.4, 16.6) from the time of response, median DOR was estimated at 22.8 months (95% CI: 9.7, not evaluable [NE]). However, this estimate was based on <10% of responders remaining at risk and should be interpreted with caution. At 12 and 18 months, 61.8% and 56.9% of patients, respectively, remained in response. The Kaplan-Meier plot of DOR per IRF is provided in Figure 5.

© Roche (2022). All rights reserved

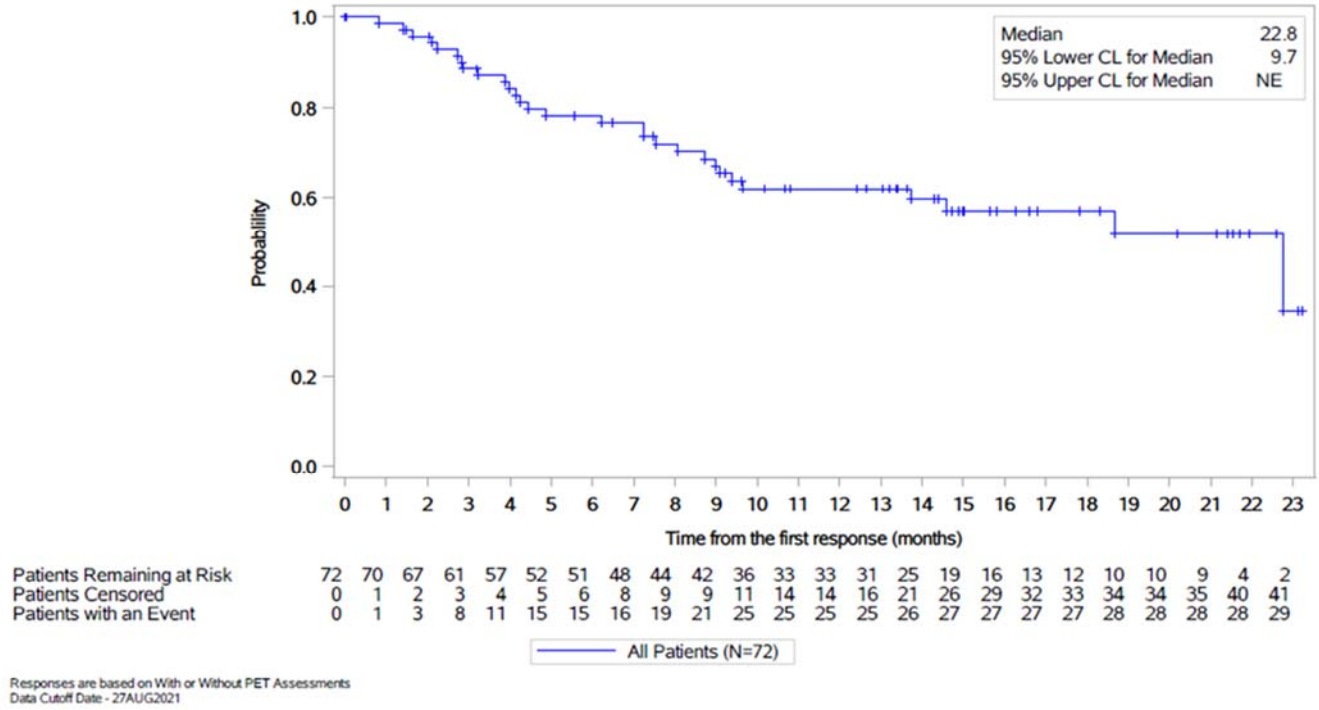


Figure 5. DOR per IRF in the GO29781 pivotal cohort (n=72 patients who achieved a response out of the total N=90)
Abbreviations: CL, confidence level; NE, not evaluable; PET, positron emission tomography

B.2.6.3.2.2 DOR per investigator

In general, investigator assessment of DOR was consistent with IRF assessment. Of the 70 patients who achieved a response as determined by the investigator [redacted] subsequently experienced disease progression or died. The median DOR was estimated at [redacted] months ([redacted]), however this estimate should be interpreted with caution as it was based on <10% of responders remaining at risk (as mentioned above, median follow-up for DOR was 14.9 months). A total of [redacted] and [redacted] of patients, respectively, remained in response at 12 and 18 months.

B.2.6.3.2.3 Duration of CR per IRF

Among the 54 patients who achieved a CR as assessed by the IRF, [redacted] patients [redacted] subsequently had disease progression. At 12 and 18 months, [redacted] and [redacted] of patients, respectively, remained in CR. [redacted] The Kaplan-Meier plot of CR duration per IRF is provided in Figure 6.

[redacted]

Figure 6. Duration of CR per IRF in the GO29781 pivotal cohort (n=54 patients who achieved a CR out of the total N=90)
Abbreviations: CL, confidence level; NE, not evaluable; PET, positron emission tomography

© Roche (2022). All rights reserved

B.2.6.3.2.4 Duration of CR per investigator

Among the 54 patients who achieved a CR as determined by the investigator, [redacted] patients ([redacted]) subsequently had disease progression. [redacted]. Of patients achieving CR per investigator assessment, [redacted] and [redacted] respectively remained in CR at 12 and 18 months.

B.2.6.3.3 PFS

B.2.6.3.3.1 PFS per IRF assessment

At the time of the data cut-off, 42/90 patients (46.7%) had a PFS event as assessed by the IRF, including 41 patients experiencing disease progression and one death. Median PFS was 17.9 months (95% CI: 10.1, NE). The 12- and 18-month PFS rates were 57.7% (95% CI: 46.9, 68.4) and 47.0% (95% CI: 34.4, 59.6). The Kaplan-Meier plot for IRF-assessed PFS is provided in Figure 7.

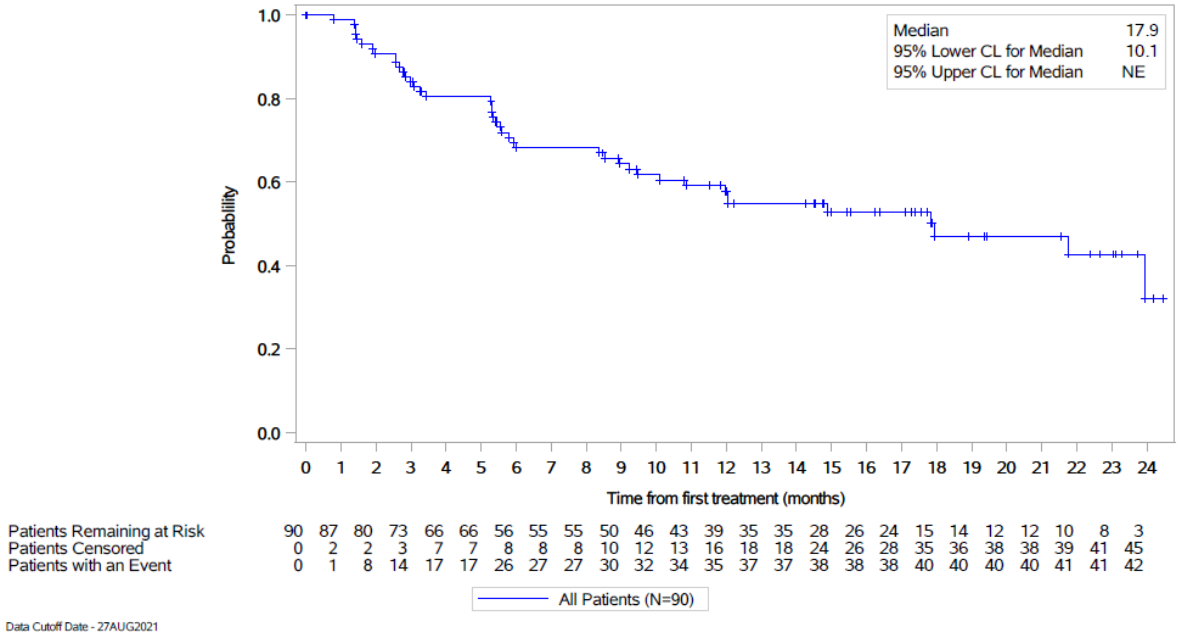


Figure 7. PFS per IRF assessment in the GO29781 pivotal cohort (N=90)
Abbreviations: CL, confidence level; NE, not evaluable

B.2.6.3.3.2 PFS per investigator assessment

PFS estimate per investigator assessment was similar to the IRF assessment. At the time of the data cut-off, [redacted] patients ([redacted]) in the had a PFS event (disease progression n=[redacted], or death n=[redacted]) and median PFS was [redacted] months (95% CI: [redacted]). The 12- and 18-month PFS rates were [redacted] and [redacted]

© Roche (2022). All rights reserved

B.2.6.3.4 OS

By the time of the data cut-off, [redacted] of the 90 patients had died. [redacted] The Kaplan-Meier plot for OS is provided in Figure 8.

[redacted]

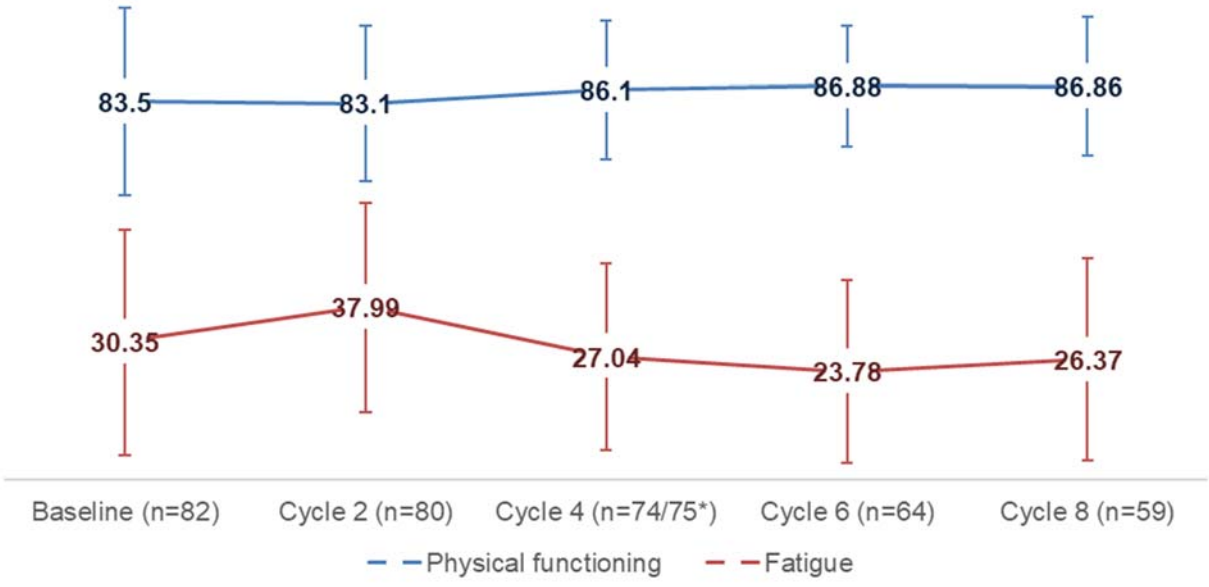
Figure 8. OS in the GO29781 pivotal cohort (N=90)

B.2.6.4 Patient-reported outcomes

Compliance with PRO assessments was high, reaching ≥75% at all scheduled assessments. PRO data are presented below until Cycle 8 only due to the low number of evaluable patients thereafter (which is expected given that only selected patients were eligible to continue treatment beyond Cycle 8, see Section B.2.3.4.1.2).

B.2.6.4.1 EORTC QLQ-C30

The baseline mean (SD) physical functioning and fatigue scores from the EORTC QLQ C30 questionnaire (n=82 evaluable patients) were 83.5 (20.7) and 30.4 (24.9) respectively, indicating normal physical functioning but slightly elevated fatigue levels. At all post-baseline assessments, the baseline scores were maintained in patients who continued to receive treatment with mosunetuzumab (Figure 9). The mean change from baseline physical functioning score at completion or discontinuation of mosunetuzumab treatment was -0.3 (SD: 19.7) and for fatigue it was -1.1 (SD: 28.4) across 68 evaluable patients.



© Roche (2022). All rights reserved

Figure 9. EORTC QLQ-C30 physical functioning and fatigue scores from baseline (prior to first mosunetuzumab infusion) through Cycle 8
*n=74 for physical functioning and 75 for fatigue
Error bars indicate the standard deviation.
Note that for fatigue lower scores indicate better outcomes, whereas for physical functioning it is higher scores that indicate better outcomes (see Section B.2.3.5.1).

The proportion of patients with a clinically meaningful improvement (defined as ≥10-point increase) in physical functioning score relative to baseline ranged from 7.3% at Cycle 2 (n=82 patients evaluable) to 15.8% at Cycle 4 (n=76 patients evaluable). At completion or discontinuation of mosunetuzumab treatment, 8 of 68 evaluable patients (11.8%) had a clinically meaningful improvement in physical functioning.

For fatigue scores, the proportion of patients achieving a clinically meaningful improvement (defined as ≥10-point increase) from baseline ranged from 23.2% at Cycle 2 (n=82 patients evaluable) to 37.9% at Cycle 6 (n=66 patients evaluable). At completion or discontinuation of mosunetuzumab treatment, 31 of 68 evaluable patients (45.6%) had a clinically meaningful improvement in fatigue.

B.2.6.4.2 FACT-Lym

The baseline mean FACT-Lym subscale score (n=81 evaluable patients) was 47.6 (SD: 8.5) indicating some burden of lymphoma-specific symptoms or concerns at baseline. At all post-baseline assessments, the baseline score was maintained in patients who continued to receive treatment with mosunetuzumab (Figure 10). Among the 67 patients evaluable, the mean change from baseline FACT-Lym score at completion or discontinuation of mosunetuzumab treatment was 1.9 (SD: 8.9).

© Roche (2022). All rights reserved

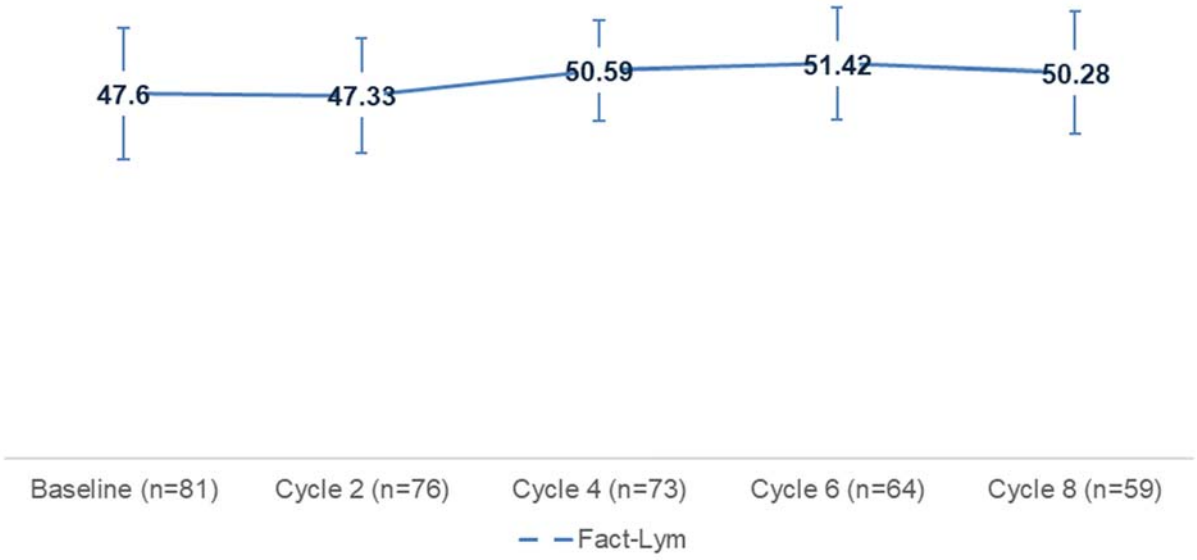


Figure 10. FACT-Lym scores from baseline (prior to first mosunetuzumab infusion) through Cycle 8
Error bars indicate the standard deviation.

The proportion of patients with a clinically meaningful improvement in lymphoma symptoms from baseline (defined as ≥3-point increase) ranged from 28.4% at Cycle 2 (n=81 evaluable patients) to 44.0% at Cycle 4 (n=75 evaluable patients). At completion or discontinuation of mosunetuzumab treatment, 28 of 67 evaluable patients (41.8%) had a clinically meaningful improvement of ≥3-points from baseline.

B.2.6.4.3 EQ-5D-5L

Baseline EQ-5D-5L scores indicated low levels of impairment in health status. Among 81 evaluable patients, the baseline mean EQ-5D-5L index utility scores for each dimension ranged from 1.2 (SD: 0.7) for self-care to 1.9 (SD: 0.9) for pain/discomfort. The baseline mean VAS score among 78 evaluable patients was 73.7 (SD: 20.2).

EQ-5D-5L index and VAS scores were maintained at all post baseline assessments in patients who continued receiving mosunetuzumab (see Figure 11 for EQ-5D VAS). The mean change from baseline to completion or discontinuation of mosunetuzumab treatment in EQ-5D-5L index utility scores ranged between -0.01 (SD: 0.7) for self-care to -0.2 for usual activities (SD: 1.3) and -0.2 (SD: 0.8) for pain/discomfort. For EQ-5D VAS score, the corresponding change from baseline was 4.2 (SD: 22.0, n=65 evaluable patients).

© Roche (2022). All rights reserved

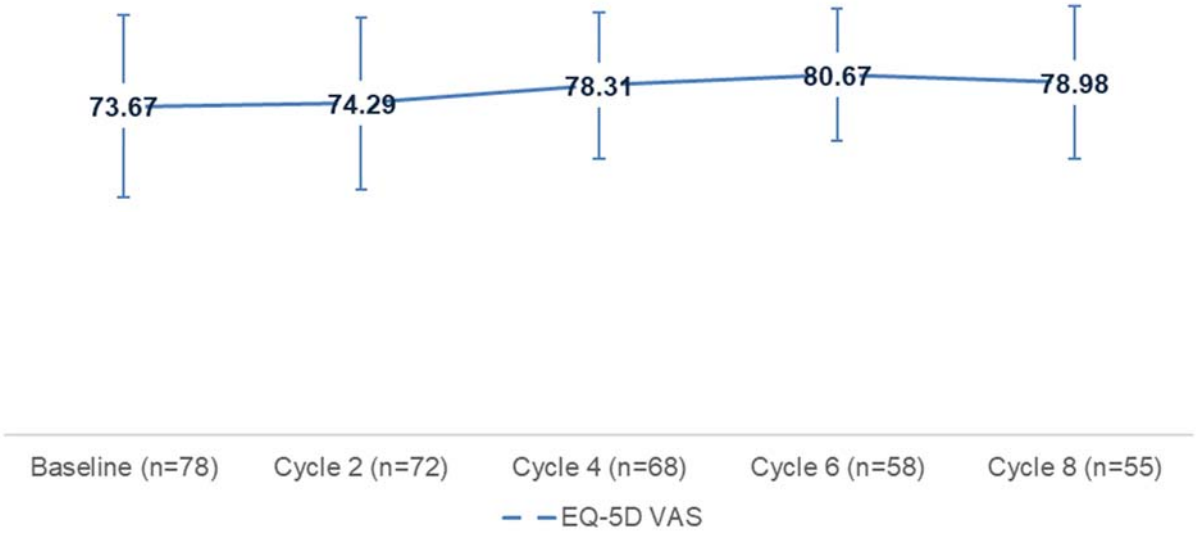


Figure 11. EQ-5D-5L VAS scores from baseline (prior to first mosunetuzumab infusion) through Cycle 8
Error bars indicate the standard deviation.

B.2.7 Subgroup analysis

CR rates and ORR for all subgroups were generally consistent with the overall results from the pivotal cohort, demonstrating that the effects of mosunetuzumab were similar across key subpopulations defined by demographic, baseline disease characteristics, prior treatment history (number of prior systemic therapies and refractory status to those prior treatments), and prognostic factors. Subgroup analyses are presented in Figure 12 below.

© Roche (2022). All rights reserved

Figure 12. Subgroup Analyses of ORR and CR Rate (per IRF assessment) in the GO29781 pivotal cohort (N=90)
Abbreviations: ALKY, alkylating; BMI, body mass index; CAR-T, chimeric antigen receptor T-cell therapy; CI, confidence interval; CR, complete response; ECOG, Eastern Cooperative Oncology Group; EZH2, Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit; FLIPI, Follicular lymphoma international prognostic index; ORR, objective response rate; PI3K, Phosphoinositide 3-kinase; PD, progression of disease

© Roche (2022). All rights reserved

B.2.8 Meta-analysis

At the time of submission, clinical evidence supporting the use of mosunetuzumab for the treatment of FL was available solely from the pivotal cohort of the GO29781 study, so no meta-analysis was performed.

B.2.9 Indirect and mixed treatment comparisons

B.2.9.1 Indirect treatment comparison methods

B.2.9.1.1 Overview of ITC methods

In the absence of head-to-head data comparing mosunetuzumab with relevant comparators described in the NICE scope, a series of ITCs was conducted to estimate the relative efficacy of mosunetuzumab (based on the pivotal cohort of the GO29781 study) and its key comparators. Given the single-arm design of GO29781, MAICs were conducted for those comparators for which only published aggregate data were available, and propensity score analyses were conducted for comparators with available IPD.

All analyses were conducted using R statistical software. For details of the ITC methodology and additional scenario results beyond those presented in this submission, please see the ITC report provided as Appendix E.

B.2.9.1.1.1 MAIC

The GO29781 pivotal cohort population was aligned in terms of inclusion/exclusion criteria to that of the comparator study and the IPD from GO29781 pivotal cohort were weighted to match prognostic factors from the comparator study (where reported). The matching-adjusted data were then used to provide an estimate of the outcomes that might have occurred if the comparator studies had included a mosunetuzumab arm. An iterative approach to the MAIC was employed to identify the most appropriate base-case analysis, which maximised the bias/variance trade-off whilst controlling for as many high priority prognostic factors as possible and, where feasible, controlling for continuous outcomes as means rather than medians or proportions.

B.2.9.1.1.2 Propensity score analysis

Propensity score analyses provide an estimate of treatment effect after accounting for differences in covariates believed to be potential prognostic factors or treatment-effect modifiers across treatment groups. The eligibility criteria of the mosunetuzumab and comparator populations were aligned prior to re-weighting the IPD or matching. The preferred target estimand was the average treatment effect (ATE) and both matching on the propensity

© Roche (2022). All rights reserved

score and inverse probability of treatment weighting (IPTW) methodologies were used to minimise imbalances between mosunetuzumab and comparator groups, as recommended in NICE DSU TSD 17[36]. The matching method resulting in better covariance balance (i.e., the one that minimised the absolute standardised mean differences and complements of overlapping coefficients for the greatest number of covariates with major emphasis on the ones with higher prognostic value) was selected as the preferred matching method for the base case scenario.

B.2.9.1.2 Data sources

Based on an SLR and feasibility assessment (see Section B.2.1 and Appendix D for details), the following ITCs were performed against the three comparators specified in the NICE scope:

- A MAIC against **rituximab plus lenalidomide (R^2)**, based on the AUGMENT trial[49]. Note that although the MAGNIFY trial[53] was also identified as potentially feasible for inclusion in a MAIC, only the most appropriate study was used where multiple studies for a given comparator were identified, and for R^2 this was determined to be AUGMENT. While an ITC against MAGNIFY was considered, the data has only been reported as an abstract, and the available information for the FL population from this trial did not allow a reliable comparison to be conducted (see the ITC report, provided as Appendix E for details). Furthermore, the criteria used to assess response in MAGNIFY were older than those used in the GO29781 trial.
- A propensity score analysis against **obinutuzumab plus bendamustine** based on the GADOLIN trial[50].
- Two analyses against **rituximab plus chemotherapy**:
    - A propensity score analysis against rituximab plus bendamustine based on CONTRALTO[51] and GO29365[52] studies.
    - A propensity score analysis against rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) was attempted based on the EORTC 20981 trial[2,3], but proved not to be feasible due to several limitations associated with small sample sizes, unavailability of information on some important prognostic factors, and important residual imbalances after adjusting for differences in the available factors (please see the ITC report, provided as Appendix E, for details).

The comparator studies included in ITCs described in this submission are summarised in Table 13. Note, that for studies with available IPD, only the size of the population

© Roche (2022). All rights reserved

corresponding to anticipated mosunetuzumab label (i.e., R/R FL after ≥2 treatment lines) is reported in the table.

Mosunetuzumab data for the ITC was derived from the pivotal cohort of GO29781 (N=90), which included patients with R/R FL who received ≥2 prior lines of systemic therapy and is described in detail in this submission.

© Roche (2022). All rights reserved

Table 13. List of performed ITCs

Comparator	Study name	Study design	Analysis population	ITC type	Results location in the submission	Likely direction of bias in the ITC
R^2	AUGMENT[49]	Phase III, multicentre, randomized trial of R^2 versus placebo + rituximab in patients with R/R FL or marginal zone lymphoma. Treatment was administered for 12 cycles with no maintenance therapy.	Patients with R/R FL (n=147)*	Unanchored MAIC	B.2.9.2.1	Favouring R^2 due to the inclusion of patients at earlier treatment line (i.e., relapsing or refractory to first-line therapy) and exclusion of patients with rituximab-refractory disease.
Obinutuzumab plus bendamustine	GADOLIN[50] (IPD used)	Open-label, randomised, phase III study of obinutuzumab + bendamustine vs bendamustine alone in patients with R/R CD20-positive indolent NHL who were refractory to rituximab. Induction treatment was administered for 6 cycles, followed by up to 2 years of obinutuzumab maintenance for patients in the obinutuzumab + bendamustine group who did not progress.	80 patients with R/R FL who had received ≥2 prior lines of treatment	Propensity score analysis	B.2.9.2.2	Favouring obinutuzumab and bendamustine for PFS and potentially OS due to the inclusion of maintenance treatment for patients not experiencing disease progression.
Rituximab plus bendamustine	G029365[52] and CONTRALTO[51] (IPD used)	CONTRALTO: open-label phase II study of venetoclax + rituximab or venetoclax + rituximab + bendamustine vs rituximab + bendamustine in patients with R/R FL. G029365: randomised phase II trial of polatuzumab vedotin + rituximab + bendamustine vs rituximab + bendamustine alone in transplant-ineligible patients with R/R FL or DLBCL Patients in the rituximab + bendamustine arm of both studies received 6 cycles of treatment.	48 patients with R/R FL who had received ≥2 prior lines of treatment	Propensity score analysis	B.2.9.2.3	Favouring rituximab and bendamustine, since even after optimal pair matching, there were notable differences in important prognostic factors such as ECOG 1 vs 0%, refractory to last therapy line, prior ASCT and time since completion of last therapy (Table 16).

© Roche (2022). All rights reserved

R-CHOP	EORTC 20981[2,3] (IPD used)	Randomised phase III trial of R-CHOP vs CHOP alone. Patents achieving CR or PR after 6 cycles of CHOP or R-CHOP were randomised to observation or rituximab maintenance for up to 2 years.	51 patients with R/R FL who had received ≥2 prior lines of treatment	Propensity score analysis	Not applicable. An adjusted analysis could not be performed due to differences in patient characteristics. Please see the ITC report (Appendix E for details and for results of a naïve comparison.	Favouring R-CHOP for PFS and potentially OS due to the inclusion of maintenance treatment for patients not experiencing disease progression. Note this ITC was not feasible to conduct due to methodological limitations.
Abbreviations: CR, complete remission; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; IPD, individual patient data; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; NHL, non-Hodgkin lymphoma; OS, overall survival; PFS, progression-free survival; PR, partial remission; (R-)CHOP, (rituximab), cyclophosphamide, doxorubicin, vincristine, and prednisone; R^2, rituximab plus lenalidomide; R/R, relapsed or refractory *Note, that 43% of all patients (FL and non-FL) in the R^2 arm of AUGMENT received ≥2 prior lines of systemic therapy						

© Roche (2022). All rights reserved

B.2.9.1.3 Outcomes included in the analyses

The outcomes of interest included OS and PFS as time-to-event endpoints, as well as ORR, CR, and treatment discontinuation due to AEs as binary endpoints. Endpoints reported in G029781 were matched to the definitions available from the comparator trials whenever possible.

B.2.9.1.4 Prognostic factors and effect modifiers

Prognostic factors and effect modifiers were classified as either high priority, low priority, or deprioritised according to clinical feedback. High-priority prognostic factors and effect modifiers included:

- Number of previous chemotherapeutic (or systemic) agents, (e.g., 3 vs. >3 [no clinically established threshold] or median, if categories not reported)
- Refractoriness to last previous therapy (yes/no)
- Refractoriness to any prior anti-CD20 antibody-containing therapy (yes/no), also used as proxy for rituximab refractoriness when needed
- Early relapse status (POD24) (yes/no)
- Prior ASCT (yes/no)
- Size of the largest lymph node lesion involved (prioritized over bulky disease, when possible)
- Bulky disease (yes/no)
- FLIPI risk group (high [≥3] versus intermediate/low [<3])
- Age (mean, or median if mean not reported, or % ≥60 years (when feasible), if neither reported)
- Ann Arbor Stage (I–II vs III–IV)
- High lactate dehydrogenase (LDH) (yes/no)
- Bone marrow involvement (yes/no, as demonstrated by bone marrow biopsy)
- Low haemoglobin (Hb) level (yes/no, e.g., <12 [or 12.5] g/dL, or <lower limit of normal [LLN])
- Double refractoriness to both a rituximab-containing regimen and an alkylating agent was considered to be of unclear priority but potentially useful for identifying a high-risk population; therefore, it was controlled for when available.

© Roche (2022). All rights reserved

If key covariates were defined differently in G029781 and the comparator trials, attempts to readjust the covariate definitions in G029781 were made, where feasible.

B.2.9.2 ITC Results

Only base cases analysis results for each ITC are presented in the following sections. Please consult the ITC report (Appendix E) for results of the scenario analyses performed.

B.2.9.2.1 Mosunetuzumab vs R^2 MAIC

B.2.9.2.1.1 Populations and baseline characteristics

The population from AUGMENT used for the MAIC included all patients with R/R FL in the R^2 arm, regardless of the number of prior therapies received (n=147). To align with the inclusion criterion of non-rituximab-refractory disease in AUGMENT, a comparison with the subgroup of GO29781 pivotal cohort patients who were not refractory to an anti-CD20 antibody was initially attempted. However, this resulted in a very small sample size after filtering and prior to any adjustment. Consequently, all analyses were performed in the full pivotal cohort of GO29781 (n=90), which introduced substantial bias against mosunetuzumab that should be taken into account when interpreting the results. Baseline characteristics before and after weighting are presented in Table 14.

Table 14. Pre- and post-weighting baseline characteristics in the mosunetuzumab vs R^2 MAIC

Variable	[redacted]	[redacted]	[redacted]
Age (mean)			
ECOG (1 vs 0) (%)			
FLIPI ≥3 (Yes) (%)			
Ann Arbor Stage III–IV (Yes) (%)			
Refractory to last line (%)			
High LDH (Yes) (%)			
Low Hb level (Yes) (%)			
Bone marrow involvement (Yes) (%)			
POD24 (Yes) (%)			
Bulky disease (Yes) (%)			
Time since completion of last therapy >2 years (yes) (%)			
Presence of B-symptoms (Yes) (%)			

Abbreviations: ASCT, autologous stem cell transplant; ECOG, Eastern Cooperative Oncology Group Performance Status; ESS effective sample size; FLIPI, Follicular lymphoma international prognostic index; Hb, haemoglobin; LDH, lactate dehydrogenase; Mosun., mosunetuzumab; POD24, Progression of disease within 24 months

© Roche (2022). All rights reserved

B.2.9.2.1.2 Response rates (per IRF assessment)

Tumour responses were assessed using the International Working Group criteria[37] in both GO29781 and AUGMENT[37].

[redacted]

B.2.9.2.1.3 PFS and OS

The Kaplan-Meier plots for PFS and OS are presented in Figure 13 and Figure 14, respectively.

[redacted]

Figure 13. PFS (per IRF assessment) in the mosunetuzumab vs R^2 MAIC
Abbreviations: R-Len, rituximab plus lenalidomide

Figure 14. OS in the mosunetuzumab vs R^2 MAIC
Abbreviations: R-Len, rituximab plus lenalidomide

B.2.9.2.1.4 Safety

With regards to safety, the OR for discontinuation due to AEs numerically [redacted] models, but the results were not statistically significant.

© Roche (2022). All rights reserved

B.2.9.2.2 Mosunetuzumab vs obinutuzumab plus bendamustine propensity score analysis

B.2.9.2.2.1 Populations and baseline characteristics

The population used for indirectly comparing mosunetuzumab with obinutuzumab plus bendamustine was a cohort of patients with R/R FL who had received at least 2 prior systemic therapies from the obinutuzumab plus bendamustine arm of the GADOLIN trial (n=80). To ensure that the patient cohorts used for the analyses were as homogeneous as possible before attempting any indirect comparisons, a filtering procedure based on applying common eligibility criteria was adopted. This involved excluding patients with ECOG PS 2 from the GADOLIN cohort to align with the eligibility criteria of the GO29781 trial, and excluding patients who were not refractory to a prior rituximab-containing regimen from the GO29781 pivotal cohort to align with the eligibility criteria of the GADOLIN trial. This resulted in 71 patients in the mosunetuzumab arm and 77 patients in the obinutuzumab plus bendamustine arm being included in the ITC.

Prior to adjustment, several baseline characteristics were imbalanced between the mosunetuzumab and obinutuzumab plus bendamustine groups, as evidenced by an absolute standardised mean difference [aSMD] >0.1 (Table 15). The balance between groups improved for many covariates following both matching on the propensity score and IPTW, but the improvement was most pronounced after IPTW, which was therefore selected as the preferred adjustment method for this comparison. The IPTW-adjusted results are presented in the remainder of this section. Please see the ITC report (Appendix E) for results adjusted based on matching on the propensity score.

© Roche (2022). All rights reserved

Table 15. Unadjusted and IPTW-adjusted baseline characteristics in the propensity score analysis of mosunetuzumab vs obinutuzumab plus bendamustine

Variable	Unadjusted				aSMD	IPTW-adjusted				aSMD
	[redacted]		[redacted]			[redacted]		[redacted]		
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Age (mean)	[redacted]									
ECOG PS (1 vs 0) (%)										
FLIPI ≥3 (Yes) (%)										
Ann Arbor Stage III/IV (Yes) (%)										
Prior therapies ≥3 (%)										
Refractory to last line (Yes) (%)										
Double refractory (yes) (%)										
POD24 (Yes) (%)										
Prior ASCT (Yes) (%)										
Size of the largest node lesion [cm] (mean)										
Low Hb (Yes) (%)										
High LDH (Yes) (%)										
Bone marrow involvement (Yes) (%)										
Presence of B symptoms (Yes) (%)										
Time since completion of last therapy [months] (mean)										

Abbreviations: ASCT, autologous stem cell transplant; ECOG, Eastern Cooperative Oncology Group Performance Status; ESS effective sample size; FLIPI, Follicular lymphoma international prognostic index; Hb, haemoglobin; LDH, lactate dehydrogenase; POD24, Progression of disease within 24 months

© Roche (2022). All rights reserved

B.2.9.2.2.2 Response rates (per IRF assessment)

In both GO29781 and GADOLIN, tumour response was assessed using the International Working Group criteria described by Cheson et al[37]. The OR for CR strongly and significantly favoured [redacted] but neither result was statistically significant.

B.2.9.2.2.3 PFS and OS

Kaplan-Meier plots of PFS and OS are available in Figure 15 and Figure 16, respectively. The HR for IRF-assessed PFS significantly favoured [redacted] but the results were not statistically significant.

B.2.9.2.2.4 Safety

For discontinuation due to AEs, the OR strongly and significantly favoured [redacted] in which the odds of discontinuation were not statistically different between the two treatments.

© Roche (2022). All rights reserved

Figure 15. PFS (per IRF assessment) in the mosunetuzumab vs obinutuzumab plus bendamustine propensity score analysis

Abbreviations: G-BENDA, obinutuzumab plus bendamustine; IPTW, inverse probability of treatment weighting; MOSUN, mosunetuzumab

Figure 16. OS in the mosunetuzumab vs obinutuzumab plus bendamustine propensity score analysis

Abbreviations: G-BENDA, obinutuzumab plus bendamustine; IPTW, inverse probability of treatment weighting; MOSUN, mosunetuzumab

© Roche (2022). All rights reserved

B.2.9.2.3 Mosunetuzumab vs rituximab plus bendamustine propensity score analysis

B.2.9.2.3.1 Populations and baseline characteristics

For the ITC between mosunetuzumab and rituximab plus bendamustine, the comparator data were derived from a combination of patients with R/R FL who had received ≥2 prior therapies enrolled in the rituximab plus bendamustine arms of CONTRALTO and GO29365 trials. To ensure optimal homogeneity of the cohorts used for the analyses before attempting any indirect comparisons, a filtering procedure based on applying common eligibility criteria was applied. This involved excluding patients with ECOG PS 2 from the CONTRALTO and GO29365 trial cohorts to align with the eligibility criteria of the GO29781 trial, and excluding patients who received >5 prior anticancer regimens from the GO29781 pivotal cohort, although this was not a formal eligibility criterion in CONTRALTO and GO29365. The filtering resulted in 81 patients in the mosunetuzumab arm and 46 patients in the rituximab plus bendamustine arm being included in the ITC.

Several potentially prognostic baseline characteristics of the CONTRALTO/GO29365 and GO29781 patient cohorts were imbalanced prior to adjustment (aSMD >0.1, Table 16). Optimal pair matching resulted in the greatest number of balanced covariates compared with other methods and was selected as the preferred adjustment method for this comparison. The result based on optimal pair matching are presented below. For results based on IPTW, please see the ITC report (Appendix E).

© Roche (2022). All rights reserved

Table 16. Unadjusted and optimal pair matching-adjusted baseline characteristics in the propensity score analysis of mosunetuzumab vs rituximab plus bendamustine

Variable	Unadjusted					Optimal pair matching-adjusted				
	[redacted]		[redacted]		aSMD	[redacted]		[redacted]		aSMD
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Age (mean)	[redacted]									
ECOG PS (1 vs 0) (%)	[redacted]									
FLIPI ≥3 (Yes) (%)	[redacted]									
Ann Arbor Stage III/IV (Yes) (%)	[redacted]									
Prior therapies ≥3 (%)	[redacted]									
Refractory to last line (Yes) (%)	[redacted]									
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]									
Double refractory (Yes) (%)	[redacted]									
POD24 (Yes) (%)	[redacted]									
Bone marrow involvement (Yes) (%)	[redacted]									
Prior ASCT (Yes) (%)	[redacted]									
Size of the largest node lesion [cm] (mean)	[redacted]									
Low Hb (Yes) (%)	[redacted]									
High LDH (Yes) (%)	[redacted]									
Time since completion of last therapy [months] (mean)	[redacted]									

Abbreviations: ASCT, autologous stem cell transplant; ECOG, Eastern Cooperative Oncology Group Performance Status; ESS effective sample size; FLIPI, Follicular lymphoma international prognostic index; Hb, haemoglobin; LDH, lactate dehydrogenase; POD24, Progression of disease within 24 months

© Roche (2022). All rights reserved

B.2.9.2.3.2 Response rates (per investigator assessment)

The criteria used for tumour response assessment differed between GO29781, where the International Working Group criteria[37] were used, and CONTRALTO and GO29365, both of which used the more recent Lugano criteria[12]. This difference in the criteria for tumour response assessment adds to the uncertainty associated with the comparison.

The OR for CR [redacted], but none of the results were statistically significant.

B.2.9.2.3.3 PFS and OS

[redacted] The Kaplan-Meier plot for PFS is available in Figure 17 and for OS in Figure 18. The HR for investigator-assessed PFS numerically favoured [redacted] For OS, the HR numerically favoured [redacted].

B.2.9.2.3.4 Safety

The odds of discontinuation due to AEs were numerically lower for [redacted].

However, none of the results were statistically significant.

© Roche (2022). All rights reserved

Figure 17. PFS (per investigator assessment) in the mosunetuzumab vs rituximab plus bendamustine propensity score analysis

Abbreviations: BR, rituximab plus bendamustine; MOSUN, mosunetuzumab

Figure 18.OS in the mosunetuzumab vs rituximab plus bendamustine propensity score analysis

Abbreviations: BR, rituximab plus bendamustine; MOSUN, mosunetuzumab

© Roche (2022). All rights reserved

B.2.9.3 Discussion of ITC results

B.2.9.3.1 Summary of results

The ITCs described above provided statistically significant evidence in support of [redacted]. In the propensity score analysis vs rituximab plus bendamustine, the evidence numerically favoured [redacted].

For ORR, [redacted]

In the PFS analyses, a numerical trend was observed favouring [redacted] 15.

In the OS analyses [redacted].

In the safety analyses comparing discontinuation rates due to AEs, the odds of discontinuation were [redacted].

Top-line results of the ITCs are visualised in Figure 19 below.

© Roche (2022). All rights reserved

[redacted]

Figure 19. Summary of ITC results
Abbreviations: AE, adverse event; CR, complete response; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; R^2, rituximab plus lenalidomide

B.2.9.3.2 Limitations and uncertainties

In the absence of head-to-head data, the ITCs described in this submission provide evidence supporting the role of mosunetuzumab as an important option for patients with R/R FL. Several limitations should, however, be taken into account when interpreting the ITC results:

- The comparison between mosunetuzumab and R-CHOP was not feasible; therefore, the only rituximab plus chemotherapy option assessed in the ITCs was rituximab plus bendamustine.
- Despite filtering by common eligibility criteria and statistical adjustment, some differences between trial population remained and it was often not possible to adjust on all predefined prognostic factors and effect modifiers.
- The effective sample sizes after adjustment were relatively small. Together with the low number of events observed for some endpoints, this may have contributed to the often wide CIs observed, limiting the interpretation of the results.
- There were some notable differences in design between GO29781 and comparator trials that could bias the ITC results against mosunetuzumab (see Table 13), including maintenance treatment in non-progression patients and enrolment of patients relapsing after first-line therapy (note that outcomes in FL worsen with the number of prior therapies received[8,35]).
    - For the MAIC vs R^2, it was not feasible to fully harmonise the eligibility criteria between AUGMENT[49] and GO29781, which introduced an important bias, as many patients in the AUGMENT trial relapsed following first-line therapy and the overall (FL and non-FL) population of the AUGMENT trial contained only 47% of patients who were 3L+ patients[49]. In contrast, the GO29781 pivotal cohort was only open to patients who had received ≥2 prior therapies. Furthermore, AUGMENT excluded rituximab-refractory patients[49], while in the GO29781 pivotal cohort 78.9% of patients were refractory to prior anti-CD20 therapy (see Section B.2.6.2). [redacted]
    - In the comparison vs GADOLIN[50], PFS was significantly better in the comparator trial compared with the GO29781 pivotal cohort. This is not

© Roche (2022). All rights reserved

surprising given that, following completion of initial treatment lasting approximately 6 months, non-progressing patients in the obinutuzumab + bendamustine arm of GADOLIN received up to 2 years of maintenance therapy with obinutuzumab[50]. In contrast, treatment duration in the GO29781 pivotal cohort was fixed at 8 or 17 cycles (see Section B.2.3.4.1.2), corresponding to approximately 6 months and 1 year, respectively. The inclusion of maintenance treatment in GADOLIN is likely to delay FL progression, resulting in favourable PFS (and potentially OS) compared with the GO29781 pivotal cohort, which is an important source of bias favouring the comparator in this ITC.

- For the comparison vs rituximab plus bendamustine, even after optimal pair matching, there were notable differences in key prognostic factors such as ECOG PS, refractory status to last therapy line, receipt of prior ASCT, and time since completion of last therapy (Table 16). [redacted]

- The data for mosunetuzumab is very immature relative to comparator data, and the conclusions from the ITC may change as longer follow-up becomes available and more events are observed for the endpoints of interest.

Overall, due to the aforementioned limitations the ITC presented is unlikely to provide a true reflection of the relative efficacy of mosunetuzumab vs the comparators specified in the NICE scope. The immaturity of GO29781 data at the time of submission and the lack of comparator in that trial further suggest the need for further data collection in the framework of the Cancer Drugs Fund (see Section 0).

B.2.9.3.2.1 Sensitivity and scenario analyses

Sensitivity analyses were performed around uncertain input values. Among the ITCs of interest to this appraisal, a sensitivity analysis was performed only for the MAIC vs R^2. This was centred around alternative values for the proportion of patients with low Hb, as the relevant value was not available from the AUGMENT trial.

In addition, a number of scenario analyses were performed to test the robustness of the ITC results. The results were generally consistent between base-case and scenario analyses, supporting the base-case conclusions presented in this submission. Please see the ITC report (Appendix E) for details of the scenario analyses performed.

© Roche (2022). All rights reserved

B.2.10 Adverse reactions

Safety data from the GO29781 study are presented below primarily for the pivotal cohort (n=90, the same population as the efficacy evaluation presented in Sections B.2.6–B.2.7). An overview of the safety profile in patients with other NHL types who received the planned label dose of mosunetuzumab is also provided for completeness. This data is reflective of the general safety profile of mosunetuzumab emerging from clinical trials and, in the absence of additional real-world data at this stage, no Appendix focused on adverse events was provided with this submission.

B.2.10.1 Adverse reactions in the pivotal cohort

B.2.10.1.1 Overview of AEs

In the pivotal cohort, the median number of mosunetuzumab cycles received was 8.0 (range: 1–17). All 90 patients experienced at least one AE, with 83 patients (92.2%) experiencing AEs considered related to mosunetuzumab treatment by the investigator. SAEs were reported in 42 (46.7%) of patients and 30 (33.3%) experienced SAEs that were considered related to treatment by the investigator. AEs occurring in ≥10% of patients are summarized in Table 17. The most common AEs included CRS ([redacted] per Lee 2014 criteria[54] and 44.4% per ASTCT 2019[55] criteria), fatigue (36.7%), and headache (31.1%). SAEs that occurred in ≥3 patients included CRS ([redacted] by Lee 2014[54] and ASTCT 2019[55]), and acute kidney injury and urinary tract infection patients ([redacted]).

Table 17. Common (occurring in ≥10% of patients) AEs in the GO29781 pivotal cohort (N=90)

Patients with at least one AE	N (%)
CRS (by Lee 2014 grade)	[redacted]
CRS (by ASTCT 2019 grade)	40 (44.4)
Fatigue	33 (36.7)
Neutropenia/neutrophil count decreased	26 (28.9)
Headache	28 (31.1)
Hypophosphataemia	24 (26.7)
Pyrexia	26 (28.9)
Hypokalaemia	17 (18.9)
Cough	16 (17.8)
Pruritus	19 (21.1)
Rash	14 (15.6)
Upper respiratory tract infection	[redacted]
Diarrhoea	15 (16.7)

© Roche (2022). All rights reserved

Nausea	15 (16.7)
Constipation	16 (17.8)
Anaemia/haemoglobin decreased	[redacted]
Insomnia	[redacted]
Hypomagnesaemia	[redacted]
Oedema peripheral	[redacted]
Dry skin	14 (15.6)
Chills	[redacted]
Dizziness	[redacted]
ALT increased	[redacted]
Back pain	[redacted]
Urinary tract infection	[redacted]
Abdominal pain	[redacted]
Arthralgia	[redacted]
Skin exfoliation	[redacted]
Abbreviations: AE, adverse event; ALT, alanine transaminase; CRS, cytokine release syndrome	

B.2.10.1.2 Grade 3–5 events

A total of 64 (70.0%) of patients experienced grade 3–4 adverse events. Most common grade ≥3 AEs (occurring in ≥5% of patients) were neutropenia/neutrophil count decreased (24 patients [26.7%]), [redacted] There was one fatal AE in a patient who received 2 cycles of mosunetuzumab (last treatment on study Day 22) and was found unresponsive in bed on study Day 60. The event was assessed by the investigator as unrelated to mosunetuzumab and the cause of death was unknown.

B.2.10.1.3 Relationship of AEs to mosunetuzumab treatment and treatment modifications due to AE

A total of 83 (92.2%) of patients experienced an AE that the investigator considered related to mosunetuzumab and 30 (33.3%) of patients experienced a mosunetuzumab-related SAE. The most frequently reported mosunetuzumab-related AEs (occurring in ≥10% of patients) were [redacted]

© Roche (2022). All rights reserved

[redacted]

Four patients (4.4%) withdrew from treatment due to an AE (CRS in two cases, and Epstein-Barr viraemia and Hodgkin's disease in 1 case each) and in 2 of those patients (2.2%), the relevant events were mosunetuzumab-related. Mosunetuzumab dose modification or dose interruptions due to AEs occurred in 34 (37.8%) patients.

B.2.10.1.4 Adverse events of special interest

Adverse events of special interest (AESI) were defined based on the evolving clinical experience with mosunetuzumab in clinical studies and included cytokine release syndrome (CRS), haemophagocytic lymphohistiocytosis, neurologic AEs, haematologic AEs, tumour lysis syndrome, tumour flare, hepatic AEs, infections, and pneumonitis or interstitial lung disease. Two key AESI types emerging in the pivotal cohort were CRS and haematologic events:

- CRS
    - Two sets of classification criteria were used to grade CRS events – Lee 2014[54] and the 2019 American Society for Transplantation and Cellular Therapy (ASTCT)[55] criteria.
    - According to both sets of grading criteria, the majority of patients who experienced CRS had grade 1–2 events.
    - The most common signs and symptoms associated with CRS are listed in Table 18.
    - [redacted].
      According to ASTCT 2019 criteria[55], 40 patients (44.4%) experienced any grade CRS and 2 patients (2.2%) experienced grade 3–4 CRS.
    - Two patients had mosunetuzumab treatment withdrawn due to CRS.
    - CRS led to dose interruption or modification of mosunetuzumab in 8 patients (8.9%) by Lee 2014 criteria[54] and in 7 patients (7.8%) by ASTCT 2019 criteria[55].
    - The median duration of a CRS event was 3 days (range: 1–29 days). All CRS events resolved by the time of the data cut-off.
    - The incidence of CRS was highest in Cycle 1, particularly following mosunetuzumab administration on Day 15 when 32 out of 88 dosed patients (36.4%) developed CRS of any grade. The proportion of patients experiencing CRS of any grade decreased in the subsequent treatment cycles, with 9 out of

© Roche (2022). All rights reserved

87 patients (10.3%) experiencing any grade CRS in Cycle 2 and just 2 patients of 83 (2.4%) having any grade CRS in Cycles 3+.

- Haematologic events
  - [redacted]
  - [redacted]
    - Haematologic events in the pivotal cohort are summarised in Table 19.

Table 18. Common signs and symptoms of CRS (occurring in ≥10% of patients with CRS) in the GO29781 pivotal cohort

	CRS per Lee 2014[54] (N=41)	CRS per ASTCT 2019[55] (N=40)
Patients with at least one AE, n (%)*		
Pyrexia	[redacted]	[redacted]
Chills	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Tachycardia	[redacted]	[redacted]
Hypoxia	[redacted]	[redacted]
Rash	[redacted]	[redacted]

Abbreviations: AE, adverse event; ASTCT, American Society for Transplantation and Cellular Therapy; CRS, cytokine release syndrome

* Multiple occurrences of signs and symptoms in one individual were counted once at the highest grade based on NCI CTCAE v4.0, excluding signs and symptoms of missing grade.

** All patients who had CRS events by ASTCT 2019 grading criteria[55] experienced pyrexia. One patient had body temperature increased Preferred Term reported; therefore, this event was not included in the pyrexia count.

Table 19. Haematologic AEs in the GO29781 pivotal cohort (N=90)

Patients with at least one haematologic AE	**N (%)**
Neutropenia/neutrophil count decreased (any grade)	[redacted]
Grade 1–2 max. severity	[redacted]
Grade 3–4 max. severity	[redacted]
Febrile neutropenia (any grade)	[redacted]
Grade 1–2 max. severity	[redacted]

© Roche (2022). All rights reserved

Grade 3–4 max. severity	[redacted]
Thrombocytopenia/platelet count decreased (any grade)	[redacted]
Grade 1–2 max. severity	[redacted]
Grade 3–4 max. severity	[redacted]
Anaemia/haemoglobin decreased (any grade)	[redacted]
Grade 1–2 max. severity	[redacted]
Grade 3–4 max. severity	[redacted]
Abbreviations: AE, adverse event; max, maximal	

B.2.10.2 Overview of safety in the broader NHL cohort

The safety profile of mosunetuzumab in patients with R/R FL included in the pivotal cohort was generally consistent with that observed in the broader cohort of patients with NHL receiving the planned label dose of mosunetuzumab (Table 20). This cohort included patients with other NHL subtypes in addition to FL who received cycle 1 step-up mosunetuzumab IV monotherapy at the dose of 1/2/60/30 mg during the dose expansion phase; see footnote to Table 20 for details.

Most common AEs in the broader NHL cohort included CRS (occurring in [redacted] of patients, see Table 20), fatigue ([redacted]), pyrexia ([redacted] neutropenia ([redacted] and headache ([redacted] Excluding fatal disease progression, SAEs occurred in [redacted]_Grade ≥3 events were reported in [redacted] and, excluding deaths from PD, [redacted]) died due to grade 5 AEs including sepsis, cholangitis, pneumonia, and one death from unknown cause (in the patient with R/R FL, see Section B.2.10.1.2 for details). The death from sepsis was considered related to treatment with mosunetuzumab.

© Roche (2022). All rights reserved

Table 20. Comparative overview of mosunetuzumab safety at the planned label dose in patients with R/R FL and the broader cohort of NHL patients

	Total NHL cohort at planned label dose (N=218)[a]	R/R FL pivotal cohort (N=90)
Total no. of patients with at least one AE	[redacted]	90 (100%)
Total no. of events	[redacted]	[redacted]
Total no. of deaths [b]	[redacted]	[redacted]
Total no. of patients withdrawn from treatment due to AE or death	[redacted]	4 (4.4%)
No. of patients with at least one:		
Fatal AE	[redacted]	2 (2.2%)
Fatal AE (not including PD)	[redacted]	[redacted]
SAE	[redacted]	42 (46.7%)
SAE (excluding Grade 5 PD)	[redacted]	[redacted]
Mosunetuzumab-related SAE	[redacted]	30 (33.3%)
Mosunetuzumab-related AE	[redacted]	83 (92.2%)
AE of Grade 3-4 [c]	[redacted]	63 (70.0%)
AE leading to withdrawal from mosunetuzumab treatment	[redacted]	4 (4.4%)
Mosunetuzumab-related AE leading to withdrawal from mosunetuzumab treatment	[redacted]	2 (2.2%)
AE leading to mosunetuzumab dose modification	[redacted]	[redacted]
AE leading to mosunetuzumab dose interruption	[redacted]	[redacted]
No. of patients with at least one CRS AE:		
Lee 2014[54] grade (any grade)	[redacted]	[redacted]
Grade 1–2 max. severity by Lee 2014 grade	[redacted]	[redacted]
Grade 3–4 max. severity by Lee 2014 grade	[redacted]	[redacted]
ASTCT 2019[55] grade (any grade)	[redacted]	40 (44.4%)
Grade 1–2 max. severity by ASTCT 2019 grade	[redacted]	38 (42.2%)
Grade 3–4 max. severity CRS by ASTCT 2019 grade	[redacted]	2 (2.2%)

Abbreviations: AE, adverse event; ASTCT, American Society for Transplantation and Cellular Therapy; CRS, cytokine release syndrome; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; NHL, non-Hodgkin lymphoma; MCL, mantle cell lymphoma; PD, progressive disease; R/R relapsed/refractory; SAE, serious AE; tFL, transformed FL

[a] Includes patients with R/R FL (n=90), DLBCL/tFL (n=88), MCL (n=25), and Richter's transformation (n=14). One patient with melanoma was enrolled in error into the DLBCL/tFL cohort and received one dose of mosun.

[b] All deaths from the start of treatment up to the data cut-off were included.

[c] Includes all patients who experienced Grade 3–4 AEs during initial treatment. It should be noted that some of these patients could have also experienced a Grade 5 event as their worst grade event.

B.2.11 Ongoing studies

The only study expected to provide data on mosunetuzumab monotherapy in R/R FL within the next 12 months is the GO29781 study described in this submission. Data from the

© Roche (2022). All rights reserved

[redacted] data cut will be available towards the end of [redacted]; updated analysis from which will be available during the appraisal process. A further data cut is planned for [redacted], with the final follow up planned to continue until [redacted].

B.2.12 Interpretation of clinical effectiveness and safety evidence

Despite recent therapeutic advances, FL remains an incurable disease for which standard of care varies substantially across the UK. Patients with R/R FL who are at third or later line of therapy are at risk of particularly poor outcomes. These patients are usually heavily pretreated, having received both anti-CD20 antibodies and chemotherapy regimens, and their disease often becomes increasingly refractory to these agents limiting their utility[8].

New treatment options that can overcome resistance to existing therapies while providing acceptable safety and tolerability are needed. As a humanised anti-CD20/CD3 bispecific antibody, mosunetuzumab offers a unique, first-in-class mechanism of action that differs from the mechanisms of currently available therapies. Therefore, mosunetuzumab can provide a novel treatment option for heavily pretreated R/R FL patients, particularly those patients who are refractory to anti-CD20 antibodies and various chemotherapy regimens.

B.2.12.1 Summary of clinical effectiveness and safety evidence

Efficacy and safety data supporting the use of mosunetuzumab in patients with R/R FL who had received at least 2 prior lines of systemic treatment was derived solely from the pivotal cohort of the phase I/II, multicentre, single-arm, open-label, dose-escalation and dose-expansion GO29781 trial. In this pivotal cohort, the primary efficacy endpoint of CR rate as assessed by IRF reached 60.0% (95% CI: 49.1, 70.2). CR rates were consistent across relevant subgroups defined by demographics, baseline disease characteristics, prior treatment, and prognostic factors, including high-risk patient populations with double refractory disease and POD24 following initial treatment. The ORR assessed by IRF was 80.0% (95% CI: 70.3, 87.7) in the overall pivotal cohort and displayed consistency across the aforementioned subgroups.

Responses to mosunetuzumab proved durable and extended well beyond the treatment duration of eight 3-weekly cycles. The median duration of CR was not reached, and an estimated 71.4% (95% CI: 57.9, 84.9) of CR patients continued to benefit from remission after 1 year. Among all responders, 61.8% (95% CI: 50.0, 73.7) were estimated to remain in remission at 1 year. One-year PFS and OS rates were estimated at 57.7% (95% CI: 46.9, 68.4) and [redacted], respectively. Median PFS was 17.9 months (95% CI: 10.1, NE), while [redacted].

© Roche (2022). All rights reserved

The EORTC QLQ-C30 questionnaire (physical functioning and fatigue scores), FACT-Lym subscale, and EQ-5D-5L questionnaire were used to assess PROs. High compliance rates [redacted] were consistently observed across the scheduled assessments. During treatment with mosunetuzumab, baseline HRQoL and health status were generally maintained.

The safety results demonstrated an acceptable tolerability and manageable safety profile of mosunetuzumab IV, both in the pivotal R/R FL cohort and in the broader cohort of patients with different NHL types who received the planned label dose of mosunetuzumab. Only 4.4% of patients in the pivotal cohort and 4.1% in the broader NHL cohort discontinued treatment due to an AE. Although CRS was relatively common, affecting 45.6% of patients in the pivotal cohort and 42.7% in the broader NHL cohort, the vast majority of events were mild or moderate (grade 1–2). With a growing experience of mosunetuzumab use among clinicians, the pre-emptive management and treatment of CRS is likely to improve, reducing the incidence of CRS relative to what was observed in this early phase trial. Reflecting the acceptable safety profile of mosunetuzumab, no hospitalisation is mandated for mosunetuzumab administration in the draft SmPC (see Appendix C).

Since the GO29781 trial does not provide comparative evidence in support of mosunetuzumab, a series of ITCs was conducted against comparators specified in the NICE scope, i.e., R^2 and rituximab plus chemotherapy (here, an ITC was feasible vs rituximab plus bendamustine, but not vs R-CHOP). An ITC vs obinutuzumab plus bendamustine was also conducted as this was listed in the NICE scope, although it should be noted that based on market share data and clinical expert advice, the company do not consider this regimen to be a relevant comparator for patients with R/R FL who had received at least 2 prior lines of systemic treatment. The analyses accounted for a set of clinically important covariates and compared CR rate, ORR, PFS, OS, and discontinuation due to AEs. In terms of methodology, MAICs were employed where only published aggregate data was available, while propensity score analyses were used to compare mosunetuzumab with those treatments for which the company had access to the IPD. [redacted]

© Roche (2022). All rights reserved

B.2.12.2 Strengths and limitations of the clinical evidence base

The pivotal cohort of the GO29781 trial provided robust data supporting mosunetuzumab use in a patient population with high unmet need and limited treatment options. In this heavily pretreated population, the median number of prior therapies was 3 (maximum of 10) and 21.1% of patients received prior ASCT. Patients were frequently refractory to commonly used therapies, including anti-CD20 antibodies (78.9%) and both an anti-CD20 antibody and an alkylating agent (53.3%). Therefore, the patients included in the pivotal GO29781 cohort would be unlikely to derive benefit from currently available therapies.

Risk factors for poor prognosis were abundant in the pivotal GO29781 trial cohort, with 44.4% of patients having a high FLIPI score, 52.2% experiencing POD24 following initial therapy, and 34.4% having bulky disease. Despite these risk factors for poor prognosis, median PFS was estimated at 17.9 months (95% CI: 10.1, NE) and the proportion of patients remaining progression-free at 18 months was estimated at 47.0% (95% CI: 34.4, 59.6). These results strongly suggest that mosunetuzumab can delay disease progression in a population of difficult-to-treat patients with poor prognosis, while maintaining HRQoL and displaying an acceptable tolerability and safety profile.

Main limitations of the GO29781 trial relate to its early-phase, non-randomised, single-arm design and the open-label nature of treatment, all of which could introduce potential bias, the direction of which is difficult to assess. In the absence of head-to-head data, ITCs were utilised to compare mosunetuzumab with the comparators listed in the NICE scope. These analyses used robust methodology; however, the effective sample sizes after adjustment were relatively small. Together with the low number of events observed for some endpoints, this may have contributed to the substantial uncertainty observed for many endpoints, limiting the interpretation of the results. Furthermore, the data for mosunetuzumab is immature relative to comparator data, and the conclusions from the ITCs may change as longer follow-up becomes available and more events are observed for the endpoints of interest.

Despite the aforementioned limitations, the availability of mosunetuzumab on the NHS can provide a novel, non-rituximab-based treatment option to heavily pretreated patients with an incurable malignancy, who would otherwise have limited therapeutic choices. The benefits that this novel treatment option can have in a patient population with high unmet need should be balanced against the clear limitations of the available evidence base. Collection of real-world data on mosunetuzumab use within a framework of a Cancer Drugs Fund managed access agreement would enable patients to start benefiting from treatment with mosunetuzumab without delay, and at the same time provide additional data that could

© Roche (2022). All rights reserved

address the uncertainties associated with the limited and short-term data that are available at present.

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

In line with the NICE health technology evaluations: the manual (2022)[56], an SLR was conducted to identify cost-effectiveness studies on the management of patients with FL. In brief, electronic database searches (Embase, MEDLINE< EconLit and Evidence Based Medicine [EBM] Reviews) were conducted in January 2022. Supplementary sources were also hand searched for completeness, including reference lists of included studies, conference proceedings, relevant additional databases and websites, and global health technology assessment (HTA) body websites. In total, 32 publications were identified (Figure 20), reporting 19 published analyses (Table 21) and 13 HTAs (Table 22). Details of the SLR can be found in the report provided as Appendix F.

Of the studies identified in the SLR, the majority (10 of the 19 original studies) described Markov models with three health states: pre-progression, progressed disease, and death. The other analysis types included partitioned survival models (PSMs), an unspecified Excel-based cohort model reconstructed in TreeAge, and a cost-minimisation model. The SLR also identified 12 HTAs, six of which were informed by Markov models and three by PSMs.

© Roche (2022). All rights reserved

Table 21. Summary list of published cost-effectiveness studies

Study name	Model structure	Intervention (s)	Patient population
Bertwistle, 2013a[57]	Partitioned survival (assumed)	• Bendamustine- rituximab • Fludarabine- rituximab	Patients with iNHL who have progressed following treatment with rituximab or a rituximab-containing regimen
Bertwistle, 2013b[58]	Partitioned survival (assumed)	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with iNHL who have progressed following treatment with rituximab or a rituximab-containing regimen
Desanvicente-Celis, 2014a[59]	Markov model	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with iNHL who have progressed following treatment with rituximab or a rituximab-containing regimen
Desanvicente-Celis, 2014b[60]	Markov model	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with iNHL who have progressed following treatment with rituximab or a rituximab-containing regimen
Erdogan-Ciftci, 2019[61]	Partitioned survival	• Obinutuzumab-bendamustine • Bendamustine monotherapy	Patients with FL who did not respond or who progressed during or up to 6 months after rituximab or a rituximab-containing regimen
Guzauskas, 2018[62]	Partitioned survival	• Obinutuzumab-bendamustine followed by obinutuzumab monotherapy • Bendamustine monotherapy	Patients with FL who relapsed after, or are refractory to, a rituximab-containing regimen
Haukaas, 2018[63]	Partitioned survival	• Obinutuzumab-bendamustine followed by obinutuzumab monotherapy • Bendamustine monotherapy	Patients with FL who relapsed after, or are refractory to, a rituximab-containing regimen
Lachaine, 2013a[64]	Markov model	• Bendamustine • Ibritumomab tiuxetan	Patients with rituximab-refractory iNHL
Lachaine, 2013b[65]	Markov model	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with relapsed iNHL and MCL

© Roche (2022). All rights reserved

Salazar, 2017[66]	Markov model	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with iNHL that has progressed during or within six months of treatment with rituximab or a regimen containing rituximab
Soini, 2011[67]	Markov model	• R-CHOP • R-CHOP-R • CHOP	Patients with relapsed/refractory FL
Soini, 2012[68]	Markov model	• Sequence 1: R-CHOP-R → R-CVP-R/B → BSC • Sequence 2: R-CHOP-R → R-CVP-R/CVP → BSC • Sequence 3: R-CHOP → R-CVP-R/B → BSC • Sequence 4: R-CHOP → R-CVP-R/CVP → BSC	Patients with grade I-IIIa FL; high tumour burden and complete/partial response to 1L chemotherapy (R-CHOP) induction
Sweetenham, 1999[69]	NA	• Fludarabine • CHOP • Rituximab	Patients with relapsed indolent B-cell NHL
Thielen, 2021[70]	Partitioned survival	• Lenalidomide-rituximab • Rituximab alone	Patients with previously treated FL
Thompson, 2005[71]	NA (trial-based)	• Ibritumomab tiuxetan • Standard rituximab (4-dose rituximab) • Standard rituximab followed by maintenance therapy (8-dose rituximab)	Patients with relapsed FL
Vandekerckhove, 2012[72]	Markov model	• Bortezomib-rituximab • Rituximab alone	Patients with biomarker-positive relapsed/refractory FL
Vijenthira, 2021[73]	Markov model	• Allogeneic SCT • Autologous SCT • Chemo-immunotherapy alone (obinutuzumab-CHOP)	Patients with early relapse of FL who are eligible for transplant; progression within 24 months of initial treatment with a rituximab-containing regimen

© Roche (2022). All rights reserved

Wehler, 2013[74]	Excel-based cohort model reconstructed in TreeAge	• Bendamustine-rituximab • Fludarabine-rituximab	Patients with relapsed iNHL
Zhang, 2020[75]	Markov model	• Lenalidomide-rituximab • Rituximab alone	Patients with relapsed or refractory indolent lymphoma
Abbreviations: BSC: best supportive care, CHOP: cyclophosphamide; doxorubicin, vincristine, and prednisolone; FL: follicular lymphoma; iNHL: indolent non-Hodgkin lymphoma, MCL: mantle cell lymphoma, R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisolone; R-CVP: rituximab, cyclophosphamide, vincristine, and prednisolone; SCT: stem cell transplant			

Table 22. Summary list of HTA submissions

Study name	Model structure	Intervention (s)	Patient population
TA629[25] England/Wales	Partitioned survival	• Obinutuzumab-bendamustine followed by obinutuzumab maintenance • Bendamustine alone	Patients with FL that did not respond or progressed up to 6 months after treatment with rituximab or a rituximab-containing regimen
TA627[24] England/Wales	Partitioned survival	• Lenalidomide-rituximab • Rituximab-chemotherapy (non-rituximab-refractory) • Obinutuzumab-bendamustine (rituximab-refractory)	Adult patients with previously treated FL or MCL
TA604[27] England/Wales	Markov model	• Idelalisib monotherapy • Chemotherapy regimens (such as cyclophosphamide- or fludarabine-containing regimens, bendamustine or chlorambucil • BSC (in patients for whom chemotherapy is unsuitable)	Patients with FL that are refractory to two prior lines of therapy
TA137[26] England/Wales	Markov model	• R-CHOP • R-CHOP-R • CHOP followed by observation	Patients with relapsed or refractory stage III or IV FL

© Roche (2022). All rights reserved

SMC2281[76] Scotland	Partitioned survival	• Lenalidomide-rituximab • Rituximab-chemotherapy (CHOP, CVP, and bendamustine) (non-rituximab-refractory) • Obinutuzumab-bendamustine (rituximab-refractory)	Adult patients with previously treated grade I-IIIa FL
SMC 1219/17[77] Scotland	Markov model	• Obinutuzumab-bendamustine followed by obinutuzumab maintenance • R-CHOP • Bendamustine alone	Patients with FL who did not respond or who progressed during or up to six months after treatment with rituximab or a rituximab-containing regimen
SMC 1039/15[78] Scotland	Markov model	• Idelalisib monotherapy • SOC (re-treatment with a range of chemotherapy and/or rituximab regimens) • BSC (scenario analysis)	Adult patients with FL that is refractory to two prior lines of treatment
SMC 171/05[79] Scotland	Markov model	• Ibritumomab tiuxetan • SOC (chemotherapy, radiotherapy, and SCT)	Adult patients with rituximab relapsed or refractory CD20+ follicular B-cell NHL
PSD March 2015[80] Australia	Single patient expected value analysis	• Idelalisib monotherapy • BSC	Adult patients with relapsed or refractory iNHL who had received at least two prior therapies
PSD November 2015[81] Australia	Single patient expected value analysis	• Idelalisib monotherapy • BSC	Patients with FL that is refractory to both rituximab and an alkylating agent
PSD March 2016[82] Australia	NR	• Idelalisib monotherapy • BSC	Patients with FL that is refractory to both rituximab and an alkylating agent
Zydelig September 2016[83] Canada	Partitioned survival	• Idelalisib monotherapy • BSC	Patients with FL who have received at least two prior systemic therapies and are refractory to both rituximab and an alkylating agent

© Roche (2022). All rights reserved

Gazyva June 2017[84] Canada	Markov model	• Obinutuzumab-bendamustine followed by obinutuzumab maintenance • Bendamustine alone	Patients with FL who have relapsed after or are refractory to a rituximab-containing regimen
Abbreviations: BSC: best supportive care; CHOP: cyclophosphamide, doxorubicin, vincristine, and prednisolone; CVP: cyclophosphamide, vincristine, and prednisolone; FL: follicular lymphoma; (i)NHL: (indolent) non-Hodgkin lymphoma, MCL: mantle cell lymphoma, R-CHOP(-R): rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisolone (followed by rituximab maintenance); PSD: Public Summary Document; SCT: stem cell transplant; SOC: standard of care			

© Roche (2022). All rights reserved

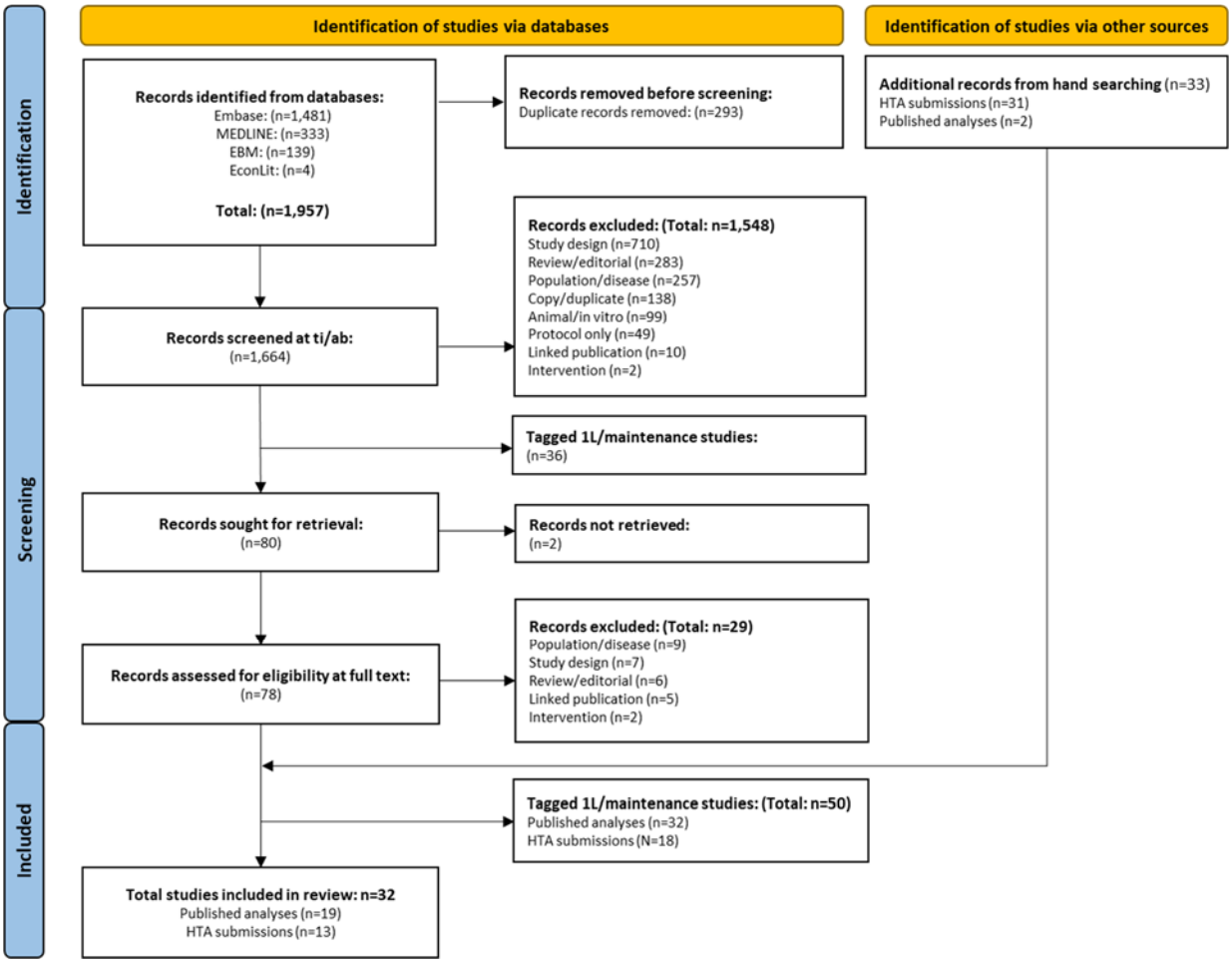


Figure 20. PRISMA flow diagram for SLR of economic evaluations

© Roche (2022). All rights reserved

B.3.2 Economic analysis

The economic case presented in this submission is based on a cost-utility analysis assessing the use of mosunetuzumab versus various active comparators (detailed in B.3.2.5) for the treatment of adult patients with R/R FL who have received at least two prior systemic therapies (hereafter referred to as third or subsequent line [3L+]). The analysis takes into account a patient access scheme (PAS) discount for mosunetuzumab (detailed in Section B.3.5.2).

The cost-effectiveness studies identified in Section B.3.1 were examined to inform the economic analysis presented in this submission. Previously published modelling approaches were mostly PSMs or Markov models with the majority of models adhering to the common oncology three-state framework (pre-progression, progressed disease, and death), regardless of modelling type, as this represents the most important clinical outcomes for patients.

PSMs are commonly used in oncology, as detailed in NICE TSD 19[85], and lend themselves to situations where transitions between all states cannot be explicitly identified and modelled, for example, where post-progression survival cannot be estimated from reported data as only PFS and OS are reported, and comparator data may not be available. It has been demonstrated that there is little difference in estimated outcomes between partitioned survival and Markov models and that the assumptions underpinning analysis are more relevant than the choice of the modelling approach[86,87]. The largest consideration is whether time to progression or death is expected to be inherently different between arms and whether the model is able to capture these endpoints appropriately[86,87]. PSMs can reflect these relevant clinical endpoints well and is appropriate where data is not available to inform alternative approaches that require more granularity[86,87]. A PSM can therefore capture long-term impact of oncology interventions in terms of both PFS and OS, which were key secondary outcomes in the GO29781 trial (see Section B.2.3.5). However, the trial's primary endpoint, response rate, is not adaptable to use in PSM.

Importantly, PSMs do not require any PFS to OS surrogacy assumptions and do not translate any PFS benefit into an OS benefit. Therefore, PFS and OS data, being taken directly from the GO2978 trial, can be independently reflected in a PSM. PFS surrogacy for OS is poor for some cancer types[88] and may be less than optimal for first-line treatment of FL[89]. Since no assessments of PFS surrogacy for OS could be identified for the 3L+ FL setting, the choice of a PSM would mean no assumptions surrounding this concept were required.

Taking into account the above considerations a *de novo* three-state PSM was built to inform decision making. This modelling approach is in line with previous TAs in the same indication[24,25] and literature identified in the related SLR.

© Roche (2022). All rights reserved

B.3.2.1 Clinical evidence used in the model

In the model, data from the GO29781 study (Sections B.2.6 and B.2.10) have been used to inform the clinical efficacy, safety and time on treatment of mosunetuzumab for the treatment of adult patients with R/R FL who have received ≥2 prior systemic therapy lines. The GO29781 study is currently the only study available to provide clinical evidence for mosunetuzumab in the intended population and can therefore be considered the best available evidence to inform the modelling. All analyses in this submission have been conducted from a National Health Service (NHS)/ Personal Social Services (PSS) perspective.

While GO29781 is the source of mosunetuzumab data for the cost-effectiveness analysis, it is a single-arm trial therefore no comparator data are available. Consequently, an ITC was required to provide comparative evidence against the potential comparators identified in the scope of this appraisal. The ITC employed a propensity score analysis for those comparators with available patient-level data, and a MAIC where only published aggregate data were available (Section B.2.9).

B.3.2.2 Patient population

Mosunetuzumab is proposed for use within the NHS in England as an alternative to any third- or later-line therapy option and irrespective of transplantation status (i.e., as a bridge to ASCT, in patients relapsing post-ASCT, and in those unsuitable for ASCT).

The population subject to the *de novo* analysis aligns with the population of the pivotal GO29781 cohort (see Section B.2.3.2) and therefore consists of adults with grade 1-3a FL that had relapsed after or failed to respond to ≥2 prior lines of systemic therapy, had an ECOG PS ≤1 and had received prior treatment with an anti-CD20-directed therapy and an alkylating agent

In the base case analysis, baseline patient parameters were derived from the baseline characteristics of pivotal cohort of patients enrolled in GO29781, as detailed in Table 23.

© Roche (2022). All rights reserved

Table 23. Baseline parameters in base case

Parameter	Mean	SE	Source
Age (years)	60.01	1.26	GO29781 Trial
Baseline body weight (kg)	81.40	1.98	GO29781 Trial
Baseline height (cm)	169.92	1.14	GO29781 Trial
Baseline BSA (m^2)	1.96	0.03	GO29781 Trial
Proportion of cohort male	61.10%	0.052	GO29781 Trial
BSA, body surface area; SE: standard error			

B.3.2.3 Model structure

A de novo partitioned survival (area under the curve [AUC]) model structure was developed representing PFS, PD, and death. These health states reflect the disease severity and clinical landmarks, as well as key distinctions in mortality, HRQoL, and the use of healthcare resources.

The economic modelling of mosunetuzumab and the relevant comparators in this indication required that comparative efficacy be pieced together from numerous sources with ITCs. Within the AUC model, health state occupancy was determined by partitioning the proportion of patients alive into PFS and PD at discrete time points based on the OS and PFS curves from the GO29781 study and the relative treatment effects derived from the ITC. The model structure is shown in Figure 21.

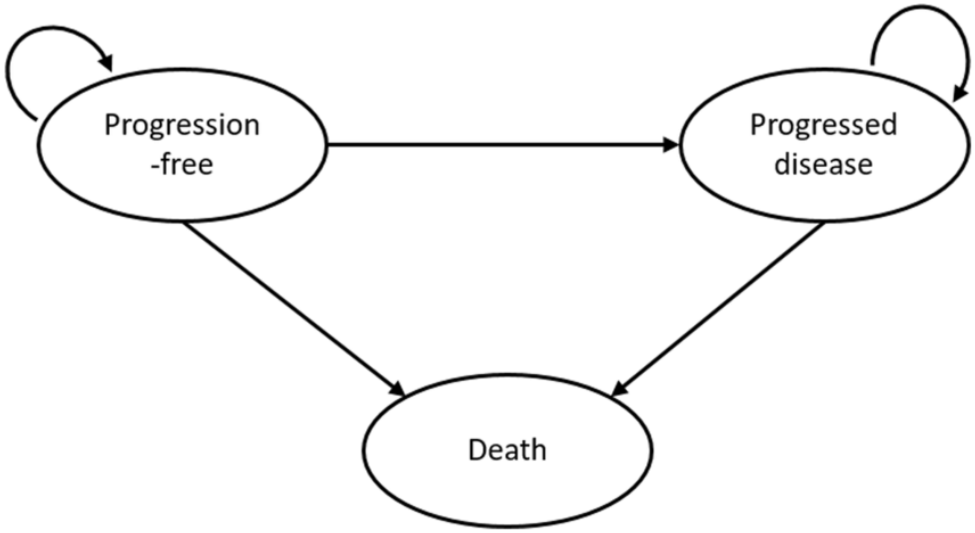


Figure 21. Model Schematic

All patients entered the model in the PFS health state and remained in this health state until their disease progressed, or they died. Once in the progressed health state, patients could either remain in the progressed health state or move to the death state. Patients in the model could not transition to an improved health state, i.e., from PD to PFS.

© Roche (2022). All rights reserved

The economic model uses a 40-year time horizon, which was expected to be sufficiently long to capture all important differences in costs or clinical outcomes between the technologies being compared as all patients in the model were expected to be in the death state by the end of 40 years. As such, the 40-year time horizon can be essentially considered equivalent of a lifetime horizon.

The model uses weekly cycles with the proportion of patients in each health state calculated after each cycle. A cycle duration of 1 week was considered appropriate for this evaluation because it enables the model to reflect differing timings of drug administrations between arms and the time scale over which patients may experience changes in their symptoms. In addition, transitions between health states can occur at any time within the cycle. In order to account for the over- or underestimation of transitions occurring at the beginning or end of the cycle, half-cycle correction was applied, in line with previous NICE technology appraisals in this disease area[24,27].

In line with the NICE Technology Evaluations Manual[56], model results are reported in terms of costs, quality-adjusted life-years (QALYs) gained, life-years (LYs) gained, net-health benefit (NHB), net-monetary benefit (NMB), and incremental cost-effectiveness ratios (ICERs).

Costs and health-related utilities were allocated by health state to calculate the weighted cost and QALYs per cycle. Cost and health outcomes were discounted at a 3.5% discount rate and, according to the NICE reference case, an NHS and PSS perspective was assumed.

B.3.2.3.1 Derivation of health state occupancy estimates

The decrease in the proportion of patients residing in the progression-free state over time (starting from 100%) was determined by parametric models fit to the PFS curves from the GO29781 data and ITC analysis. The PFS curves indicate, for each time point, the proportion of patients who have not progressed or died.

PD accommodates all patients who have experienced disease progression but have not yet died. The proportion of patients in this state was calculated as the difference between the proportion of living patients and the proportion of patients who were both living and pre-progression. The transitions into and out from the progression health state were thus not modelled explicitly, a defining feature of PSMs.

Death was modelled as an absorbing state meaning that all patients eventually enter this state and cannot leave it. The transition rate of patients from the progression-free and progressed disease heath states into the death state was determined by parametric models fit to the OS curves derived from the GO29781 trial and the relevant comparator data. OS curves indicate

© Roche (2022). All rights reserved

the proportion of patients who are alive at a given time point or, equivalently, the proportion of patients who die during a model cycle dependent on the time since treatment initiation. Estimates of OS and PFS were made by fitting independent parametric curves using the weights obtained from the ITC. In addition, this application prevents proportions of the cohort remaining alive indefinitely which is not a reasonable assumption in this indication. Clinicians supported this assumption and felt that the resulting OS curves reflected the survival that they would expect to see in clinics.

All cause-mortality (ACM) was included in the model to maintain external validity with the hazard applied such that survival never exceeds that of an age- and sex- matched population.

B.3.2.3.2 Derivation of treatment line occupancy

The time to off-treatment (TTOT) KM data from the GO29781 trial were used directly in the model to estimate treatment discontinuation with mosunetuzumab. KM data from the trial was followed for 12 months but it was restricted in the model such that it did not exceed PFS at any time.

For all other treatments, the TTOT was set equal to the selected parametric distribution for PFS if treatment until progression was selected, or capped at the treatment-specific maximum number of cycles if relevant.

While patients remained progression free, they could be on or off treatment. Once in the PD health state, it was assumed that patients would move to a further line of treatment. Subsequent therapy use was informed by clinical advice. Further, it was assumed that the subsequent therapies comprise all possible therapies that patients may receive either sequentially or concurrently. The cost of subsequent therapy was applied as a one-off cost on progression, taking into consideration the mean treatment duration and the proportion likely to take each available therapy. This is detailed further in Section B.3.5.4.

In previous TAs where immunotherapies have had stopping rules[24,90,91], treatment effect waning has been applied; this is a common approach in modelling immunotherapies. However, only 5.6% of patients were still receiving mosunetuzumab in the GO29781 trial in the cycle before the maximum of 17 cycles (approximately one year) was reached, with 67.7% of patients receiving mosunetuzumab at that point still alive and progression-free. Further, the comparator treatments also have stopping rules that affect a far greater proportion of patients in the model (93.2% for rituximab plus bendamustine and 54.5% for obinutuzumab with bendamustine). As such, the application of treatment effect waning for mosunetuzumab, despite the presence of a maximum cycle number in the trial, was not considered necessary. As a far greater percentage of patients were affected by the comparator treatment stopping

© Roche (2022). All rights reserved

rules than were affected by the stopping rule for mosunetuzumab, inclusion of treatment waning in the model would have resulted in more favourable cost-effectiveness results for mosunetuzumab.

B.3.2.3.3 Outcome measures

The primary model output is the incremental cost-effectiveness ratio (ICER) expressed as incremental costs per quality-adjusted life-year (QALY) gained. The model provides an overview of other health economic outcomes such as total QALYs, costs, NMB, NHB, and life-years associated with each treatment in total and in a disaggregated form.

B.3.2.4 Comparison of the *de novo* analysis with previous appraisals

An SLR was undertaken to evaluate modelling approaches for R/R FL to identify relevant literature, including previous technology appraisals (TAs). Table 24 provides a comparison of the current submission versus several previous appraisals for FL.

The *de novo* analysis followed precedent from existing submissions as well as the NICE reference case. A lifetime horizon was used to capture all potential costs and benefits and efficacy and utility data were derived from the key trial or sourced from the literature when trial data were not suitable.

© Roche (2022). All rights reserved

Table 24. Previous relevant technology appraisals

	Previous appraisals				Current appraisal	
Factor	**TA629**	**TA627**	**TA604**	**TA137**	**Chosen values**	**Justification**
Time horizon	Lifetime (25 years)	Lifetime (40 years)	Lifetime (38 years)	Lifetime (30 years)	40 years	Lifetime horizon is appropriate when modelling a chronic disease such as FL.
Treatment waning effect?	Not mentioned	5-year treatment effect with a hard cut-off. Additional time points tested in scenario analysis as no evidence to suggest appropriate time.	The multifaceted nature of the approach to cost-effectiveness analyses means different survival assumptions and structural approaches to survival are considered, within the context of the clinical data limitations at hand. Parametric survival analysis of clinical endpoints is central to each approach	Not explicitly discussed	No treatment waning effect	Treatment effect waning due to a stopping rule would affect only a small percentage of patients receiving mosunetuzumab and much larger percentages of patients receiving comparator treatments. Its exclusion therefore generates conservative cost effectiveness results for mosunetuzumab.
Source of utilities	Published literature (Wild et al [2006][92])	AUGMENT trial	Published literature (Pettengell et al [2008][15]; linked to Wild et al [2006][92])	Oxford Outcomes Utility Study (Wild et al [2006][92])	PFS – 0.80 PPS – 0.75 GO29781 data	Trial data was used where possible, but data was immature for PPS and so estimates were sourced from published evidence.

© Roche (2022). All rights reserved

Source of costs	NR	MIMS, eMIT, NHS reference costs (2017/2018), PSSRU 2018, published literature and previous TAs	MIMS, eMIT, NHS reference costs (2016/2017), PSSRU 2017, published literature and previous TAs	NHS reference costs (2004), TOPS FU303	eMIT, BNF, NHS reference costs (2019/2020), PSSRU 2021.	Consistent with the NICE reference case
Abbreviations: BNF: British National Formulary; eMIT: drugs and pharmaceuticals electronic market information tool; FL: follicular lymphoma; MIMS: Monthly Index of Medical Specialties; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; NR: not reported; PFS: progression-free survival; PPS: post-progression state; PSSRU: Personal Social Services Research Unit; TAs: technology appraisals						

© Roche (2022). All rights reserved

B.3.2.5 Intervention technology and comparators

The health economic model was developed to compare the cost-effectiveness of mosunetuzumab versus:

- Lenalidomide with rituximab;
- Rituximab in combination with bendamustine, representing rituximab in combination with chemotherapy;
- Obinutuzumab with bendamustine followed by obinutuzumab maintenance.

These three comparators were considered to be the most relevant to the decision problem (Section B.1.1) based upon feedback from eight clinical experts at an Advisory Board where the consensus was that these treatments covered 90-100% of patients treated for FL in the 3L setting. Market research suggests that the majority of patients in a 3L setting would receive rituximab in combination with either chemotherapy or lenalidomide[93] and clinicians confirm this (Advisory Board data on file). The IPSOS Oncology Monitor 2022 report covering the period February 2021 to January 2022 indicates that rituximab in combination with lenalidomide is the most commonly used option (used by 33% people)[93]. The majority of treatments are rituximab containing regimens[93]. However, the treatment landscape is somewhat fragmented and the combination of bendamustine and obinutuzumab represents a small portion (■) of the landscape[93] where the other options may not be appropriate, with clinical experts also noting that this regimen is used infrequently in the 3L+ setting. As such, this comparator is included for completeness although the company does not consider obinutuzumab plus bendamustine to be a relevant comparator.

Best supportive care (BSC) was not included as a comparator as it is considered by clinicians to be a palliative approach. The age standardised survival rate over five years for FL in the 3-L setting in England is estimated to be 83.0%[1]. As such, palliative approaches were not considered to be of relevance to this submission, which was confirmed by clinical experts at an Advisory Board (Data on File).

It was not possible to make robust estimates of the comparative efficacy of R-CHOP and mosunetuzumab in the supporting ITC (see Section B.2.9). As such, rituximab in combination with bendamustine was presented as an alternative considered to represent rituximab with chemotherapy as closely as the data allowed.

© Roche (2022). All rights reserved

B.3.3 Clinical parameters and variables

B.3.3.1 Evidence synthesis

Evidence to describe the characteristics of the patient population and the effectiveness of mosunetuzumab was primarily derived from the GO29781 trial, a Phase I/II, multi-centre, open-label dose escalation and expansion single arm study. Comparator efficacy was informed by an SLR followed by an ITC, as described in Sections B.2.1 and B.2.9, respectively. Full details of the ITC are provided in Appendix E.

B.3.3.2 Survival analysis approach

All analyses were completed in R Software version 3.6.3 using flexsurv package, hazards were visualised and assessed with the muhaz package. Briefly, published Kaplan-Meier (KM) curves were digitised. By combing the scraped data with the number at risk it was possible to estimate the individual patient data for each comparator by using an algorithm proposed by Guyot et al, 2012[94].

The data used for all outcomes and arms is derived from the ITC; the adjusted mosunetuzumab KM and the comparator unadjusted KM data. Initially, proportional hazards were assessed for each set of reconstructed comparator data and the mosunetuzumab data to determine the suitability of the application of hazard ratios (HRs) and model choices. As described in Sections B.3.3.2.1 to B.3.3.2.3, in some cases it was not appropriate to accept the proportional hazard assumption. As such, independent parametric models were fit to each OS and PFS outcome for the respective comparator (unadjusted) and mosunetuzumab (adjusted). Fitting independent models is recommended, regardless of the proportional hazards assessment as, if proportional hazards are warranted, the independent models should reflect this regardless[95]. This was done for all comparators and outcomes (aside from TTOT) to ensure a consistent and conservative approach. However, as the assumption of proportional hazards did not need to be rejected for all arms and outcomes, and because the data contains limited events, the HRs generated by the ITC are included in the model to facilitate scenario analyses where this is considered a plausible scenario. A more robust assessment of the proportional hazards assumption may be possible as more data with longer follow up comes available. The results of these scenarios are shown in Section B.3.11.3.2.

Extrapolation beyond the clinical follow-up period for each treatment data was performed by fitting the following parametric distributions to the observed data:

- Exponential
- Weibull

© Roche (2022). All rights reserved

- Log-normal
- Log-logistic
- Gompertz
- Generalised gamma

These parametric extrapolations can be used directly for the entire time horizon of the model. A fully parametric approach was preferred due to data scarcity and pragmatism and so other model types were evaluated.

The base case, parameters for each treatment were selected in line with recommendations in TSD 14[96]. Firstly, the six distributions considered were ranked based upon the Akaike Information Criterion (AIC) statistic, noting that distributions that within 5 points of each other should be considered as effectively indistinguishable. The AIC ranking was followed by graphical assessment of the visual fit of the distribution to the adjusted (mosunetuzumab) and unadjusted (comparator) data and assessment of the empirical hazard data to see if it was suggestive of specific distributions (such as a constant hazard suggesting an exponential). Distributions that were poor visual fits or produced clearly implausible projections were discounted, with the remaining distribution with the lowest AIC statistic chosen in the base case. The chosen distributions were validated for long-term plausibility by eight clinical experts at an Advisory Board and supported by analysis of real-world data (RWD). A Bayesian analysis was conducted which uses the 5-year survival as a prior in the parametric extrapolations. This is used to constrain parameter estimates and their uncertainty towards realistic outcomes; this is described more fully in Section B.3.3.3.1.

B.3.3.2.1 Rituximab plus lenalidomide

B.3.3.2.1.1 Progression Free Survival

The comparison between mosunetuzumab and rituximab plus lenalidomide (R^2) is informed by the MAIC adjusted mosunetuzumab population (ESS n=32.9) and the unadjusted R^2 population (n=147) as presented in Section B.2.9.1.2 and Table 13.

It was not feasible to fully harmonise the inclusion and exclusion criteria between the AUGMENT (R^2) trial and the GO29781 trial as this led to unacceptably low ESS numbers. This introduced an important bias as many patients in the AUGMENT trial had only one previous line of treatment compared to those in the GO29781 trial. The overall (FL and non-FL) population of the AUGMENT trial contained only 47% of patients who were 3L+ patients[49]. In addition, AUGMENT included patients whose disease was not refractory to rituximab, but the entire GO29781 pivotal cohort, including patients refractory to an anti-CD20 antibody, had to be used in the analysis to ensure adequate patient numbers (see Section B.2.9.2.1).

© Roche (2022). All rights reserved

Therefore, their expected outcomes are thought to be bias against mosunetuzumab for this reason and it is important to view results of the efficacy estimates with this in mind.

Figure 22 displays the KM for mosunetuzumab and R^2 and shows that mosunetuzumab is consistently estimated to be slightly below R^2. This is not unexpected as there are notable differences in the underlying populations and it is reasonable to assume that those in the AUGMENT trial (R^2) would experience improved PFS, and likely OS. The log negative log plot (Figure 23) shows some convergence in the latter time periods. This, combined with the Schoenfeld test (p= 0.0287) would not allow acceptance of the proportional hazards assumption.

[redacted]

Figure 22. PFS Kaplan Meier for mosunetuzumab (adjusted) and R^2 (unadjusted)

[redacted]

Figure 23. PFS log negative log plot for mosunetuzumab (adjusted) and R^2 (unadjusted)

The parametric choice for mosunetuzumab was based on an overall assessment of all the individual ITC weighted mosunetuzumab populations. A single distribution for mosunetuzumab was chosen for all comparisons, as it was assumed that there would not be a critically different shape of the hazard based on differences in the populations.

AIC and BIC statistics were calculated for the six distributions considered (Table 25). For mosunetuzumab, the exponential or generalised gamma distribution were the highest ranked distribution with the log-normal the second ranked distribution, but all other distributions fell within five points of the generalised gamma distribution. For R^2, the log-normal was the highest ranked distribution with the log-normal and generalised gamma and log-logistic being within five points. Analysis of survival and hazard plots (Figure 24) suggested that the shape of the hazard in the KM data in both treatment arms indicates a concave shaped parametric hazard, which is compatible with log-normal and log-logistic models. Clinical experts at the Advisory Board considered that both the log-normal and log-logistic model produced the most plausible PFS estimates for mosunetuzumab. Taking the above factors into account the log-normal distribution was chosen for both mosunetuzumab and R^2 base case.

[redacted]

Figure 24. PFS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and R^2 (unadjusted)

© Roche (2022). All rights reserved

Table 25. AIC and BIC for PFS (R^2 and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for RL)						
Parameter 1	0.0264	0.049	3.6012	0	0.0482	0.0474
Parameter 2		0.7675	2.0379	0.0009	0.86	-0.0857
Parameter 3				-139.5818		
AIC	100.264	101.292	99.968	98.603	100.862	100.312
BIC	102.764	106.292	104.968	106.102	105.862	105.311
AIC Ranking	3	6	2	1	5	4
BIC Ranking	1	6	2	5	4	3
Rituximab + Lenalidomide						
Parameter 1	0.0185	0.0091	3.5833	14.4075	0.006	0.0175
Parameter 2		1.2195	1.2279	4.4146	1.4444	0.0041
Parameter 3				-0.3479		
AIC	560.896	560.173	554.943	556.403	557.017	562.795
BIC	563.886	566.153	560.924	565.374	562.998	568.776
AIC Ranking	5	4	1	2	3	6
BIC Ranking	3	5	1	4	2	6

B.3.3.2.1.2 Overall Survival

Figure 25 presents the KM plots for mosunetuzumab (adjusted) and $R^{2.}$ (unadjusted). Follow up for both OS and PFS was longer for R^2 than mosunetuzumab. Mosunetuzumab data, though immature, indicated a potentially more horizontal future trajectory compared with R^2. Examination of the log negative log hazard plots (Figure 26) demonstrated an early crossing for OS, although Schoenfeld residuals and corresponding test (p= 0.706) indicated that there was no reason to reject the proportional hazards assumption.

Figure 25. OS Kaplan Meier for mosunetuzumab (adjusted) and R^2 (unadjusted)

Figure 26. OS log negative log plot for mosunetuzumab (adjusted) and R^2 (unadjusted)

AIC and BIC statistics were calculated for the six distributions considered (Table 26). For mosunetuzumab, the exponential was the highest ranked distribution with the log-normal the second ranked distribution, but all other distributions were within five points of the exponential

© Roche (2022). All rights reserved

distribution. For R^2, the generalised gamma or exponential was the highest ranked distribution with the log-normal the second ranked but again all distributions were within five points of the highest ranked generalised gamma distribution. As was the case with PFS, analysis of survival and hazard plots (Figure 27) suggested that the shape of the hazard in the KM data in the R^2 treatment arm indicates a concave shaped parametric hazard, which is compatible with log-normal and log-logistic models. This judgement is not as easy to make in the mosunetuzumab arm. Clinical experts at the Advisory Board considered that that they expected little difference in OS for mosunetuzumab and any comparator. In addition, the Bayesian Analysis validation exercise (B.3.3.3.1) indicated that linearly decreasing or constant hazards might be the most appropriate. Taking the above factors into account, and the discordance of the most appropriate modelling approach, the exponential model was chosen as the base case distribution for mosunetuzumab, while R^2 is represented by the Weibull distribution. Though not the most highly ranking according to fit statistics, all models are within 5 points by both measures indicating that there is little agreement from this measure. Other models are considered in scenario analysis to examine the impact of this choice (B.3.11.3.1).

[redacted]

Figure 27. OS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and R^2 (unadjusted)

Table 26. AIC and BIC for OS (R^2 and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for RL)						
Parameter 1	0.004	0.001	5.0798	17.4849	0.0009	0.0029
Parameter 2		1.4796	1.4965	5.3556	1.5208	0.0359
Parameter 3				-0.2284		
AIC	30.808	32.474	32.365	34.356	32.45	32.708
BIC	33.308	37.474	37.364	41.856	37.449	37.708
AIC Ranking	1	4	2	6	3	5
BIC Ranking	1	4	2	6	3	5
Rituximab + Lenalidomide						
Parameter 1	0.0026	0.0007	5.5598	1001.959	0.0007	0.0022
Parameter 2		1.3684	1.5151	0.0186	1.4081	0.011
Parameter 3				-3.516		
AIC	154.62	155.484	154.293	153.402	155.346	156.462
BIC	157.611	161.465	160.274	162.373	161.327	162.443
AIC Ranking	3	5	2	1	4	6
BIC Ranking	1	4	2	5	3	6

© Roche (2022). All rights reserved

B.3.3.2.2 Rituximab plus bendamustine

B.3.3.2.2.1 Progression Free Survival

As noted in Section B.2.9, it was not feasible to perform a robust ITC between mosunetuzumab and R-CHOP. Rituximab in combination with chemotherapy was therefore represented in the model by rituximab and bendamustine (RB) as it was feasible to perform a comparison between mosunetuzumab and this regimen. The RB arm was informed by pooling relevant patients from the CONTRALTO and GO29365 trials (see Section B.2.9.1.2).

In order to ensure that the patient cohorts used for analysis were as homogenous as possible, patients were filtered such that they matched those in the GO29781 trial and resulted in 46 matching patients in the RB arm and 46 in the mosunetuzumab arm. Even after optimal pair matching, there were notable differences in key prognostic factors such as ECOG 1 vs 0%, refractory to last therapy line, prior ASCT and time since completion of last therapy (Table 16) meaning results of this analysis should be considered with these caveats in mind.

As with R^2, a longer follow up time was available for RB than for mosunetuzumab PFS and OS (Figure 28). Mosunetuzumab PFS tracks slightly underneath RB until approximately 15 months where a crossing is evident. The log negative log curves confirm that the hazard is not parallel, with crossing in the latter period (Figure 29). The Schoenfeld test did not require the proportional hazards assumption to be rejected (p=0.151) though the late crossing meant it was deemed sensible to fit independent models.

[redacted]

Figure 28. PFS Kaplan Meier for mosunetuzumab (adjusted) and RB (unadjusted)

[redacted]

Figure 29. PFS log negative log plot for mosunetuzumab (adjusted) and RB (unadjusted)

AIC and BIC statistics were calculated for the six distributions considered (Table 27). For mosunetuzumab and RB, the lognormal or exponential was the highest ranked distribution but all other distributions were within five points of the log-normal distribution. Analysis of survival and hazard plots (Figure 30) suggested that for RB the shape of the hazard indicates a concave shaped parametric hazard, which is compatible with log-normal and log-logistic models. Clinical experts at the Advisory Board considered that both the log-normal and log-logistic model produced the most plausible PFS estimates for mosunetuzumab. Taking the

© Roche (2022). All rights reserved

above factors into account, the log-normal base case distribution chosen for both mosunetuzumab and RB.

[redacted]

Figure 30. PFS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and RB (unadjusted)

Table 27. AIC and BIC for PFS (RB and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for RB)						
Parameter 1	0.0323	0.0247	3.0861	8.4038	0.0215	0.0302
Parameter 2		1.0991	1.3982	2.9739	1.2552	0.0082
Parameter 3				-0.3679		
AIC	287.447	289.1	287.278	289.000	288.738	289.377
BIC	289.842	293.889	292.067	296.183	293.527	294.166
AIC Ranking	2	5	1	4	3	6
BIC Ranking	1	4	2	6	3	5
Rituximab + Bendamustine						
Parameter 1	0.0328	0.0179	3.0087	5.1002	0.0089	0.0315
Parameter 2		1.1914	1.1437	11.9623	1.5747	0.003
Parameter 3				0.2728		
AIC	225.309	226.305	225.268	227.074	224.52	227.284
BIC	227.137	229.962	228.926	232.56	228.177	230.942
AIC Ranking	3	4	2	5	1	6
BIC Ranking	1	4	3	6	2	5

B.3.3.2.2.2 Overall survival

The KM data for RB and mosunetuzumab suggest potentially different trajectories, with mosunetuzumab displaying potential for a more horizontal trajectory than that observed with RB (Figure 31). The log negative log plot shows early crossing, but this combined with the Schoenfeld test (p=0.614) did not give reason to reject the proportional hazards assumption (Figure 32). For consistency with other treatments, it was considered appropriate to fit models independently.

[redacted]

Figure 31. OS Kaplan Meier for mosunetuzumab (adjusted) and RB (unadjusted)

© Roche (2022). All rights reserved

[redacted]

Figure 32. OS log negative log plot for mosunetuzumab (adjusted) and RB (unadjusted)

AIC and BIC statistics were calculated for the six distributions considered (Table 28). For mosunetuzumab, the exponential was the highest ranked distribution with the log-normal the second ranked distribution, but all other distributions were within five points of the exponential distribution. For OB, the generalised gamma was the highest ranked distribution with the exponential the second ranked but again all distributions were within five points of the highest ranked generalised gamma distribution. Analysis of survival and hazard plots (Figure 33) suggested that the shape of the hazard in the KM data in both treatment arms indicates a concave shaped parametric hazard, which is compatible with log-normal and log-logistic models. Clinical experts at the Advisory Board considered that they expected little difference in OS for mosunetuzumab and any comparator. In addition, the Bayesian Analysis validation exercise (B.3.3.3.1) indicated that linearly decreasing or constant hazards might be the most appropriate. Taking the above factors into account the exponential model was chosen in the base case to represent mosunetuzumab and the Weibull distribution was chosen for RB. The Weibull model estimates a shape parameter below one which represents a decreasing hazard, in line with the Bayesian Analysis conclusions. The fit statistics for all models are within five points for both measures, suggesting little difference between them. To address the impact of this decision, scenario analysis is conducted to consider alternative models (B.3.11.3.1) and the impact of assuming proportional hazards as there was no reason to reject this initially (B.3.11.3.2).

[redacted]

Figure 33. OS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and RB (unadjusted)

Table 28. AIC and BIC for OS (RB and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for RB)						
Parameter 1	0.0051	0.0037	5.4666	4.4594	0.0035	0.0054
Parameter 2		1.1138	1.9137	0.2252	1.147	-0.006
Parameter 3				-0.6166		
AIC	87.651	89.566	89.172	90.944	89.506	89.643
BIC	90.046	94.355	93.961	98.127	94.295	94.431
AIC Ranking	1	4	2	6	3	5
BIC Ranking	1	4	2	6	3	5

© Roche (2022). All rights reserved

Rituximab + Bendamustine						
Parameter 1	0.007	0.0078	5.0189	1.97E+82	0.0073	0.0073
Parameter 2		0.9666	1.8766	0.0008	1.0194	-0.0026
Parameter 3				-134.596		
AIC	86.396	88.385	87.855	84.555	88.381	88.389
BIC	88.224	92.042	91.512	90.04	92.038	92.047
AIC Ranking	2	5	3	1	4	6
BIC Ranking	1	5	3	2	4	6

B.3.3.2.3 Obinutuzumab with bendamustine

B.3.3.2.3.1 Progression Free Survival

As described in Section B.2.9 the comparison between obinutuzumab and bendamustine (OB) was informed by the propensity score analysis and results in mosunetuzumab being informed by 32.6 relevant patients after matching and OB being informed by 47.2 relevant patients from the GADOLIN trial.

As with other analyses presented, there were notable differences to be adjusted for between the GADOLIN and GO29781 populations and even after full matching, important differences were noted between age, Ann Arbor Stage III/IV %, refractory to last line % and double refractory % (Table 15).

Figure 34 presents PFS KM plots for mosunetuzumab and OB. Follow up for OB was longer than for mosunetuzumab for both OS and PFS. The log negative hazard plots indicate that it is likely the proportional hazards assumption may hold for both outcomes with no convergence and equidistance for most time points for PFS (Figure 35). The Schoenfeld residual test (p=0.335) and plots confirm that there is no reason to reject the assumption of proportional hazards in this arm. For consistency with other treatments, it was considered appropriate to fit models independently. The impact of this assumption is addressed in scenario analysis (Section B.3.11.3.2).

Figure 34. PFS Kaplan Meier for mosunetuzumab (adjusted) and OB (unadjusted)

Figure 35. PFS log negative log plot for mosunetuzumab (adjusted) and OB (unadjusted)

© Roche (2022). All rights reserved

AIC and BIC statistics were calculated for the six distributions considered (Table 29). For mosunetuzumab, the lognormal distribution was the highest ranked distribution with all other distributions within five points. For OB, the Gompertz was the highest ranked distribution but again all distributions were within five points. Analysis of survival and hazard plots (Figure 36) suggests all curves are reasonable fits to the KM data. Taking the above factors into account, the log-normal distribution was chosen for mosunetuzumab base case and the Gompertz model for OB base case.

[redacted]

Figure 36. PFS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and OB (unadjusted)

Table 29. AIC and BIC for PFS (OB and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for OB)						
Parameter 1	0.0536	0.0493	2.4917	3.9684	0.0404	0.0598
Parameter 2		1.033	1.2934	0.8802	1.2895	-0.0148
Parameter 3				-0.7841		
AIC	290.985	292.931	288.507	288.938	290.858	292.722
BIC	293.247	297.457	293.032	295.726	295.384	297.247
AIC Ranking	4	6	1	2	3	5
BIC Ranking	2	6	1	4	3	5
Obinutuzumab + Bendamustine						
Parameter 1	0.0251	0.0139	3.2483	0.0016	0.0085	0.0196
Parameter 2		1.1654	1.1727	0.6457	1.4489	0.0109
Parameter 3				1.6007		
AIC	437.044	437.446	439.485	439.412	441.579	436.806
BIC	439.388	442.134	444.173	446.443	446.267	441.494
AIC Ranking	2	3	5	4	6	1
BIC Ranking	1	3	4	6	5	2

B.3.3.2.3.2 Overall Survival

The available data to inform OB survival was considerably longer than for mosunetuzumab. However, the plot KM data shows that it is likely that survival with mosunetuzumab could be improved when compared to that of survival with OB (Figure 37). The log negative log plot shows parallel hazards over time (Figure 38) and the Schoenfeld test indicated no reason to reject the proportional hazards assumption (p=0.953).

[redacted]

Figure 37. OS Kaplan Meier for mosunetuzumab (adjusted) and OB (unadjusted)

© Roche (2022). All rights reserved

Figure 38. OS log negative log plot for mosunetuzumab (adjusted) and OB (unadjusted)

AIC and BIC statistics were calculated for the six distributions considered (Table 30). For mosunetuzumab and OB, the exponential was the highest ranked distribution with the log-normal the second ranked distribution, but all other distributions were within five points of the exponential distribution for both treatments. For OB, the generalised gamma was the highest ranked distribution with the log-normal the second ranked, but all distributions scored within five points. Analysis of survival and hazard plots (Figure 39) suggested that the shape of the hazard in the KM data in both treatment arms indicates a concave shaped parametric hazard, which is compatible with log-normal and log-logistic models. Clinical experts at the Advisory Board considered that they expected little difference in OS for mosunetuzumab and any comparator. Taking the above factors into account the log-normal base case distribution chosen for OB and exponential for mosunetuzumab.

Figure 39. OS hazard and survival plots for distributions considered for mosunetuzumab (adjusted) and OB (unadjusted)

Table 30. AIC and BIC for OS (OB and mosunetuzumab)

	Exponential	Weibull	Log Normal	Generalised gamma	Log Logistic	Gompertz
Mosunetuzumab (adjusted for OB)						
Parameter 1	0.007	0.0031	4.7888	56.7605	0.0029	0.0055
Parameter 2		1.2805	1.6125	35.9933	1.3324	0.0246
Parameter 3				-0.0948		
AIC	93.266	94.773	94.571	96.567	94.716	95.089
BIC	95.528	99.299	99.096	103.355	99.242	99.614
AIC Ranking	1	4	2	6	3	5
BIC Ranking	1	4	2	6	3	5
Obinutuzumab + Bendamustine						
Parameter 1	0.0081	0.0081	4.4992	94.6939	0.0061	0.0081
Parameter 2		1.0012	1.559	123.9072	1.139	0.0002
Parameter 3				0.0593		
AIC	379.806	381.806	381.128	383.12	381.956	381.805
BIC	382.15	386.493	385.816	390.151	386.643	386.493
AIC Ranking	1	4	2	6	5	3
BIC Ranking	1	3	2	6	5	4

© Roche (2022). All rights reserved

B.3.3.3 Validation of survival curves applied in the economic evaluation

It was noted that the data sources available to inform clinical efficacy were often of limited follow up. This is common in indications where the population of interest is small. It is necessary to be pragmatic when evaluating extrapolations made from data that is collected in populations with either limited population numbers or follow up. As such, in addition to clinical advice on distribution selection for OS and PFS, the distribution choices in the base case were validated by a comparison with real world data (RWD) on OS.

B.3.3.3.1 Bayesian Meta-Analysis of OS in RWD cohorts

Roche performed a meta-analysis and Bayesian analysis of OS, leveraging priors from RWD sources.

The frequentist approach that underpins the base case efficacy can be sensitive to small changes in event times as the surviving proportion at follow up is high and there is often a high amount of censoring. This can lead to low consistency in estimates between the parametric models and infeasible estimates made by some (as noted in Sections B.3.3.2.1– B.3.3.2.3).

Eight RWD cohorts[97-104] were identified in 3L+ FL patients which had long follow up times. These can be used within a Bayesian framework to obtain realistic extrapolations. Longer follow up data also allows for more complex hazard functions to be examined and consider uncertainty in the models. The identified RWD cohorts were used to inform the prior which was then integrated into the available clinical data. Specifically, the estimate of OS at 5 years from the RWD cohorts was used as a prior for a random effects meta-analysis made in a Bayesian framework. Independent parametric models were fit to each respective comparator arm to align with the analysis presented in Section B.3.3.2 and the assumption of non-proportional hazards. No adjustment for heterogeneity was made in order to preserve the reflection of the underlying population and provide the most realistic representation of the population in England.

Data from the LEO and NLCS studies were informed by patients that were tracked since the start of their therapy. Data from the remaining sources were estimated by digitising OS at each therapy line, and pooling to obtain OS at 3L+. Limitations with this technique are noted; patients who were alive at multiple lines would be considered independent patients and within-patient correlation cannot be accounted for. As individual patient data could not be made available, this limitation is acknowledged, and results should be caveated with this in mind. As this technique resulted in a larger than expected number of patients at 3L, sampling was performed from the expected number to avoid resulting models being unrealistically optimistic

© Roche (2022). All rights reserved

and to reflect a more probable population sample size. Patient numbers at each line were available from ReCORD-FL and so this was used to match patients to treatment line.

From examining these data, it was possible to evaluate the shape of the underlying hazard to understand whether RWD supports a parametric model for OS that indicates an increasing, decreasing or constant hazard. Hazards were estimated by bootstrapping smoothed hazards with the muhaz package in R (Figure 40). The hazards show that six sources[98,100-104] indicate decreasing hazards and two[97,99] constant hazards for OS.

Figure 40. Hazard function for RWD cohorts at 3L+

The same data was used to inform a meta-analysis of survival at five and eight years. The model assumes that survival at the respective time point is log-normally distributed, and a Bayesian random effects meta-analysis was performed considering 20,000 iterations (burn in 5,000). Priors were uninformative to ensure that the data were allowed to dominate estimate.

The analysis estimated a mean survival of 60% at 5 years, with a 95% credibility interval of 49% - 84% and 53% at 8 years, with a 95% credibility interval of 40% - 85%.

The meta-analysis provided OS estimates that can be used for Bayesian extrapolations. These extrapolations assumed a normal prior on survival at five years and were performed for Exponential, Weibull, Log-Normal, Log-Logistic and Gompertz distributions. Estimation was done in R with the rjags package and considered 5,000 iterations with three chains. The method employed is in line with those discussed in the NICE Decision Support Unit Technical Support Document 21[105] and other recent publications[106-108]. These models represent realistic OS outcomes and are used predominantly to validate the extrapolations made from robust ITC analyses.

The RWD supports the use of OS models that represent constant or linearly decreasing hazards such as the exponential used in the base case. In addition, estimates align with those predicted by the economic model thus increasing certainty in the efficacy underpinning the decision problem.

B.3.3.4 All-cause mortality

The model included age and sex-adjusted mortality based on information from UK life tables. These values were included in every cycle in addition to the disease-related mortality values and were applied multiplicatively. While some form of double counting is likely to occur, this effect applies equally to all comparators and is likely to have a minimal impact on predicted survival (and hence cost-effectiveness).

© Roche (2022). All rights reserved

The model also contains an option to model all-cause mortality using the distribution of ages observed in the GO29781 trial. This method essentially allows for the change in population heterogeneity over time by resetting population demographics to at discrete time points (each cycle) to account for those subjects who have experienced a death event[109]. This option is available in the CEM and shown as a scenario.

B.3.3.5 Treatment discontinuation

In the base-case, the GO29781 trial KM data on treatment duration (TTOT) was directly used in the model, limited to not exceed PFS. As TTOT data was complete from GO29781, there was no need to fit a distribution to the KM data and, as treatment stops at 12 months, there was no need for curve fitting for extrapolation. Using the KM TTOT data directly removes adding an unnecessary level of uncertainty resulting from curve fitting. For all other treatments, the time to off-treatment (TTOT) was set to be equal to the selected parametric distribution for PFS as no TTOT curves were available; these were capped at the treatment-specific maximum number of cycles if relevant. Base case estimates of TTOT for all modelled treatments are presented in Table 31.

[redacted]Table 31. Base case estimates for TTOT

	Mosunetuzumab	RB (Rituximab	RB (Bendamustine)	OB (Obinutuzumab)	OB (Bendamustine)	R^2 (Rituximab)	R^2 (Lenalidomide)
Model results, time on treatment							
Mean number (cycles)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean time (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median time (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Proportion still on treatment							
0 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

© Roche (2022). All rights reserved

12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations: OB, obinutuzumab and bendamustine; RB, rituximab and bendamustine; R^2, rituximab and lenalidomide							

B.3.3.6 Adverse events

Adverse events (AEs) are an inevitable consequence of any intervention and, to reflect this, were applied in the model affecting costs and QALYs accrued with each intervention. Only treatment-related AEs with a severity grade of 3 or higher that occurred in greater than 2% of the population were considered in the model (see

Table 32) to reflect those events that are most likely to impact cost-effectiveness.

Incidence rates from mosunetuzumab and comparator trials were converted into per-cycle equivalents based on number of patients experiencing an event and follow-up using standard formulae; relevant model inputs are summarised in Table 33. It was not feasible to perform a safety ITC and so the reported durations and incidence of AE was used directly as reported from the respective sources.

Table 32. Adverse events considered in the model

Adverse event, no. of events (no. of patients with event)	Number of Grade 3-5 TRAE in mosunetuzumab arm	Number of Grade 3-5 TRAE in the RB arm[51]	Number of Grade 3-5 TRAE in the OB arm[50]	Number of Grade 3-5 TRAE in the R^2 arm[49]
Alanine aminotransferase increased	4 (4)	0	0	3 (3)
Anaemia	4 (3)	1 (1)	15 (15)	8 (8)
Aspartate aminotransferase increased	3 (3)	0	0	0
Cytokine release syndrome	2 (2)	0	0	0
Hypokalaemia	0	1 (1)	0	0
Hypophosphatemia	16 (11)	0	0	0
Lymphopenia	3 (3)	0	0	0
Neutropenia	20 (13)	14 (14)	64 (64)	88 (88)
Neutrophil count decreased	17 (7)	0	0	0

© Roche (2022). All rights reserved

Rash erythematous	2 (2)	0	0	0
Thrombocytopenia	0	3 (3)	21 (21)	4 (4)
Tumour flare	2 (2)	0	0	0
Upper respiratory tract infection	2 (2)	0	4 (4)	0
Diarrhoea	0	0	2 (2)	5 (5)
Febrile neutropenia	0	3 (3)	0	0
Vomiting	0	0	4 (4)	0
Infections	0	0	0	4 (4)
Leukopenia	0	2 (2)	0	12 (12)
Cutaneous reaction	0	0	0	10 (10)
Abbreviations: OB, obinutuzumab and bendamustine; RB, rituximab and bendamustine; R2, rituximab and lenalidomide; TRAE, treatment-related adverse events				

Table 33. Adverse event probabilities derived for use in the model

Adverse event	Per cycle rate in mosunetuzumab arm	Per cycle rate in RB arm	Per cycle rate in OB arm	Per cycle rate in R^2 arm
Alanine aminotransferase increased	0.0006	0.0000	0.0000	0.0001
Anaemia	0.0006	0.0003	0.0008	0.0004
Aspartate aminotransferase increased	0.0004	0.0000	0.0000	0.0000
Cytokine release syndrome	0.0003	0.0000	0.0000	0.0000
Hypokalaemia	0.0000	0.0003	0.0000	0.0000
Hypophosphatemia	0.0022	0.0000	0.0000	0.0000
Lymphopenia	0.0004	0.0000	0.0000	0.0000
Neutropenia	0.0028	0.0036	0.0036	0.0041
Neutrophil count decreased	0.0024	0.0000	0.0000	0.0000
Rash erythematous	0.0003	0.0000	0.0000	0.0000
Thrombocytopenia	0.0000	0.0008	0.0012	0.0002
Tumour flare	0.0003	0.0000	0.0000	0.0000
Upper respiratory tract infection	0.0003	0.0000	0.0002	0.0000
Diarrhoea	0.0000	0.0000	0.0001	0.0002
Febrile neutropenia	0.0000	0.0008	0.0000	0.0000
Vomiting	0.0000	0.0000	0.0002	0.0000
Infections	0.0000	0.0000	0.0000	0.0002
Leukopenia	0.0000	0.0005	0.0000	0.0006
Cutaneous reaction	0.0000	0.0000	0.0000	0.0005
Abbreviations: OB, obinutuzumab and bendamustine; RB, rituximab and bendamustine; R^2, rituximab and lenalidomide				

© Roche (2022). All rights reserved

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality-of-life studies

An SLR was conducted to identify studies evaluating HRQoL in the target population. Further details of the SLR can be found in the report provided as Appendix G. The SLR identified six studies that reported HRQoL data, two of which were available as full text manuscripts and four as abstracts. These findings demonstrate the sparsity of literature on utilities available in this indication. Disease status appears to be a primary factor influencing HRQoL, with patients who experience progressive disease, those on later lines of therapy, and those receiving maintenance therapy showing a clear deterioration in HRQoL. All identified studies reported substantial differences in utility by disease status and therapy lines. In all studies except one (Choi et al. 2015[110]), utilities were elicited directly from patients.

Three of the identified studies reported UK-specific utilities[16,63,111], of which one used the preferred three-level EQ-5D questionnaire[63] and one did not specify the EQ-5D version used[111].

In addition, relevant TAs were investigated for appropriate evidence and thirteen were found from CADTH, NICE, SMC and PBAC submissions. In most, the source of health state utility values was not reported (or available publicly). Of the remaining that provided the relevant data, sources included the AUGMENT trial[49] (NICE TA627[24], SMC2281[76]), Pettengel et al 2008[15] (NICE TA604[27]), and Wild et al 2006[92], which informed the majority of TAs either entirely or in part (SMC 1039/15[78], SMC 1219/17[77], NICE TA 137[26],NICE TA604[27], NICE TA629[25]). It is important to note that, while relevant, not all identified utility values from TAs are specifically for 3L+ FL patients. The values presented by Wild (2006)[92] have been used in submissions to NICE and SMC to represent relapsed or refractory patients, and specifically patients who are refractory to two previous lines. Full detail of these studies can be seen in Appendix G.

B.3.4.2 Health-related quality-of-life data from clinical trials

HRQoL in the GO29781 trial was assessed using the EORTC QLQ-C30 v3.0 questionnaire and the 15-item FACT-Lym subscale. Health status was assessed using the EQ-5D-5L questionnaire. PRO analyses by scheduled visits focused on patients who were still receiving initial treatment with mosunetuzumab at the time of PRO questionnaire administration. Typically, it was observed that following the Cycle 8 assessment timepoint, less than 25% of the baseline patient population within remained evaluable for PRO analyses.

© Roche (2022). All rights reserved

PRO data was collected in GO29781 pre-infusion, and then at the first day of every other cycle, starting at cycle 2 (see also Section B.2.3.3). An additional collection was made at the end of treatment completion (regardless of reason).

The baseline mean EQ-5D-5L index utility scores for each dimension in the pivotal cohort reflected a low level of impairment in their health status at baseline. Both EORTC QLQ-C30 and FACT-Lym subscale indicated similar, although highlighted slightly elevated fatigue levels and burden of lymphoma specific symptoms, respectively (Section B.2.6.4). As EQ-5D is the preferred HRQoL measure for NICE, subsequent sections focus on this measure.

B.3.4.2.1 GO29781 HRQoL data analysis

All analysis of EQ-5D-5L data in the trial was conducted in R, version 3.6.3.

Utility values associated with mosunetuzumab (Table 34) were informed by the results of analysis from the GO29781 study. Utilities were measured using the EQ-5D-5L instrument and valued with the UK specific tariff in line with NICE recommended methods[112]. These were mapped to the EQ-5D-3L instrument; further detail is provided in Section B.3.4.3.

At baseline, 83 observations of 90 were available and a mean of 0.767 (SE 0.17) was measured indicating a reduced utility for patients compared to an age-matched general population.

As only a small number of patients were available from the trial to inform analysis, a pragmatic approach was taken, and health state utilities were calculated for PFS and PPS. Estimates by progression status are informed by the date of progression for each patient unless it cannot be assigned due to censoring, in which case it is considered unknown and is not included in analysis.

Linear mixed regression models on post baseline utilities, controlling for centralised baseline utilities, using random intercepts for each patient were used. This approach was taken as it is considered robust to violations of distributional assumptions [113]. Results for are shown in Table 34.

A brazier age-adjusted health state utility value coefficient was also applied (Table 35). This age-adjustment is a linear estimation of how utility changes in the general population as a function of sex and age. In this model, the linear function was used to calculate a multiplier, corresponding to proportional utility loss as a function of age, which was used in the final calculation of QALYs for each cycle in each treatment model.

Table 34. Utility estimates from GO29781

State	Utility Value (SE)	95% CI

© Roche (2022). All rights reserved

PFS	0.804 (0.011)	(0.782 - 0.826)
PPS	0.75 (0.02)	(0.712 - 0.788)

Table 35. Brazier age-adjusted coefficients

Parameter	Estimate (SE)	95% CI
(Intercept)	0.55149 (0.24527)	(0.11776 - 1.06449)
sexM	0.03502 (0.03319)	(-0.02984 - 0.0955)
age	0.01094 (0.00836)	(-0.00694 - 0.0262)
age2	-0.00011 (0.00007)	(-0.00025 - 0.00004)

B.3.4.3 Mapping

As trial data were collected using the EQ-5D-5L instrument, it was necessary to map this to the preferred three level instrument. All analysis was done in line with recommendations made by the Decision Support Unit (DSU)[114]. Analysis was performed with the R code made available with these recommendations[112].

B.3.4.4 Adverse reactions

It was not possible to conduct an ITC for safety outcomes due to data sparsity. As such, the information relating to AEs contained within the CEM and reported in this document is taken directly from literature and represents a naïve comparison with mosunetuzumab. Utility data was available from the GO29781 trial, but not for any of the other treatment arms modelled. The PFS values estimated from this trial analysis are considered to represent the HRQL experienced by patients when they are pre-progression and are further considered to account for any potential adverse reactions. Therefore, it was not considered sensible to include specific disutilities for any adverse reactions as this would constitute double counting.

A conservative assumption was made that the HRQL experienced in all arms is consistent and most related to the health state rather than toxicity. The most impactful AEs are evident early after treatment onset (such as CRS, which occurred predominantly after the Day 15 dose in Cycle 1, see Section B.2.10.1.4) and these will be captured within the PFS health state measurement.

While it may have been possible to collect disutility estimates for some AEs experienced, these were not collected within a comparative trial. It was therefore considered that including disutilities and combining these with rates from a naïve comparison, would introduce unnecessary uncertainty to the decision problem.

© Roche (2022). All rights reserved

B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

In the base case, the health state utility values from GO29781 were used for all arms and therefore, it was assumed that the utility of patients in each treatment arm is comparable (Table 36). It is acknowledged that values from literature differ, particularly with respect to the PPS health state. As such, scenarios are presented where health state utility is represented by the values reported by Wild et al. 2006[92] and Cognet et al. 2015[111] (Section B.3.11.3.3).

Table 36. Base case utility values and scenario utility values

Scenario	State	Utility values	Standard Error
Base case	PFS	0.804	0.01
	PPS	0.75	0.02
Scenario (Wild et al 2006)	PFS	0.81	0.02
	PPS	0.62	0.06
Scenario (Cognet et al 2015)	PFS	0.71	0.30
	PPS	0.51	0.35

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

B.3.5.1 Published costs and resources studies

An SLR was conducted to identify studies describing the costs and resource use associated with the management of patients with FL. In brief, electronic database searches (Embase, MEDLINE, Evidence Based Medicine [EBM], and EconLit) were conducted in January 2022. Supplementary sources were hand searched for completeness, including reference lists of included studies, conference proceedings, relevant additional databased and websites, and global HTA body websites. In total, 13 publications were identified (5 full publications and 8 conference abstracts). Details of the SLR can be found in the report provided as Appendix H.

Of the studies identified in the SLR, five were retrospective cohort studies[115-119], three were retrospective real-world studies[120-122], three were cost analyses[123-125], one was a retrospective database analyses[126], and one was an economic evaluation[127].

The patient populations considered across the 13 studies included patients with R/R indolent NHL (n=4)[122-124,126], patients with FL who were previously untreated/initiating 1L therapy (data reported by treatment line or by relapse status) (n=3)[119,120,127], patients with R/R FL in the 3L+ setting (n=1)[125], patients with R/R FL, refractory to rituximab and an alkylating agent (n=1)[121], adult patients with grade I/II FL (note: allogeneic SCT, autologous SCT, and rituximab groups consisted entirely of patients receiving treatment in the 2L+ setting only) (n=1)[116], patients with

© Roche (2022). All rights reserved

aggressive or indolent NHL (data reported for progressed [2L+] and non-progressed [1L] disease) (n=1)[118], patients with FL or marginal zone lymphoma (MZL) (data reported by treatment line [1L-4L]) (n=1)[117] and patients with FL with or without progression (n=1)[115]. All 13 included studies reported direct medical costs[115-127]. Other outcomes reported across the studies included: cost drivers (n=9)[115,116,118,120-123,127]; healthcare resource use (n=5)[115,116,118,120,123]; length of stay (LOS) (n=2)[120,121]; and indirect costs (n=1)[124]. No studies were identified which reported direct non-medical costs associated with FL/ indolent NHL in the 2L+ setting.

B.3.5.2 Intervention and comparators' costs and resource use

B.3.5.2.1 Mosunetuzumab costs

The costs of mosunetuzumab, including drug procurement (Table 37) and administration (Table 38), were applied in each cycle, based on acquisition and administration costs. Administration of mosunetuzumab was assumed to take place under supervision at hospital and has been conservatively costed as prolonged infusion, first attendance for all appointments taking place in line with the dosing schedule. Vial sharing was assumed not to happen as small vials are unlikely to be suited to this practice. Clinicians advised that they would not routinely do this with smaller vials (below 30mg).

Mosunetuzumab was assumed to be taken as in the GO29781 trial; 1mg on day 1, 2mg on day 8 and 60mg on day 15 of the first cycle. In cycle 2, 60mg was taken on the first day. In each subsequent cycle, 30mg was taken on the first day.

Table 37. Mosunetuzumab dosing and acquisition

Dosing	1/2/60/60/30mg
Dose per cycle	As above
Cost (excluding PAS)	£6,600 (30mg); £220 (1mg)
Cost per dose (excluding PAS)	1mg: £220 2mg: £440 60mg: £13,200 30mg: £6,600
Administration costs	See Table 38
Total (excluding PAS)	1mg: £220 2mg: £440 60mg: £13,200 30mg: £6,600
Abbreviations: PAS: patient access scheme	

© Roche (2022). All rights reserved

Table 38. Administration costs for mosunetuzumab

Component	National cost collection for the NHS	Cost	Inflated costs
Administration	Daycase and Reg Day/Night: Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance (SB14Z)	£403.84	NHS Reference Costs 2019 to 2020[128]

B.3.5.2.2 Patient Access Scheme

A PAS has been applied, comprising a simple discount of [redacted] from the mosunetuzumab list price. In order to best replicate the true economic impact of a positive recommendation for mosunetuzumab, the economic evaluation presented in this submission applies the PAS in the base case analysis (Table 39).

Table 39. Acquisition costs of mosunetuzumab following application of PAS

Vial Size	No PAS	PAS
1mg	£220	[redacted]
30mg	£6,600	[redacted]

B.3.5.3 Comparator costs

Comparator dosing and schedule was estimated in accordance with BNF recommendations and assumes no vial sharing where applicable (Table 40).

Rituximab was assumed to be given at 375mg/m^2 on days 1, 8, 15 and 22 in the first cycle and on the first day of cycles 2–5 when given with lenalidomide. In the same regimen, lenalidomide was assumed to be given at a dose of 20mg daily for days 1–21 of each 28-day cycle. Administration costs were assumed to be the same as for mosunetuzumab (Table 38) for the first cycle and then costed as subsequent elements of a chemotherapy cycle for all subsequent administrations (Table 42). Administration of lenalidomide was assumed to require supervision and so was costed as delivery of exclusively oral chemotherapy. Rituximab and chemotherapy was assumed to comprise rituximab with bendamustine (RB). In this regimen, rituximab was assumed to be given at 375mg/m^2 every 21days. Bendamustine was given at 90 mg/m^2 on days 1 and 2. Cycles were assumed to be 21 days. Administration costs were assumed to be the same as for mosunetuzumab (Table 38) for the first cycle and then costed as subsequent elements of a chemotherapy cycle for all subsequent administrations (Table 42).

In line with the obinutuzumab SmPC[129], obinutuzumab was assumed to be given at a fixed dose of 1,000mg on days 1,8 and 15 in the first cycle. In cycles 2–6, a fixed dose of 1,000mg was given on day 1. Cycles 2–6 were 28-day cycles. In cycles 7–18, the cycle length was

© Roche (2022). All rights reserved

assumed to be 56 days and a fixed dose of 1,000mg was given on the first day. Bendamustine was given at 90 mg/m^2 on days 2 and 3 in cycle 1 and at 90 mg/m^2 days 1 and 2 cycles 2–6. Cycles were assumed to be 28 days in line with recommendations from an international consensus panel[130]. Administration costs were assumed to be the same as for mosunetuzumab (Table 38) for the first cycle and then costed as subsequent elements of a chemotherapy cycle for all subsequent administrations (Table 42).

Table 40. Comparator dosing and acquisition

Comparator	Unit Cost	Source
Rituximab (200mg)	£314.33	BNF[131]
Rituximab (500mg)	£785.84	BNF[131]
Bendamustine (25mg)	£27.55	eMIT[132]
Bendamustine (100mg)	£65.56	eMIT[132]
Obinutuzumab (1,000mg)	£3,312.00	BNF[133]
Lenalidomide (20mg)	£4,168.50	BNF[134]

Table 41. Comparator cost per cycle

Comparator	Cost per cycle
Rituximab	£1,349.89
Bendamustine	£257.05
Obinutuzumab	£3,312.00
Lenalidomide	£1,389.50

Table 42. Comparator administration costs

Component	National cost collection for the NHS	Cost	Inflated costs
Administration (first appointment)	Daycase and Reg Day/Night: Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance (SB14Z)	£403.84	NHS Reference Costs 2019 to 2020[128]
Administration (Subsequent appointments)	Daycase and Reg Day/Night: Subsequent Elements of Chemotherapy Cycle (SB15Z)	£339.46	NHS Reference Costs 2019 to 2020[128]
Administration (oral treatment)	Daycase and Reg Day/Night: Deliver Exclusively Oral Chemotherapy (SB11Z)	£210.79	NHS Reference Costs 2019 to 2020[128]
Abbreviations: NHS: National Health Service			

B.3.5.4 Treatment costs at subsequent lines of therapy

Once patients in the model discontinued their initial treatment line after progression, they were assumed to be eligible for all other treatments available at third and subsequent lines of FL

© Roche (2022). All rights reserved

treatment. These are represented in the model as a pool of treatments that can be taken in any order after discontinuation from any arm. Subsequent therapy use was derived from a survey of the clinicians who attended the Clinical Advisory Board (Data on File). The post discontinuation therapy cost was applied once to the proportion who move from the PFS to PPS health state each cycle. This takes into account the mean duration of treatment, the proportion assumed to use each treatment option and the cost.

Table 43 shows the dosing and acquisition costs for post-discontinuation therapies. Cost. Dosing and associated costs for any regimens that are considered direct comparators, but may also be used post discontinuation are detailed in Section B.3.5.3. Table 44 and Table 45 show the total weekly post-discontinuation costs for each therapy regimen, and total cost post discontinuation for mosunetuzumab and all included comparators, respectively. Table 46 shows the mean duration and percentage share of each therapy class included in the post-discontinuation treatment for mosunetuzumab, RB, RL and OB.

Cyclophosphamide, doxorubicin and vincristine, taken alongside rituximab and assumed to represent part of the R-Chemo regimen; these were assumed to be given at doses of 750mg/m^2, 500mg/m^2, and 1.4mg/m^2, respectively, on the first day of each 21-day cycle. Prednisolone was assumed to be given at a fixed dose of 100mg per day for the first five days of each 21-day cycle. These drugs, with the exception of doxorubicin, also constituted the R-CVP therapy at the same dosage.

R-Chemotherapy costs were assumed to be average cost of RB, R-CHOP, and R-CVP

Other non-rituximab containing chemotherapies were assumed to be the average costs of OB and bendamustine.

Administration costs were assumed to be the same as for mosunetuzumab (Table 38) for the first cycle and then costed as subsequent elements of a chemotherapy cycle for all subsequent administrations (Table 42).

Table 43. Post-discontinuation therapies dosing and acquisition

Comparator	Unit Cost	Source
Prednisolone (20mg)	£3.30	eMIT[132]
Vincristine (1mg)	£12.71	eMIT[132]
Vincristine (2mg)	£6.48	eMIT[132]
Vincristine (5mg)	£329.50	BNF[135]
Doxorubicin (10mg)	£2.83	eMIT[132]
Doxorubicin (50mg)	£7.09	eMIT[132]
Doxorubicin (200mg)	£20.02	eMIT[132]
Cyclophosphamide (500mg)	£8.23	eMIT[132]

© Roche (2022). All rights reserved

Cyclophosphamide (1,000mg)	£13.55	eMIT[132]
Cyclophosphamide (2,000mg)	£27.50	eMIT[132]

Table 44. Weekly treatment costs for post-discontinuation

Treatment	Total cost	Comments
R^2	£1,874.16	Average cost of R^2
R-Chemo	£860.25	Mean cost of R-B, R-CHOP, R-CVP
Other (non-rituximab containing chemo)	£823.29	Mean cost of O-B, bendamustine
Palliative care	£0.00	Assumed £0 (end of life cost is applied elsewhere)
Trials	£0.00	Assumed £0 (funded by sponsors)

Table 45. Total post-discontinuation costs

Treatment	Total cost
Mosunetuzumab	£40,570.93
Rituximab + Bendamustine	£40,570.93
Obinutuzumab + Bendamustine	£40,570.93
Rituximab + Lenalidomide	£40,570.93

Table 46. Proportion assumed to take each subsequent therapy by arm

Therapy class	Therapy	Mean duration in weeks	% on Mosun	% on RB	% on OB	% on RL
R^2	R^2	52.00	35%	35%	35%	35%
R-Chemo	R-Chemo	21.73	25%	25%	25%	25%
Other (non-rituximab chemo)	Other (non-rituximab chemo)	21.73	10%	10%	10%	10%
Palliative care	Palliative care	21.73	10%	10%	10%	10%
Trials	Trials	21,73	20%	20%	20%	20%

B.3.5.5 Supportive and terminal care costs

Supportive care costs were applied to each model cycle a patient was alive. These costs were different between the progression-free survival and post-progression health states and were independent of treatment arm (Table 47 and

© Roche (2022). All rights reserved

Table 48). They are therefore considered to represent health care resource use that is specific to disease status rather than treatment arm.

A microcosting approach to supportive care costs was taken to determining the resources used in supportive care for each health state or event. Resource use for PFS was extracted from TA243[31] and discussed with clinicians who felt that there may be some overestimation of resource use and suggested alternative values instead (those used in this submission). These resource estimates were then costed using NHS reference costs or appropriate inflation.

In the progressed health state, resource use was extracted from TA604[27] and discussed with clinicians who felt this resource use was no longer representative usage in their experience. Clinicians provided alternative estimates based on their experiences and the resources used were costed using NHS references costs or appropriate inflation.

The disease progression event was assumed to be as reported in TA604[27], which was validated by clinicians as being an appropriate representation of care. A similar approach was taken with terminal care costs, which were inflated to the current year.

The costs applied for supportive care are reported in Table 47 to Table 50. Table 47 and

© Roche (2022). All rights reserved

Table 48 show supportive care costs associated with the PFS and PPS health state respectively. The totals for each health state or event are shown in Table 49. Table 50 shows the one-off costs associated with disease progression. This one-off cost was applied in the cycle that progression takes place.

© Roche (2022). All rights reserved

Table 47. Supportive care costs associated with progression-free state

Component	Resource use per month	Resource use per week	% patients	Unit Cost	Total cost	Cost Source
Haematologist led	1	0.23	100%	£171.18	£39.35	NHS cost collection 2019/20 CL WF01A: 303 (Clinical haematology, non-admitted face-to-face attendance follow-up).[128]
Full blood count	1	0.23	100%	£7.20	£1.66	TA243[31] inflated using PSSRU 2021
Patient history/physical exam	1	0.23	100%	£7.12	£1.64	
Full profile (U&E, LFT, calcium)	1	0.23	100%	£19.60	£4.51	
Serum IgG, IgA, IgM and electrophoresis	1	0.23	100%	£28.77	£6.61	
LDH test	1	0.23	100%	£14.55	£3.34	
CT scans	0.08	0.01	100%	£115.21	£2.21	NHS cost collection 19/20: RD27Z computerised tomography scan of more than three areas[128]
Abbreviations: U&E, urea and electrolytes; LFT, liver function tests; Ig, immunoglobin; LDH, lactate dehydrogenase, CT, computed tomography						

© Roche (2022). All rights reserved

Table 48. Supportive care costs associated with progression state

Component	Frequency per year	Frequency per week	% patients	Unit cost	Total cost	Cost Source
Haematologist led	1	0.23	100%	£171.18	£39.35	NHS cost collection 2019/20 CL WF01A: 303 (Clinical haematology, non-admitted face-to-face attendance follow-up)[128]
Full blood count	1	0.23	100%	£7.20	£1.66	TA243[31] inflated using PSSRU 2021
Patient history/physical exam	1	0.23	100%	£7.12	£1.64	
Full profile (U&E, LFT, calcium)	1	0.23	100%	£19.60	£4.51	
Serum IgG, IgA, IgM and electrophoresis	1	0.23	100%	£28.77	£6.61	
LDH test	0	0	0%	£0	£0	NHS cost collection 19/20: RD27Z computerised tomography scan of more than three areas[128]
Abbreviations: U&E, urea and electrolytes; LFT, liver function tests; Ig, immunoglobin; LDH, lactate dehydrogenase, CT, computed tomography						

© Roche (2022). All rights reserved

Table 49. Supportive care cost

Component	Mean cost used in the model (£)
Progression-free state	59.32
Progression state	57.11
One off progression cost	220.78

Table 50. Supportive care one-off costs associated with progression state

Component	% Patients	Unit cost	Total Cost
Radiological assessments	100	£77.31	£77.31
Biopsy	25	£563.62	£140.91
Blood test	100	£2.56	£2.56

At the time of death, a one-off terminal care cost was applied (Table 51). This cost was sourced from the King's Fund report 2008[136] and inflated to cost year 2021. The cost reflects additional resources that a patient may require in the last months and weeks of life, specifically palliative care and additional care.

Table 51. Terminal care cost

Component	Mean cost used in the model (£)	Source
Terminal care cost	£6,707.78	King's Fund Report[136]

B.3.5.6 Adverse reaction unit costs and resource use

The costs of AEs during the time on treatment were calculated based on the average number of treatment-related AEs per patient per week in the relevant trial (Section B.3.3.6) and the unit cost of these AEs (Table 52). Costs were assumed in line with relevant recent technology appraisals and costed using the most recent reference costs. All costs and assumed treatment were validated by clinicians.

Table 52. Costs of AEs included in the model

Event	Mean cost used in the model (£)	Source*
Alanine aminotransferase increased	£125.44	NHS Reference Costs 2019/20: WF01A - Non-Admitted Face-to-Face Attendance, Follow-up
Anaemia	£3,674.74	NHS Reference Costs 2019/20: weighted average of SA03G to SA03H, NEL
Aspartate aminotransferase increased	£125.44	NHS Reference Costs 2019/20: WF01A - Non-Admitted Face-to-Face Attendance, Follow-up

© Roche (2022). All rights reserved

Cytokine release syndrome	£1,619.80	Hospitalisation costs: NHS Reference Costs 2019/20: weighted average of medical adult patients (unspecified specialty) XC01Z to XC07Z, Critical Care, multiplied by mean days hospitalised for G3-4 CRS in GO29781 (median duration 6 days as per trial)
	£102.40	Tocilizumab: List price (BNF)[137] multiplied by average number of tocilizumab doses administered in patients with G3-4 CRS in GO29781 (8mg/kg per dose, weighted average of 1.5)
	£11,050.03	Total cost for CRS event (assuming [redacted] days hospitalised and weighted average doses of tocilizumab received ([redacted]) based on mean weight of cohort (81kg), requiring [redacted] vials)
Hypokalaemia	£356.81	NHS Reference Costs 2019/20: Weighted average of KC05H to KC05N, day case
Hypophosphatemia	£348.17	NHS Reference Costs 2019/20: KC05N - Fluid or Electrolyte Disorders, without Interventions, with CC Score 0-1, NES
Lymphopenia	£457.41	NHS Reference Costs 2019/20: Weighted average of SA08G to SA08J, day case
Neutropenia	£2,042.93	NHS Reference Costs 2019/20: SA08J - Other Haematological or Splenic Disorders, with CC Score 0-2, NEL
Neutrophil count decreased	£2,042.93	NHS Reference Costs 2019/20: SA08J - Other Haematological or Splenic Disorders, with CC Score 0-2, NEL
Rash erythematous	£441.46	NHS Reference Costs 2019/20: Weighted average of JD07E-K, NES
Thrombocytopenia	£3,573.86	NHS Reference Costs 2019/20: Weighted average of SA12G to SA12K, NEL
Tumour flare	£0.00	No costs assumed
Upper respiratory tract infection	£1,794.53	NHS Reference Costs 2019/20: Weighted average of WH07A to WH07G across NEL, NES, DC
Diarrhoea	£1,685.34	NHS Reference Costs 2019/20: FD01J - Gastrointestinal Infections without Interventions, with CC Score 0-1, NEL
Febrile neutropenia	£6,933.22	NICE guidelines NG52; Appendix A (A.3.2.3); inflated to 2019/20 based on PSSRU
Infections	£1,794.53	NHS Reference Costs 2019/20: Weighted average of WH07A to WH07G across NEL, NES, DC
Vomiting	£1,922.00	NHS Reference costs 19/20: FD10M -Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2, NEL
Infections	£1,922.44	NHS Reference costs 19/20: FD10M -Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2, NEL
Leukopenia	£4,205.55	NHS Reference Costs 2019/20: SA31E - Malignant Lymphoma, including Hodgkin's and Non-Hodgkin's, with CC Score 2-3, NEL
Cutaneous reaction	£612.78	NHS Reference Costs: 2019/20: SA31E - Malignant Lymphoma, including Hodgkin's and Non-Hodgkin's, with CC Score 2-3, NES
*Please see reference 128 for NHS Reference costs 2019/2020		

© Roche (2022). All rights reserved

The probability of events was combined with the cost of each AE in each treatment arm (see Table 53). These costs were then applied in the model to the proportion who remain on treatment in each cycle.

Table 53. Adverse event costs per cycle

Drug regimen	Cost per model cycle (weekly) (£)
Mosunetuzumab	17.34
Rituximab and Bendamustine	18.53
Obinutuzumab and Bendamustine	15.71
Rituximab and Lenalidomide	13.66

B.3.5.7 Miscellaneous unit costs and resource use

No additional costs were considered in this analysis.

B.3.6 Severity

This indication is not expected to require any adjustments to the value of a QALY in line with the NICE Methods Manual. As can be seen in Table 54 none of the analysis are expected to meet the thresholds detailed by NICE.

Table 54: QALY shortfall analysis

Expected total QALYs for the general population	Assumed current treatment	Total QALYs expected for people living with the condition, under current treatment	Absolute QALY shortfall	Proportional QALY shortfall
12.34	RB	6.27	6.07	49.21%
	OB	6.19	6.15	49.85%
	R^2	7.63	4.71	38.19%

B.3.7 Uncertainty

Due to data sparsity and immaturity, there is some uncertainty regarding the efficacy estimates included within the economic model. Data sparsity and immaturity are common obstacles in indications where there are small patient numbers and this situation highlights the requirement for treatments that provide alternative options for patients.

Though all analyses were conducted in line with NICE DSU recommended methods, the underlying populations informing the ITC were not perfectly matched. Further, the relaxation of inclusion criteria for patients included in the ITC biased against mosunetuzumab in the

© Roche (2022). All rights reserved

estimates of outcomes as many of the patient populations in the comparator populations were not as treatment experienced as those included in the GO29781 trial.

When results of the ITC were shown to clinicians at a Clinical Advisory Board (Data on File), all advisors commented that the comparisons from the ITC, for R^2 in particular, were not what they would expect to see in practice and they were concerned that it did not reflect the clinical benefit observed with mosunetuzumab in the GO29781 study. The advisors noted that there were still imbalances between the matched populations, as described in Sections B.2.9 and B.3.3.2.1, which would bias towards the comparators. Consequently, the ITC presented is likely to underestimate the clinical benefit that mosunetuzumab offers over the comparators specified in the scope, which is important to consider when interpreting the cost-effectiveness estimates. It is important to acknowledge however that the robustness of the ITC may improve as further data cuts become available and more events for the endpoints of interest are observed.

Given the limitations and immaturity of the current evidence base, a further non-interventional study is planned (0) which aims to increase certainty around the clinical estimates that have been included in this analysis. Collecting comparative data is challenging in indications with small patient numbers and due consideration must be given to patient access while data collection is ongoing.

B.3.8 Managed access proposal

The immaturity of GO29781 data at the time of submission and the lack of comparator in that trial warrant the need for further data collection in the framework of the Cancer Drugs Fund. The most appropriate source of future data will come from future analyses of the GO29781 pivotal cohort, with follow up planned to continue until [redacted] and [redacted] 55

It is proposed that the multi-national, prospective non-interventional study will be supported by retrospective analyses of mosunetuzumab use in the Haematological Malignancy Research Network (HMRN) and Systemic Anti-Cancer Therapy (SACT) databases. The Company currently has an agreement with the University of York for retrospective analyses of

© Roche (2022). All rights reserved

previous treatment patterns and outcomes in follicular lymphoma and it plans to update the agreement to perform retrospective analyses in the managed access data collection plan.

Table 55. Future Data Collection: proposed, prospective, multi-national NIS

Study	[redacted]
Study design	[redacted]
Population	[redacted]
Intervention(s)	[redacted]
Comparator(s)	[redacted]
Outcomes	[redacted]
Indicate if study used in the NICE economic model	[redacted]
Trial start date	[redacted]
Data cut submitted to NICE	[redacted]
Anticipated data cut after a period of managed access	[redacted]

Table 56. Future data collection: HMRN registry

Registry	Haematological Malignancy Research Network (HMRN)
Type of registry	Ongoing population-based cohort
Population	All patients who use systemic-anti cancer therapies across all NHS England trusts
Relevant data items collected	The HMRN region covers the former two adjacent UK Cancer Networks with a total population of 3.8 million (Yorkshire and the Humber & Yorkshire Coast Cancer Networks) and collects detailed information about all haematological malignancies diagnosed in the region
Data analysis	The company will not have access to the patient data, but will receive de-personalised summary data
Governance	All necessary governance arrangements through agreement with the University of York and the local Clinical Haematology Network.
Indicate if registry previously used within a NICE managed access	Yes

© Roche (2022). All rights reserved

Table 57. Future data collection: SACT database

Registry	Systemic Anti-Cancer Therapy (SACT)
Type of registry	Mandated dataset as part of the Health and Social Care Information Standards
Population	All patients who use systemic-anti cancer therapies across all NHS England trusts
Relevant data items collected	Time to progression Treatment duration Overall survival
Data analysis	The company will not have access to the NHS Digital patient data, but will receive de-personalised summary data
Governance	All necessary governance arrangements through SACT, and other datasets brought together by NHS Digital, have been established with NHS Trusts and NHSE&I.
Indicate if registry previously used within a NICE managed access	Yes

B.3.9 Summary of base-case analysis inputs and assumptions

B.3.9.1 Summary of base-case analysis inputs

A summary of all values, and their respective distributions, used in the base case analysis is presented in Table 58

Table 58. Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution	Section
Baseline parameters			
Baseline parameters	Table 23	None	B.3.2.2
Survival and progression functions			
PFS – R-Len	Table 25	Distribution specific	B.3.3.2.1.1
OS – R-Len	Table 26	Distribution specific	B.3.3.2.1.2
PFS - RB	Table 27	Distribution specific	B.3.3.2.2.1
OS - RB	Table 28	Distribution specific	B.3.3.2.2.2
PFS – OB	Table 29	Distribution specific	B.3.3.2.3.1
OS - OB	Table 30	Distribution specific	B.3.3.2.3.2**Error! Reference source not found.**
All-cause mortality	**Error! Reference**	None	B.3.3.4

© Roche (2022). All rights reserved

	source not found.		
TTOT	Table 31	Distribution specific for mosunetuzumab. Fixed or related to PFS for comparators.	B.3.3.5
Clinical parameters			
Adverse event rates	Table 32	Normal	B.3.3.6
Utilities			
GO29781 – utility estimates	Table 34	Beta	B.3.4.2.1
Brazier age-adjusted coefficients	Table 35	None	B.3.4.2.1**Error! Reference source not found.**
Costs			
Mosunetuzumab – dosing and acquisition	Table 37	None	B.3.5.2.1
Mosunetuzumab - administration costs	Table 38	Generalised gamma	B.3.5.2.1
Comparators – dosing and acquisition	Table 40	None	B.3.5.3
Comparators - costs per cycle	Table 41	None	B.3.5.3
Comparators - administration costs	Table 42	None	B.3.5.3
Post discontinuation therapies – dosing and acquisition	Table 43	None	B.3.5.4
Post discontinuation - weekly treatment costs	Table 44	None	B.3.5.4
Post-discontinuation costs	Table 45	Generalised gamma	B.3.5.4**Error! Reference source not found.**
Proportion assumed to take subsequent therapy	Table 46	None	B.3.5.4**Error! Reference source not found.**
Supportive care costs - PFS	Table 47	None	B.3.5.5**Error! Reference source not found.**
Supportive care costs – progression state	Table 48	None	B.3.5.5**Error! Reference source not found.**
Supportive care costs	Table 49	Generalised gamma	B.3.5.5**Error! Reference source not found.**
Supportive care one-off costs	Table 50	Generalised gamma	B.3.5.5**Error! Reference source not found.**

© Roche (2022). All rights reserved

Terminal care costs	Table 51	Generalised gamma	B.3.5.5**Error! Reference source not found.**
Adverse event costs	Table 52	Log Normal	B.3.5.6**Error! Reference source not found.**

B.3.9.2 Assumptions

During the construction of economic models, it is necessary to make some assumptions, both structural and related to model inputs. The assumptions underlying the economic model presented in this submission (Table 59) were tested, where possible, in the sensitivity analyses described in Section B.3.11.

Table 59. Summary of model assumptions

Topic	Assumption	Justification/reason
ITC	In using ITC methods, it is assumed that there is sufficient overlap between trial populations	Data sparsity is a considerable problem in indications where there are low patient numbers. The ITC was conducted in line with recommended methods and population matching was performed with as close a population as reasonable without impacting the viability of estimates though it is acknowledged that there may be some bias against mosunetuzumab in the presented analysis. This was considered unavoidable given the limitations of the available data.
Efficacy	Efficacy generated from the ITC represents the likely comparative estimates that will be realised in practice	Though the efficacy outputs generated by the ITC is considered to bias against mosunetuzumab, the efficacy estimates included in the economic model are considered to be the most robust source of data available at this time.
Treatment effect	No treatment waning applied after treatment cessation.	Treatment waning was not included as the majority of patients taking mosunetuzumab had completed their regimen within the observed period.
Utilities	Same utility values applied to all treatment arms	No evidence was available to suggest that the HRQL experienced by patients on comparator therapies would differ when compared with those taking mosunetuzumab. Further, incidence and type of adverse events experienced are similar between arms.
Dosing	Cheapest combination of vial sizes will be administered	This assumption is in line with the reference case though it is acknowledged that in practice, it may

© Roche (2022). All rights reserved

		be necessary to use more expensive options sometimes.
Vial sharing	No vial sharing is considered	This assumption was validated by clinicians who were interviewed.
Abbreviations: HRQL; health-related quality of life, ITC: indirect treatment comparison		

B.3.10 Base-case results

B.3.10.1 Base-case incremental cost-effectiveness analysis results

Table 60 presents the base case cost-effectiveness results for mosunetuzumab with the proposed PAS discount. Mosunetuzumab is shown to be cost-effective at a £20,000 threshold versus R^2. Mosunetuzumab is shown to be cost saving versus OB with a slight QALY decrement. When compared to R^2 and RB, mosunetuzumab is estimated to be associated with a QALY and LY gain for patients with some additional cost associated.

© Roche (2022). All rights reserved

Table 60. Deterministic Base Case Cost-effectiveness Results with PAS discount

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosun vs R^2								
Mosun	[redacted]	11.07	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	0.71	[redacted]	£16,103	[redacted]
Mosun vs RB								
Mosun	[redacted]	9.92	[redacted]					
RB	[redacted]	8.56	[redacted]	[redacted]	1.35	[redacted]	£51,148	[redacted]
Mosun vs OB								
Mosun	[redacted]	8.30	[redacted]					
OB	[redacted]	8.49	[redacted]	[redacted]	-0.19	[redacted]	£10,397 (Cost saving)	[redacted]

© Roche (2022). All rights reserved

B.3.11 Exploring uncertainty

B.3.11.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed for 1,000 iterations. In each iteration, the model inputs were randomly drawn from the specific distributions, summarised in Table 58.

The median probabilistic incremental costs and QALYs gained from mosunetuzumab with the PAS discount considered for 1,000 iterations are given in Table 61. The pairwise cost-effectiveness acceptability curves are presented in Figure 41. Assuming a WTP threshold of £20,000 per QALY gained, the probability of mosunetuzumab being the most cost-effective treatment option, against R^2 is [redacted] When comparing to RB and OB, at the same cost-effectiveness threshold, mosunetuzumab's likelihood of cost-effectiveness is [redacted] and [redacted] respectively. The incremental results of each iteration in the PSA are displayed in Figure 42. The results from the probabilistic analysis are in line with those of the deterministic analysis in terms of the estimated QALY and LY gains and the estimated incremental costs demonstrating that the deterministic base case results are likely to represent the average experience per person treated with mosunetuzumab.

© Roche (2022). All rights reserved

Table 61. Probabilistic cost-effectiveness results with PAS discount

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R^2							
Mosun	[redacted]	10.57	[redacted]				
R^2	[redacted]	10.16	[redacted]	[redacted]	0.42	[redacted]	£23,342
Mosun vs RB							
Mosun	[redacted]	9.75	[redacted]				
RB	[redacted]	8.31	[redacted]	[redacted]	1.44	[redacted]	£48,609
Mosun vs OB							
Mosun	[redacted]	8.22	[redacted]				
OB	[redacted]	8.47	[redacted]	[redacted]	-0.25	[redacted]	£7,863

© Roche (2022). All rights reserved

Figure 41. Cost-effectiveness acceptability curve

Figure 42. Incremental cost-effectiveness plane

B.3.11.2 Deterministic sensitivity analysis

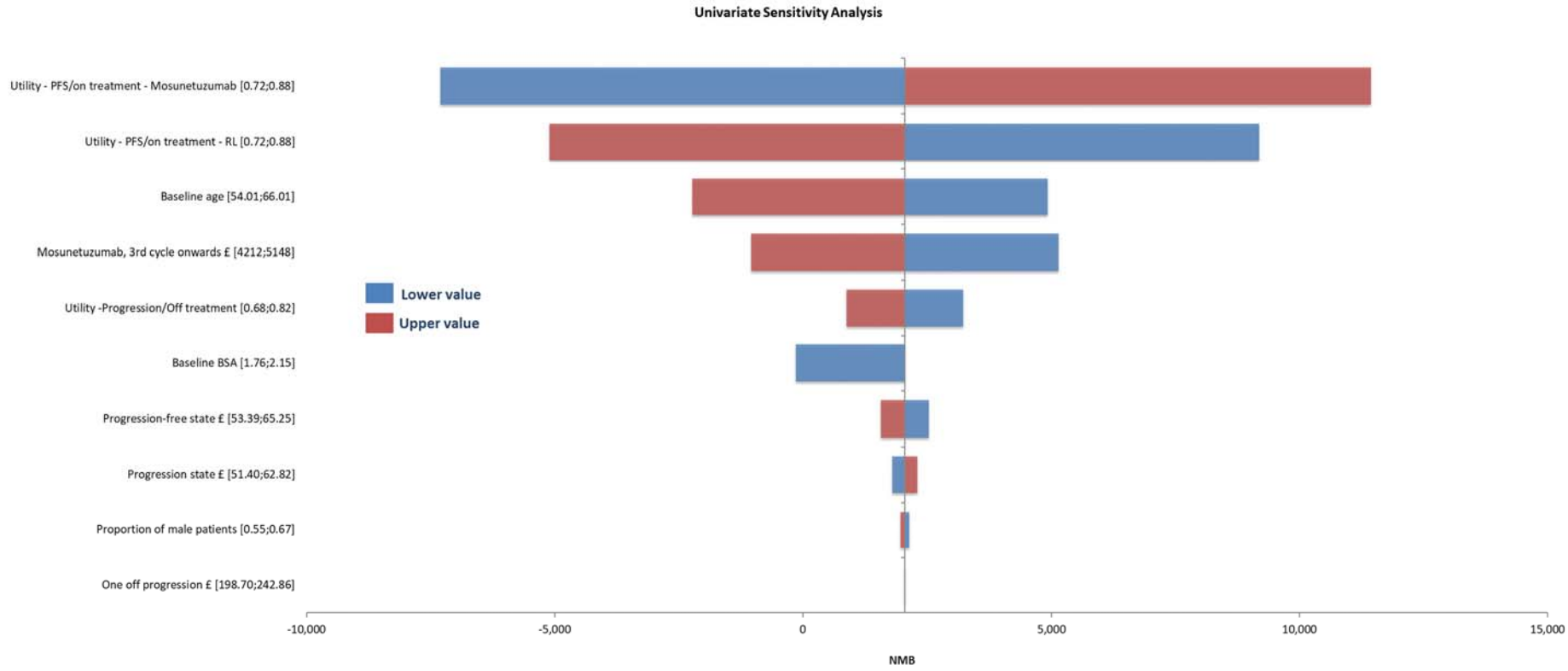
Figure 43 to **Error! Reference source not found.** present the ten most influential parameters on cost-effectiveness with descending sensitivity when mosunetuzumab is compared to R^2, RB and OB, respectively.

The parameters that had the largest impact on the results for mosunetuzumab versus R^2 was utility assigned to the PFS health state for mosunetuzumab for both the intervention and comparator, and the baseline age. The NMB ranged from -£7,309 to £11,438, reflecting the substantial uncertainty around the cost-effectiveness conclusion. Note that the unintuitive result regarding baseline body surface area (BSA) is not erroneous, it is due to the dosing algorithms used and the varying cost per mg for different vial sizes.

A similar pattern was seen in the comparisons to RB and OB. The NMB ranged from -£35,635 to -£24,809 when compared to RB and -£5,748 to £3,025, when compared to OB. Other important parameters were the cost of mosunetuzumab after the loading doses and to a lesser extent, the health state costs. All influential parameters are related to the time spent in each health state, which is a usual driver in partitioned survival modelling and more so where there is uncertainty in the efficacy estimates as these will inherently impact the health state occupancy. As such, to conclude that the value assigned to these states is influential is expected.

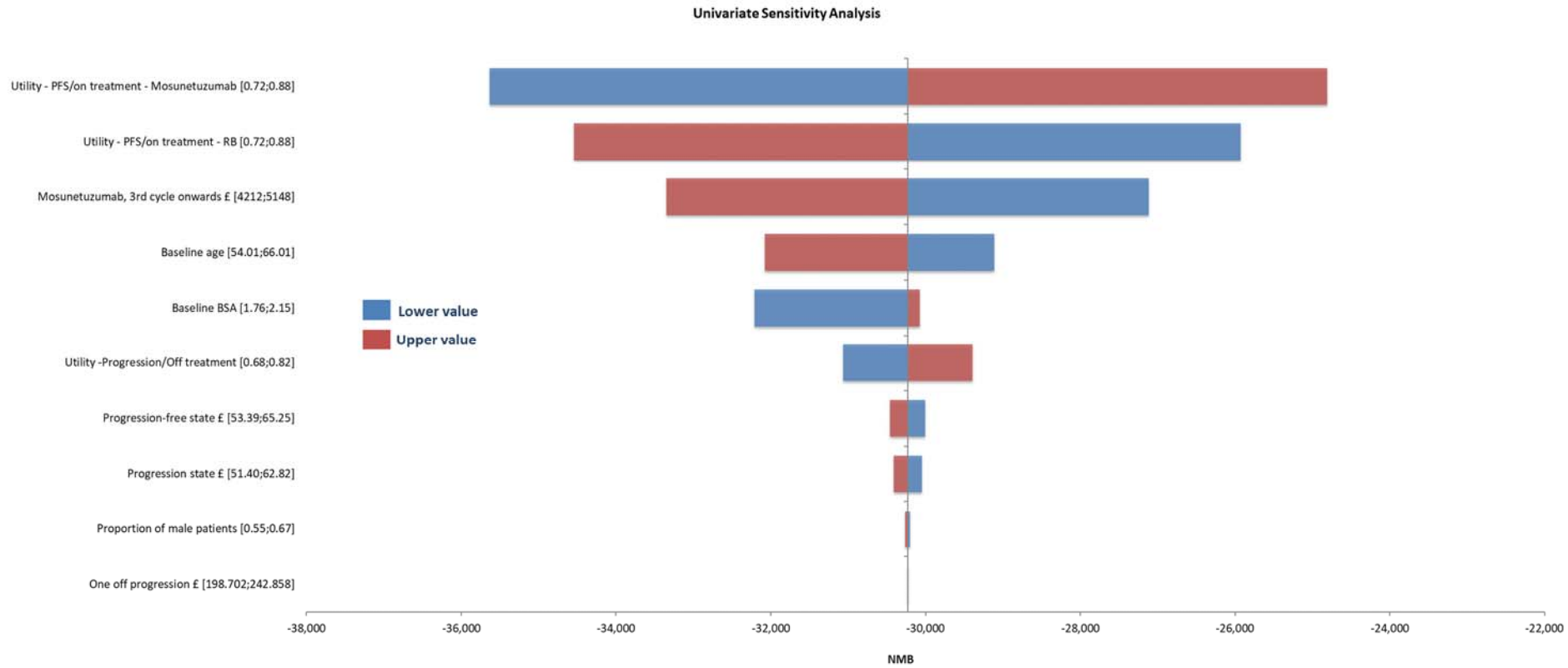
© Roche (2022). All rights reserved

Figure 43. Tornado diagram showing OWSA results on NMB – Mosun vs R^2



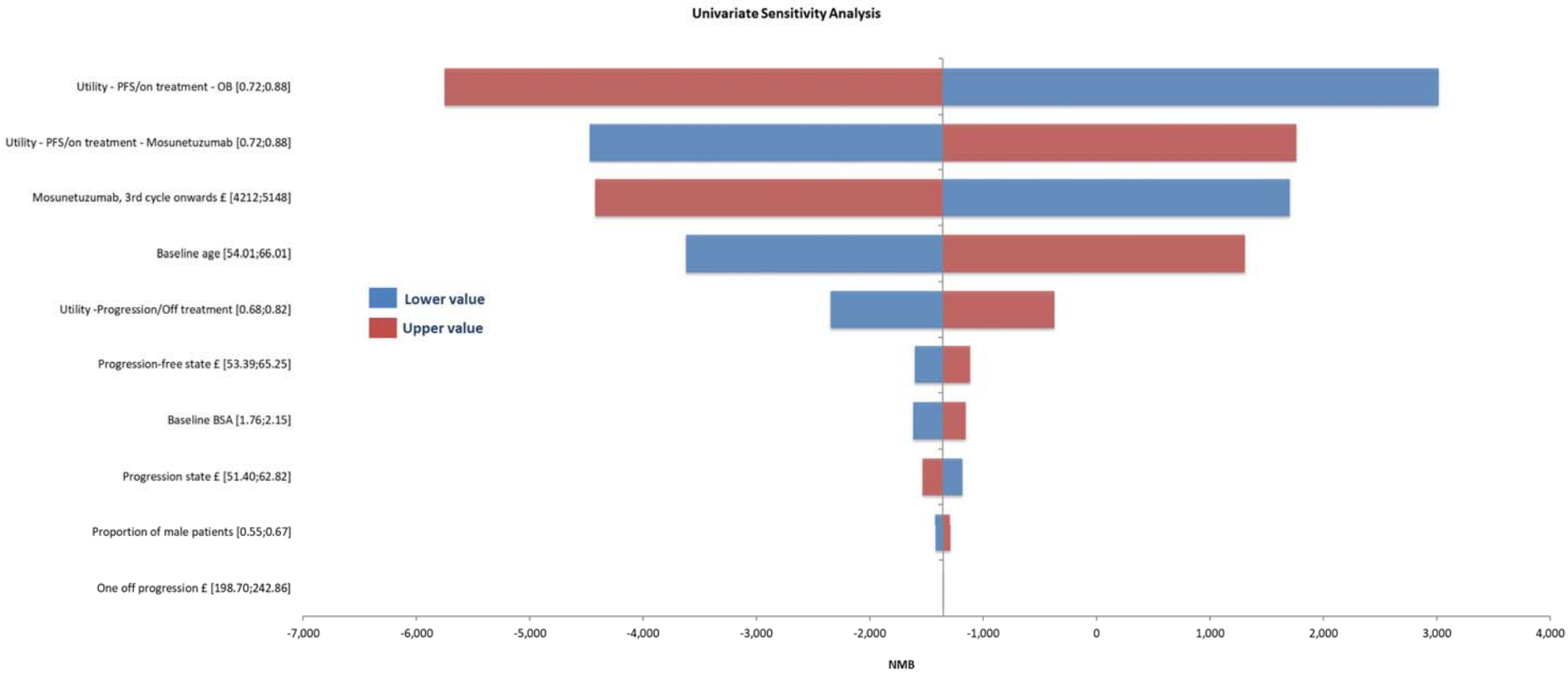
© Roche (2022). All rights reserved

Figure 44. Tornado diagram showing OWSA results on NMB – Mosun vs RB



© Roche (2022). All rights reserved

Figure 45.Tornado diagram showing OWSA results on NMB – Mosun vs OB



© Roche (2022). All rights reserved

Summary of sensitivity analyses results

Sensitivity analysis is of particular importance where data is sparse and there is potential for the decision to be subject to uncertainty. The deterministic sensitivity analysis determines which parameters exert the most influence over these findings. These show that the utility value of living with pre-progressed FL is of particular influence. Similarly, health state costs are shown to be influential. These results are expected because they demonstrate that health state occupancy is the key driver, itself determined by efficacy estimates. Though there is some challenge with limited data in indications with small populations, the probabilistic sensitivity analysis estimates that the deterministic results are likely to be reliable and that mosunetuzumab offers a cost-effective alternative to R^2 and offers cost saving when compared to OB.

B.3.11.3 Scenario analysis

B.3.11.3.1 Parametric Survival Models

The base case parametric models chosen to represent OS and PFS outcomes for each arm were developed and selected in line with NICE DSU guidelines. However, it is acknowledged that for the most part this is subject to each analyst's judgement. This introduces structural uncertainty. As such, the model results are shown when other selections are made to address the impact of these choices.

Table 62 shows that the assumed distribution for mosunetuzumab OS influence the decision though most parametric models would result in mosunetuzumab being considered cost effective vs R^2. Similarly, all but one of the parametric models considered would result in mosunetuzumab being considered cost effective vs R^2, with results versus the other comparators remaining stable (Table 63).

Results are sensitive where alternative models are considered to represent R^2 OS, (Table 64) though all alternative parametric models to represent PFS would result in mosunetuzumab being considered cost effective (Table 65).

Aside from the log models, alternative parametric models to represent RB OS (Table 66) retained similarity with the base case ICER and alternative PFS models exerted almost no change from the base case (Table 67). In most scenarios shown using alternative parametric models for OB OS and PFS (Table 68 and Table 69 respectively), the base case result was stable to these assumptions. Where exponential and Weibull models were used to represent OB OS, mosunetuzumab was estimated to be cost effective.

© Roche (2022). All rights reserved

These analyses demonstrate that assumptions about the structural form of OS events exert considerable influence over the decision problem. This uncertainty can be resolved through further data collection.

© Roche (2022). All rights reserved

Table 62: Scenario Analysis Results: Alternative parametric distributions for Mosunetuzumab OS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£303,719	£4,512 (Cost Saving)	£1,577
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£25,777	£2,939	£25,354
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£18,912	£3,236	£11,054
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£45,115	£5,621 (Cost Saving)	Dominated
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£26,337	£4,286 (Cost Saving)	£2,381

Table 63: Scenario Analysis Results: Alternative parametric distributions for Mosunetuzumab PFS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Log normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£53,901	£12,724 (Cost Saving)	£23,361
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£54,627	£11,220 (Cost Saving)	£18,955
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£48,599	£21,620 (Cost Saving)	£16,060
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,349	£12,649 (Cost Saving)	£16,213
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£54,729	£17,740 (Cost Saving)	£12,442

© Roche (2022). All rights reserved

Table 64: Scenario Analysis Results: Alternative parametric distributions for R^2 OS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	Dominated
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	Dominated
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£768 (Cost Saving)
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	Dominated
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£8,431

Table 65: Scenario Analysis Results: Alternative parametric distributions for R^2 PFS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Log normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£15,891
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£11,975
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£19,774
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£15,464
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397	£14,312

© Roche (2022). All rights reserved

Table 66: Scenario Analysis Results: Alternative parametric distributions for RB OS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£43,902	£10,397	£16,103
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£77,416	£10,397	£16,103
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£10,397	£16,103
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£83,935	£10,397	£16,103
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£214,498	£10,397	£16,103

Table 67: Scenario Analysis Results: Alternative parametric distributions for RB PFS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£49,930	£10,397	£16,103
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£48,990	£10,397	£16,103
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£50,117	£10,397	£16,103
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£50,995	£10,397	£16,103
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£49,589	£10,397	£16,103

© Roche (2022). All rights reserved

Table 68: Scenario Analysis Results: Alternative parametric distributions for OB OS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£2,522	£16,103
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£2,530	£16,103
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£13,318 (Cost Saving)	£16,103
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£42,976 (Cost Saving)	£16,103
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£2,621	£16,103

Table 69: Scenario Analysis Results: Alternative parametric distributions for OB PFS

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	Dominated	£16,103
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£9,748 (Cost Saving)	£16,103
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	Dominated	£16,103
Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£11,119 (Cost Saving)	£16,103
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	Dominated	£16,103

© Roche (2022). All rights reserved

B.3.11.3.2 Proportional Hazards Assumption

As detailed in Section B.3.3.2, the Schoenfeld residuals and log negative log plots were used to make judgement on whether the proportional hazards assumption was suitable for each outcome and arm relative to mosunetuzumab data. The decision was not consistent for all arms and outcomes, but a consistent and conservative approach to not apply the proportional hazards assumption was preferred. The impact of this approach and the assumption of non-proportional hazards is examined in scenarios where the proportional hazards generated by the ITC are used. These are shown in Table 70.

The results of these analyses show that where the proportional hazards assumption is accepted, mosunetuzumab would be considered cost effective versus both OB and RB. As the log negative log plots for R^2 outcomes indicate that it is unlikely the proportional hazards assumption be accepted, even with further data availability, these were not considered a plausible or useful scenario to investigate. Where data is sparse, it can be difficult to reliably assess whether the assumption of proportional hazards holds, though in many cases there was no reason to reject the hypothesis and these scenarios aim to reflect the situation where more data is available. A more robust assessment of the proportional hazards assumption may be possible as more data with longer follow up comes available.

© Roche (2022). All rights reserved

Table 70: Scenario Analysis Results: Proportional hazard assumption accepted

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
PH assumption RB	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£21,790	£10,397	£16,103
PH assumption OB	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£7,534	£16,103

© Roche (2022). All rights reserved

B.3.11.3.3 Source of health state utility estimates

It was considered most appropriate to use HRQL data collected in the GO29781 trial as this represents the group of people most closely aligned with the intended population that mosunetuzumab would be used in. However, GO29781 did not include all relevant comparators and so there is uncertainty as to whether the HRQL values estimated from this group may fully represent the other arms. As such, literature values were sourced and model results where these inform the health state utilities are shown in Table 71.

In all comparisons, the cost-effectiveness results are moderately influenced by choice of health state utility values, but mosunetuzumab remains a cost-effective alternative to R^2 when using the alternative utility values.

© Roche (2022). All rights reserved

Table 71: Scenario Analysis Results: Alternative health state utility sources

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYs	Costs	QALYs	Costs	QALYs	Costs			
GO29781 (Off/On progression)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Wild et al 2006	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£55,309	£6,414 (Cost Saving)	£13,437
Cognet et al 2015	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£64,299	£6,524 (Cost Saving)	£14,518

© Roche (2022). All rights reserved

B.3.11.3.4 Mortality

The impact of using an alternative method to accommodate background mortality is shown in Table 72. This method uses the age distribution of those in the GO29781 to better represent the distribution of patient demographics over time, as detailed by Felizzi, Manevy and Maoi[109].

The results of this analysis demonstrate that when the cohort has the same age distribution as the GO29781 trial cohort, there is a larger incremental QALY gain for mosunetuzumab versus RB and R^2, compared with the base case results. When compared with OB, the incremental QALY loss is larger than is estimated in the base case.

© Roche (2022). All rights reserved

Table 72: Scenario Analysis Results: Alternative methods for background mortality

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYS	Costs	QALYS	Costs	QALYS	Costs			
Average cohort age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
Age distribution from trial	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£46,934	£7,200 (Cost Saving)	£12,583

© Roche (2022). All rights reserved

B.3.11.3.5 Confidential discounts for comparators

Where it is known that confidential discounts are in place for comparators, the NICE user manual for the submission template recommends presenting scenarios with a range of potential discounts to aid decision making. The only comparator anticipated to be subject to a confidential discount is OB which is owned by Roche. As such, the agreed discount is known. A scenario using this discount is shown in Table 73.

These results suggest that mosunetuzumab is dominated when compared with OB with a confidential discount applied. The results of other comparisons are also marginally influenced by the OB confidential discount, on account of this treatment's use as a subsequent therapy in the model but results remain stable with respect to the base case estimates.

© Roche (2022). All rights reserved

Table 73: Scenario Analysis Results: Confidential discount for OB

Base case	inc. vs RB		inc. vs OB		inc. vs RL		ICER vs RB	ICER vs OB	ICER vs R^2
	QALYS	Costs	QALYS	Costs	QALYS	Costs			
OB without PAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,148	£10,397 (Cost Saving)	£16,103
Scenarios									
OB with PAS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£51,105	Dominated	£15,959

© Roche (2022). All rights reserved

B.3.12 Subgroup analysis

No subgroup analysis has been conducted for this decision problem.

B.3.13 Benefits not captured in the QALY calculation

Clinical advice to the company was that there is no accepted standard of care for third line and subsequent FL treatment and that clinical practice in England is highly variable. Patients with FL who are heavily pre-treated and often refractory to multiple available therapies represent a population in which there is a substantial unmet need for novel therapeutic treatment options.

As a humanised anti-CD20/CD3 bispecific antibody, mosunetuzumab offers a unique, first-in-class mechanism of action that differs from the mechanisms of currently available therapies, providing an innovative new treatment option that can overcome resistance to existing therapies while providing acceptable safety and tolerability. Clinical advice to the company indicated that the ORR of 80% and CR rate of 60% observed in the GO29781 pivotal cohort were high, which is a considerable strength, especially that the patients enrolled in this cohort represented a high-risk population.

Currently available data, while immature, positions mosunetuzumab as a step-change in the treatment of FL at third line and beyond. In particular, mosunetuzumab offers a much-needed novel treatment option for heavily pretreated R/R FL patients, particularly those who are refractory to anti-CD20 antibodies and various chemotherapy regimens.

B.3.14 Validation

B.3.14.1 Validation of cost-effectiveness analysis

The model was subject to an internal quality assurance procedure, which included technical validation and cross-validation. Any issues or errors noted in the reviews were documented and addressed in the final version of the models. The technical validation ensured that there were no functional errors in the model calculations and was completed by an analyst who was not involved with the development of the model. Cross-validation involved providing a comparison between the results of the model developed for this submission and other published models in order to increase confidence in the results generated by the model.

© Roche (2022). All rights reserved

B.3.15 Interpretation and conclusions of economic evidence

B.3.15.1 Strengths in the modelling approach

The model uses the available evidence in the most intuitive way and all data has been analysed in line with NICE DSU recommendations. Where ITC populations were not completely aligned, population inclusion criteria were expanded conservatively so that the comparator estimates were not biased towards mosunetuzumab. Estimates from the model have been extensively validated; the outcomes were shown and discussed with clinical experts at an advisory board meeting and structural decisions were supported by analysis of RWE.

Extensive sensitivity analysis has been undertaken which supports the deterministic base case and examines the impact of structural uncertainty, parameter uncertainty and parameter precision. Probabilistic analysis shows that for on average, mosunetuzumab represents a cost-effective option for treatment. The deterministic results indicate that versus R^2, the comparator occupying the largest proportion of the market, mosunetuzumab is cost-effective. When compared to RB, mosunetuzumab would not be considered to be cost-effective, although this comparator was only included as it was not possible to make a robust comparison with rituximab in combination with other chemotherapies. RB itself is estimated to be used by only [redacted] of patients. When compared with OB, mosunetuzumab is estimated to result in marginal QALY decrement (equivalent to approximately 2 months) although this should be framed with consideration of the ITC limitations which are likely to bias in favour of OB. Furthermore, with a market share of just [redacted], the company do not consider OB to be a relevant comparator, with data against this regimen only included for completeness with the NICE scope.

B.3.15.2 Limitations in the modelling approach

As discussed throughout the submission documentation, the primary limitation of the economic model concerns the evidence base and is a feature of collecting data in an indication with a small population. All models require some simplification from real life and any simplifications and assumptions made have been documented. Given the imbalances between matched populations, small sample sizes and bias against mosunetuzumab with the ITCs due to comparator populations including those at earlier lines with better prognosis, the current results are unlikely to provide a true reflection of the relative efficacy of mosunetuzumab. As such, the current cost-effectiveness estimates can be considered to be conservative against mosunetuzumab; further follow up from upcoming data cuts may improve the robustness of these estimates.

© Roche (2022). All rights reserved

B.3.15.3 Conclusions

Despite therapeutic advances in FL, there remains an unmet need, particularly among patients who have received two or more prior lines of systemic therapy given that survival and duration of remission worsen significantly as patients progress through multiple lines of therapy. There is no established standard of care at third and subsequent treatment lines, with therapeutic choices dependent on prior treatments received.

As a first-in-class CD20/CD3 bispecific antibody, mosunetuzumab represents an innovative, alternative rituximab and chemotherapy-free treatment option, particularly for patients whose disease is resistant to multiple agents or who do not tolerate chemotherapy, and who therefore have limited treatment options. Due to the aforementioned limitations of the evidence base, the cost-effectiveness estimates of mosunetuzumab vs the comparators specified in the NICE scope is likely to be underestimated. The robustness of the GO29781 data will improve as further data cuts become available, but overall the current analysis suggest the need for further data collection in the framework of the Cancer Drugs Fund.

© Roche (2022). All rights reserved

B.4 References

1. Public Health England. Cancer survival in England for patients diagnosed between 2014 and 2018, and followed up until 2019. https://www.gov.uk/government/statistics/cancer-survival-in-england-for-patients-diagnosed-between-2014-and-2018-and-followed-up-until-2019
2. van Oers MH KR, Marcus RE, Wolf M, Kimby E, Gascoyne RD, Jack A, Van't Veer M, Vranovsky A, Holte H, van Glabbeke M, Teodorovic I, Rozewicz C, Hagenbeek A. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood*. 2006;108(10):3295-3301.
3. van Oers MH, Van Glabbeke M, Giurgea L, et al. Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. *J Clin Oncol*. Jun 10 2010;28(17):2853-8. doi:10.1200/JCO.2009.26.5827
4. Juárez-Salcedo LM, Conde-Royo D, Quiroz-Cervantes K, Dalia S. Use of anti-CD20 therapy in follicular and marginal zone lymphoma: a review of the literature. *Drugs Context*. 2020;9doi:10.7573/dic.2019-9-3
5. Sun LL, Ellerman D, Mathieu M, et al. Anti-CD20/CD3 T cell-dependent bispecific antibody for the treatment of B cell malignancies. *Sci Transl Med*. May 13 2015;7(287):287ra70. doi:10.1126/scitranslmed.aaa4802
6. Aldoss I, Bargou RC, Nagorsen D, Friberg GR, Baeuerle PA, Forman SJ. Redirecting T cells to eradicate B-cell acute lymphoblastic leukemia: bispecific T-cell engagers and chimeric antigen receptors. *Leukemia*. Apr 2017;31(4):777-787. doi:10.1038/leu.2016.391
7. Dada R. Diagnosis and management of follicular lymphoma: A comprehensive review. *European Journal of Haematology*. 2019;103(3):152-163. doi:https://doi.org/10.1111/ejh.13271
8. Rivas-Delgado A, Magnano L, Moreno-Velázquez M, et al. Response duration and survival shorten after each relapse in patients with follicular lymphoma treated in the rituximab era. *Br J Haematol*. Mar 2019;184(5):753-759. doi:10.1111/bjh.15708
9. Haematological Malignancy Research Network. *Treatment Patterns and Outcomes in Follicular Lymphoma with a Focus on Patients who have Received at least Two Prior Lines of Systemic Therapy*. 8 April 2022.
10. Montoto S, Davies AJ, Matthews J, et al. Risk and Clinical Implications of Transformation of Follicular Lymphoma to Diffuse Large B-Cell Lymphoma. *Journal of Clinical Oncology*. 2007/06/10 2007;25(17):2426-2433. doi:10.1200/JCO.2006.09.3260
11. McNamara C, Montoto S, Eyre TA, et al. The investigation and management of follicular lymphoma. *British Journal of Haematology*. 2020;191(3):363-381. doi:https://doi.org/10.1111/bjh.16872
12. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. Sep 20 2014;32(27):3059-68. doi:10.1200/jco.2013.54.8800
13. Dreyling M, Ghielmini M, Rule S, et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. Mar 2021;32(3):298-308. doi:10.1016/j.annonc.2020.11.008
14. UK CR. Follicular lymphoma. 4 January 2022. https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/follicular-lymphoma
15. Pettengell R, Donatti C, Hoskin P, et al. The impact of follicular lymphoma on health-related quality of life. *Ann Oncol*. Mar 2008;19(3):570-6. doi:10.1093/annonc/mdm543
16. Wang H, Smith A, Yu G, et al. UK Utility Elicitation in Patients with Follicular Lymphoma. Article. *Value in Health*. 10/20 2017;20(9):A449. doi:10.1016/j.jval.2017.08.286
17. Janssen B, Szende A. Population Norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, eds. *Self-Reported Population Health: An International Perspective based on EQ-5D*. Springer Netherlands; 2014:19-30.

© Roche (2022). All rights reserved

18. NHS Digital. Cancer Registration Statistics, England 2019. Cancer incidence data tables. https://digital.nhs.uk/data-and-information/publications/statistical/cancer-registration-statistics/england-2019
19. Haematological Malignancy Research Network. Prevalence. 7 January 2022. https://hmrn.org/statistics/Prevalence
20. NHS Digital. Cancer Registration Statistics, England 2019. Cancer mortality data tables. https://digital.nhs.uk/data-and-information/publications/statistical/cancer-registration-statistics/england-2019
21. Office for National Statistics. Clinical commissioning group population estimates (National Statistics) mid-2019. 4 May 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/clinicalcommissioninggroupmidyearpopulationestimates
22. Casulo C, Byrtek M, Dawson KL, et al. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. *J Clin Oncol*. Aug 10 2015;33(23):2516-22. doi:10.1200/jco.2014.59.7534
23. Casulo C. Prognostic factors in follicular lymphoma: new tools to personalize risk. *Hematology Am Soc Hematol Educ Program*. 2016;2016(1):269-276. doi:10.1182/asheducation-2016.1.269
24. National Institute for Health and Care Excellence. Technology appraisal guidance [TA627]: Lenalidomide with rituximab for previously treated follicular lymphoma. https://www.nice.org.uk/guidance/ta627
25. National Institute for Health and Care Excellence. Technology appraisal guidance [TA629]: Obinutuzumab with bendamustine for treating follicular lymphoma after rituximab. https://www.nice.org.uk/guidance/ta629
26. National Institute for Health and Care Excellence. Technology appraisal guidance [TA137]: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. https://www.nice.org.uk/Guidance/TA137
27. National Institute for Health and Care Excellence. Technology appraisal guidance [TA604]: Idelalisib for treating refractory follicular lymphoma. 15 September 2021. https://www.nice.org.uk/guidance/ta604
28. Brice P, Bastion Y, Lepage E, et al. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the Groupe d'Etude des Lymphomes Folliculaires. Groupe d'Etude des Lymphomes de l'Adulte. *J Clin Oncol*. Mar 1997;15(3):1110-7. doi:10.1200/jco.1997.15.3.1110
29. National Institute for Health and Care Excellence. NICE guideline [NG52]: Non-Hodgkin's lymphoma: diagnosis and management. https://www.nice.org.uk/guidance/ng52/chapter/Recommendations#management-of-follicular-lymphoma
30. Prettyjohns M, Hoskin P, McNamara C, Linch D. The cost-effectiveness of immediate treatment or watch and wait with deferred chemotherapy for advanced asymptomatic follicular lymphoma. *Br J Haematol*. Jan 2018;180(1):52-59. doi:10.1111/bjh.14990
31. National Institute for Health and Care Excellence. Technology appraisal guidance [TA243]: Rituximab for the first-line treatment of stage III-IV follicular lymphoma. https://www.nice.org.uk/guidance/ta243
32. Solal-Céligny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. Sep 1 2004;104(5):1258-65. doi:10.1182/blood-2003-12-4434
33. National Institute for Health and Care Excellence. Technology appraisal guidance [TA513]: Obinutuzumab for untreated advanced follicular lymphoma. https://www.nice.org.uk/guidance/ta513
34. National Institute for Health and Care Excellence. Technology appraisal guidance [TA226]: Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma. https://www.nice.org.uk/guidance/ta226
35. Link BK, Day BM, Zhou X, et al. Second-line and subsequent therapy and outcomes for follicular lymphoma in the United States: data from the observational National LymphoCare Study. *Br J Haematol*. Feb 2019;184(4):660-663. doi:10.1111/bjh.15149

© Roche (2022). All rights reserved

36. Faria R, Hernandez Alava A, M,, Manca A, Wailoo A, J. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data. *Available from http://wwwnicedsuorguk*. 2015;
37. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. Feb 10 2007;25(5):579-86. doi:10.1200/jco.2006.09.2403
38. Dreyling M, Santoro A, Mollica L, et al. Phosphatidylinositol 3-Kinase Inhibition by Copanlisib in Relapsed or Refractory Indolent Lymphoma. *J Clin Oncol*. Dec 10 2017;35(35):3898-3905. doi:10.1200/JCO.2017.75.4648
39. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. Mar 3 1993;85(5):365-76. doi:10.1093/jnci/85.5.365
40. Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual (3rd Edition). European Organisation for Research and Treatment of Cancer. 18 January 2022. https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf
41. Giesinger JM, Kuijpers W, Young T, et al. Thresholds for clinical importance for four key domains of the EORTC QLQ-C30: physical functioning, emotional functioning, fatigue and pain. *Health Qual Life Outcomes*. 2016;14:87-87. doi:10.1186/s12955-016-0489-4
42. Webster K, Cashy D, Cella D. Measuring the quality of life (QOL) in patients with non-Hodgkin's lymphoma (NHL): the functional assessment of cancer therapy. 01/01 2005;14
43. EuroQoL. EQ-5D-5L | About. 18 January 2022. https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/
44. Dreyling M, Morschhauser F, Bouabdallah K, et al. Phase II study of copanlisib, a PI3K inhibitor, in relapsed or refractory, indolent or aggressive lymphoma. *Ann Oncol*. Sep 1 2017;28(9):2169-2178. doi:10.1093/annonc/mdx289
45. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. Jan 1998;16(1):139-44. doi:10.1200/jco.1998.16.1.139
46. Carter GC, Liepa AM, Zimmermann AH, Morschhauser F. Validation of the Functional Assessment of Cancer Therapy–Lymphoma (FACT-LYM) in Patients with Relapsed/Refractory Mantle Cell Lymphoma. *Blood*. 2008;112(11):2376-2376. doi:10.1182/blood.V112.11.2376.2376
47. Hlubocky FJ, Webster K, Cashy J, Beaumont J, Cella D. The Development and Validation of a Measure of Health-Related Quality of Life for Non-Hodgkin's Lymphoma: The Functional Assessment of Cancer Therapy—Lymphoma (FACT-Lym). *Lymphoma*. 2013/02/20 2013;2013:147176. doi:10.1155/2013/147176
48. 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*. Jun 1998;52(6):377-84. doi:10.1136/jech.52.6.377
49. Leonard JP, Trneny M, Izutsu K, et al. AUGMENT: A Phase III Study of Lenalidomide Plus Rituximab Versus Placebo Plus Rituximab in Relapsed or Refractory Indolent Lymphoma. Clinical Trial, Phase III Journal Article Multicenter Study Randomized Controlled Trial Research Support, Non-U S Gov't. *J Clin Oncol*. May 10 2019;37(14):1188-1199. doi:10.1200/JCO.19.00010
50. Sehn LH, Chua N, Mayer J, et al. Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN): a randomised, controlled, open-label, multicentre, phase 3 trial. *Lancet Oncol*. Aug 2016;17(8):1081-1093. doi:10.1016/S1470-2045(16)30097-3
51. Zinzani PL, Flinn IW, Yuen SLS, et al. Venetoclax-rituximab with or without bendamustine vs bendamustine-rituximab in relapsed/refractory follicular lymphoma. *Blood*. Dec 3 2020;136(23):2628-2637. doi:10.1182/blood.2020005588
52. Sehn LH, Kamdar M, Herrera AF, et al. Randomized phase 2 trial of polatuzumab vedotin (pola) with bendamustine and rituximab (BR) in relapsed/refractory (r/r) FL and DLBCL. *Journal of Clinical Oncology*. 2018;36(15_suppl):7507-7507. doi:10.1200/JCO.2018.36.15_suppl.7507
53. Andorsky DJ, Coleman M, Yacoub A, et al. MAGNIFY phase IIIb interim analysis of induction R2 followed by maintenance in relapsed/refractory indolent NHL. *Journal of Clinical Oncology*. 2020;38(15_suppl):8046-8046. doi:10.1200/JCO.2020.38.15_suppl.8046

© Roche (2022). All rights reserved

54. Lee DW, Gardner R, Porter DL, et al. Current concepts in the diagnosis and management of cytokine release syndrome. *Blood*. Jul 10 2014;124(2):188-95. doi:10.1182/blood-2014-05-552729
55. Lee DW, Santomasso BD, Locke FL, et al. ASTCT Consensus Grading for Cytokine Release Syndrome and Neurologic Toxicity Associated with Immune Effector Cells. *Biol Blood Marrow Transplant*. Apr 2019;25(4):625-638. doi:10.1016/j.bbmt.2018.12.758
56. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. Accessed 8 February 2022, https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation
57. Bertwistle D, Munakata J, Wehler E, et al. The Cost-Effectiveness of Bendamustine-Rituximab Versus Fludarabine-Rituximab for Patients with Indolent Non-Hodgkin’s Lymphoma (INHL) Who Have Progressed Following Treatment with Rituximab or a Rituximab-Containing Regimen in Colombia. *Value in Health*. 2013;16(7):A408. doi:10.1016/j.jval.2013.08.492
58. Bertwistle D, Munakata J, Wehler E, et al. The Cost-Effectiveness of Bendamustine-Rituximab Versus Fludarabine-Rituximab for Patients with Indolent Non-Hodgkin's Lymphoma (INHL) Who Have Progressed Following Treatment with Rituximab or a Rituximab-Containing Regimen in Mexico. *Value in Health*. 2013;16(7):A411. doi:10.1016/j.jval.2013.08.508
59. Desanvicente-Celis Z, Obando CA, Chaves M, Gonzalez L, Muschett D. Cost-Effectiveness Analysis of Bendamustine-Rituximab Treatment Compared with Fludarabine-Rituximab Treatment, in Patients with Indolent Non-Hodgkin's Lymphoma in Costa Rica. *Value in Health*. 2014;17(3):A85. doi:10.1016/j.jval.2014.03.497
60. Desanvicente-Celis Z, Obando CA, Rodriguez E, Gonzalez L, Muschett D. Cost-Effectiveness Analysis of Bendamustine-Rituximab Treatment Compared with Fludarabine-Rituximab Treatment, in Patients with Indolent Non-Hodgkin's Lymphoma in Panama. *Value in Health*. 2014;17(3):A83. doi:10.1016/j.jval.2014.03.484
61. Erdogan-Ciftci E, Yanik L, Akyol Ersoy B, Krivasi T. PCN233 COST EFFECTIVENESS OF OBINUTUZUMAB PLUS BENDAMUSTINE VS BENDAMUSTINE ALONE FOR PATIENTS WHO DID NOT RESPOND OR PROGRESSED DURING OR AFTER RITUXIMAB OR A RITUXIMAB CONTAINING REGIMENT IN TURKEY. *Value in Health*. 2019;22:S481. doi:10.1016/j.jval.2019.09.428
62. Guzauskas GF, Masaquel A, Reyes C, Bernaards C, Krivasi T, Veenstra DL. Cost-effectiveness of obinutuzumab plus bendamustine followed by obinutuzumab monotherapy for the treatment of follicular lymphoma patients who relapse after or are refractory to a rituximab-containing regimen in the US. *J Med Econ*. Oct 2018;21(10):960-967. doi:10.1080/13696998.2018.1489254
63. Haukaas FS, Ohna A, Krivasi T. Cost-Effectiveness of Obinutuzumab in Combination with Bendamustine Followed by Obinutuzumab Maintenance versus Bendamustine Alone in Treatment of Patients with Rituximab-Refractory Follicular Lymphoma in Norway. *Appl Health Econ Health Policy*. Aug 2018;16(4):569-577. doi:10.1007/s40258-018-0401-y
64. Lachaine J, Beauchemin C, Mathurin K, Aissa F. Cost-effectiveness of bendamustine versus ibritumomab tiuxetan in the treatment of rituximab-refractory indolent non-Hodgkin's lymphoma in Canada. *Value Health*. 2013a;16(3):A141.
65. Lachaine J, Beauchemin C, Mathurin K, Aissa F. Cost-effectiveness of bendamustine +rituximab versus fludarabine +rituximab in the treatment of relapsed indolent non-Hodgkin's and mantle cell lymphomas in Canada. *Value Health*. 2013b;16(3):A141.
66. Salazar A. Economic evaluation of bendamustine in patients with relapsed indolent non-Hodgkin's lymphoma in Mexico. *Value Health*. 2017;20(5):A215.
67. Soini EJO, Martikainen JA, Nousiainen T. Treatment of follicular non-Hodgkin's lymphoma with or without rituximab: cost-effectiveness and value of information based on a 5-year follow-up. *Ann Oncol*. May 2011;22(5):1189-1197. doi:10.1093/annonc/mdq582
68. Soini EJ, Martikainen JA, Vihervaara V, Mustonen K, Nousiainen T. Economic evaluation of sequential treatments for follicular non-hodgkin lymphoma. *Clinical therapeutics*. Apr 2012;34(4):915-925.e2. doi:10.1016/j.clinthera.2012.02.019
69. Sweetenham J, Hieke K, Kerrigan M, et al. Cost-minimization analysis of CHOP, fludarabine and rituximab for the treatment of relapsed indolent B-cell non-Hodgkin's lymphoma in the U.K. *Br J Haematol*. Jul 1999;106(1):47-54. doi:10.1046/j.1365-2141.1999.01515.x

© Roche (2022). All rights reserved

70. Thielen FW, Kersten MJ, Kuizenga P, et al. Cost-effectiveness of lenalidomide plus rituximab versus rituximab monotherapy in patients with previously treated follicular lymphoma: a societal view. *Expert review of anticancer therapy*. Sep 2 2021:1-12. doi:10.1080/14737140.2021.1971520
71. Thompson S, Erasmus MC. Cost-effectiveness of 90Y-ibritumomab tiuxetan (90Y-Zevalin) versus rituximab monotherapy in patients with relapsed follicular lymphoma. *Blood* 2005;106(11)
72. Vandekerckhove S, Lamotte M, Iannazzo S, Chirita O. Cost utility analysis of bortezomib in a biomarker positive subgroup of relapsed and refractory folicular lymphoma. *Value Health*. 2012;15(7):A427.
73. Vijenthira A, Kuruvilla J, Prica A. Cost-effectiveness analysis of allogeneic versus autologous stem cell transplant versus chemo-immunotherapy for early relapse of follicular lymphoma within 2 years of initial therapy. *Bone marrow transplantation*. May 13 2021;doi:10.1038/s41409-021-01327-5
74. Wehler E, Bilir SP, Bertwistle D, Leyva V, Munakata J. Validating an indolent non-hodgkin's lymphoma (NHL) cost-effectiveness analysis model using two software tools: Key implementation considerations and results. *Value Health*. 2013;16(7):A589.
75. Zhang PF, Xie D, Wen F, Li Q. Lenalidomide plus rituximab Vs rituximab alone in relapsed or refractory indolent lymphoma: A cost-effectiveness analysis. *Cancer medicine*. Aug 2020;9(15):5312-5319. doi:10.1002/cam4.3121
76. Scottish Medicines Consortium (SMC). SMC2281 Lenalidomide in combination with rituximab (anti-CD20 antibody) for the treatment of adult patients with previously treated FL (grade I-IIIa), October 2020 [online]. Available at: https://www.scottishmedicines.org.uk/medicines-advice/lenalidomide-revlimid-full-smc2281/ [Last accessed 15/09/21].
77. Scottish Medicines Consortium (SMC). 1219/17 Obinutuzumab in combination with bendamustine for the treatment of patients with FL who did not respond or who progressed during or up to six months after treatment with rituximab or a rituximab-containing regimen, March 2017 [online]. Available at: https://www.scottishmedicines.org.uk/medicines-advice/obinutuzumab-gazyvaro-fullsubmission-121917/ [Last accessed 15/09/21].
78. Scottish Medicines Consortium (SMC). 1039/15 Idelalisib monotherapy for the treatment of adult patients with FL that is refractory to two prior lines of treatment, May 2015 [online]. Available at: https://www.scottishmedicines.org.uk/medicines-advice/idelalisib-zydelig-fullsubmission-103915/ [Last accessed 15/09/21].
79. Scottish Medicines Consortium (SMC). 171/05 Ibritumomab tiuxetan (Zevalin®) for the treatment of adult patients with rituximab relapsed or refractory CD20+ follicular B-cell NHL, resubmission, July 2007 [online]. Available at: https://www.scottishmedicines.org.uk/medicines-advice/ibritumomab-tiuxetan-zevalin-resubmission-17105/ [Last accessed 15/09/21].
80. Pharmaceutical Benefits Advisory Committee (PBAC). Idelalisib for iNHL in patients with progressive disease despite previous treatment, PSD March 2015 [online]. Available at: https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2015-03/Files/idelalisib-2-psd-march-2015.pdf [Last accessed 15/09/21].
81. Pharmaceutical Benefits Advisory Committee (PBAC). Idelalisib for FL that is refractory to both rituximab and an alkylating agent, resubmission, PSD November 2015 [online]. Available at: https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2015-11/files/idelalisib-follicular-lymphoma-psd-november-2015.pdf [Last accessed 15/09/21].
82. Pharmaceutical Benefits Advisory Committee (PBAC). Idelalisib for FL that is refractory to both rituximab and an alkylating agent, minor resubmissions, PSD March 2016 [online]. Available at: https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2016-03/files/idelalisib-fl-psd-march-2016.pdf [Last accessed 15/09/21].
83. Canadian Agency for Drugs and Technolologies in Health (CADTH). Idelalisib (Zydelig) for follicular lymphoma, September 2016 [online]. Available at: https://www.cadth.ca/zydelig-follicular-lymphoma-details [Last accessed 15/09/21].
84. Canadian Agency for Drugs and Technolologies in Health (CADTH). Obinutuzumab (Gazyva) for follicular lymphoma, June 2017 [online]. Available at: https://www.cadth.ca/gazyva-follicular-lymphoma-details [Last accessed 15/09/21].

© Roche (2022). All rights reserved

85. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU TECHNICAL SUPPORT DOCUMENT 19: PARTITIONED SURVIVAL ANALYSIS FOR DECISION MODELLING IN HEALTH CARE: A CRITICAL REVIEW. Accessed 9 February 2022, http://nicedsu.org.uk/technical-support-documents/partitioned-survival-analysis-tsd/
86. Smare C, Lakhdari K, Doan J, Posnett J, Johal S. Evaluating Partitioned Survival and Markov Decision-Analytic Modeling Approaches for Use in Cost-Effectiveness Analysis: Estimating and Comparing Survival Outcomes. *PharmacoEconomics*. 2020/01/01 2020;38(1):97-108. doi:10.1007/s40273-019-00845-x
87. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of Survival Probabilities for Use in Cost-effectiveness Analyses: A Comparison of a Multi-state Modeling Survival Analysis Approach with Partitioned Survival and Markov Decision-Analytic Modeling. *Med Decis Making*. May 2017;37(4):427-439. doi:10.1177/0272989x16670617
88. Cooper K, Tappenden P, Cantrell A, Ennis K. A systematic review of meta-analyses assessing the validity of tumour response endpoints as surrogates for progression-free or overall survival in cancer. *British Journal of Cancer*. 2020/11/01 2020;123(11):1686-1696. doi:10.1038/s41416-020-01050-w
89. Bachy E, Cerhan JR, Salles G. Early progression of disease in follicular lymphoma is a robust correlate but not a surrogate for overall survival. *Blood Adv*. 2021;5(6):1729-1732. doi:10.1182/bloodadvances.2020003797
90. National Institute for Health and Care Excellence. Technology appraisal guidance [TA531]: Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer. https://www.nice.org.uk/guidance/ta531
91. National Institute for Health and Care Excellence. Technology appraisal guidance [TA713]: Nivolumab for advanced non-squamous non-small-cell lung cancer after chemotherapy. https://www.nice.org.uk/guidance/ta713
92. Wild D, Walker M, Pettengell R, Lewis G. PCN62 UTILITY ELICITATION IN PATIENTS WITH FOLLICULAR LYMPHOMA. *Value in Health*. 2006;9(6):A294. doi:10.1016/S1098-3015(10)63491-2
93. IPSOS. *Ipsos Healthcare Global Oncology Monitor*. 2022.
94. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology*. 2012/02/01 2012;12(1):9. doi:10.1186/1471-2288-12-9
95. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. *Med Decis Making*. Apr 2014;34(3):343-51. doi:10.1177/0272989x13497998
96. Latimer N. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA. http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf
97. Linton K, Julian C, Gibb A, et al. Treatment Patterns and Outcomes in Patients with Relapsed/Refractory Follicular Lymphoma Who Received Third Line Therapy at the Christie NHS Foundation Trust in the UK. *Blood*. 2021/11/23/ 2021;138:5011. doi:https://doi.org/10.1182/blood-2021-148490
98. Batlevi CL, Sha F, Alperovich A, et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. *Blood Cancer Journal*. 2020/07/17 2020;10(7):74. doi:10.1038/s41408-020-00340-z
99. Lymphoma Epidemiology of Outcomes. About LEO. Accessed 25 March 2022, https://leocohort.org/about-leo/
100. Friedberg JW, Taylor MD, Cerhan JR, et al. Follicular lymphoma in the United States: first report of the national LymphoCare study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009;27(8):1202-1208. doi:10.1200/JCO.2008.18.1495
101. Salles G, Schuster SJ, Fischer L, et al. A Retrospective Cohort Study of Treatment Outcomes of Adult Patients with Relapsed or Refractory Low-Grade Follicular Lymphoma (ReCORD-FL). *Blood*. 2021;138(Supplement 1):1349-1349. doi:10.1182/blood-2021-145870

© Roche (2022). All rights reserved

102. Haematological Malignancy Research Network. About HMRN. 21 April 2022. https://hmrn.org/about

103. International Organization of Medical Oncology. IOMedico 21 April 2022. https://www.iomedico.com/en/iom/numbers-data-facts/

104. Palomba ML, Ghione P, Patel AR, et al. A Comparison of Clinical Outcomes from Updated Zuma-5 (Axicabtagene Ciloleucel) and the International Scholar-5 External Control Cohort in Relapsed/Refractory Follicular Lymphoma (R/R FL). *Blood*. 2021/11/23/ 2021;138:3543. doi:https://doi.org/10.1182/blood-2021-147608

105. Rutherford M, Lambert PC, Sweeting MJ, et al. NICE DSU TECHNICAL SUPPORT DOCUMENT 21: Flexible Methods for Survival Analysis. https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD21-Flex-Surv-TSD-21_Final_alt_text.pdf

106. Cooney P. Incorporating Clinical Opinion to Survival Extrapolations: An Applied Example. Accessed 25 March 2022, https://r-hta.org/talk/cooney-2020/philip-cooney-rhta2020.pdf

107. Cope S, Ayers D, Zhang J, Batt K, Jansen JP. Integrating expert opinion with clinical trial data to extrapolate long-term survival: a case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia. *BMC Med Res Methodol*. Sep 2 2019;19(1):182. doi:10.1186/s12874-019-0823-8

108. Birnie R, Ouwens M, Heeg B, Dyer M. NOVEL APPROACHES TO COMMON CHALLENGES IN MODELLING OF SURVIVAL OUTCOMES FOR COSTEFFECTIVENESS ANALYSES IN ONCOLOGY. Accessed 25 March 2022, https://www.ispor.org/docs/default-source/presentations/91714pdf.pdf?sfvrsn=a5b2756f_0

109. Felizzi F, Manevy F, Di Maio D. POSA74 Accounting for Population Heterogeneity Over-Time in Estimating Costs and Outcomes in Cohort-Based Health-Economic Evaluations. *Value in Health*. 2022/01/01/ 2022;25(1, Supplement):S47-S48. doi:https://doi.org/10.1016/j.jval.2021.11.219

110. Choi I, Park S, Song J, Jun JH, Suh D. Utilities for Health States In Patients With Relapsed/Refractory Non-Hodgkin Lymphoma And Factors Influending Utility Values. *Value in Health*. 2015;18(3):A208. doi:10.1016/j.jval.2015.03.1206

111. Cognet M, Taieb V, Bec M, Gauther A, Collins G. Utility values for patients with indolent non-hodgkin lymphoma (INHL) in top five european countries (EU-5). *Haematologica*. 2015;1:578.

112. DSU N. Mapping EQ-5D-5L to 3L. Accessed 25 March 2022, https://nicedsu.sites.sheffield.ac.uk/methods-development/mapping-eq-5d-5l-to-3l

113. Schielzeth H, Dingemanse NJ, Nakagawa S, et al. Robustness of linear mixed-effects models to violations of distributional assumptions. https://doi.org/10.1111/2041-210X.13434. *Methods in Ecology and Evolution*. 2020/09/01 2020;11(9):1141-1152. doi:https://doi.org/10.1111/2041-210X.13434

114. Hernandez Alava M, Wailoo A, Pudney S. METHODS FOR MAPPING BETWEEN THE EQ-5D-5L AND THE 3L FOR TECHNOLOGY APPRAISAL. REPORT BY THE DECISION SUPPORT UNIT. https://drive.google.com/file/d/1ehLw3-Mo7WyHFiwyijI4Ur5kCAPTEAPe/view

115. Beveridge R, Satram-Hoang S, Sail K, et al. Economic impact of disease progression in follicular non-Hodgkin lymphoma. *Leuk Lymphoma*. Nov 2011;52(11):2117-23. doi:10.3109/10428194.2011.592623

116. van Agthoven M, Kramer MH, Sonneveld P, et al. Cost analysis of common treatment options for indolent follicular non-Hodgkin's lymphoma. *Haematologica*. Oct 2005;90(10):1422-32.

117. Clancy Z, Bell E, Chastek B, Bunner S, Copher R, Arekapudi S. PCN118 TREATMENT PATTERNS AND HEALTH CARE COSTS IN PATIENTS WITH FOLLICULAR LYMPHOMA (FL) OR MARGINAL ZONE LYMPHOMA (MZL). *Value in Health*. 2020;23:S43. doi:10.1016/j.jval.2020.04.1609

118. Engel-Nitz N, Dacosta Byfield S, Bancroft T, et al. What Is the Cost of Disease Progression Among Patients with Indolent or Aggressive Non-Hodgkin Lymphoma (NHL)? *Blood*. 2016;128(22):5957-5957. doi:10.1182/blood.V128.22.5957.5957

119. Shoji A, Yamaguchi S, Oya N, Hatake K. PCN239 MEDICAL COST OF RITUXIMAB-CONTAINING THERAPY IN PATIENTS WITH FOLLICULAR LYMPHOMA IN JAPAN: R-B IS

© Roche (2022). All rights reserved

LESS EXPENSIVE THAN R-CHOP. *Value in Health*. 2019;22:S482. doi:10.1016/j.jval.2019.09.434

120. Fowler NH, Chen G, Lim S, Manson S, Ma Q, Li FY. Treatment Patterns and Health Care Costs in Commercially Insured Patients with Follicular Lymphoma. *Journal of health economics and outcomes research*. 2020;7(2):148-157. doi:10.36469/jheor.2020.16784

121. van der Jagt R, Hamm C, Fleury I, et al. Treatment Patterns and Health/Economic Outcomes in Relapsed/Refractory Follicular Lymphoma: A Real-World Study. *Blood*. 2017;130(Supplement 1):5632-5632. doi:10.1182/blood.V130.Suppl_1.5632.5632

122. Lambert-Obry V, van der Jagt R, Hamm C, et al. PCN150: TREATMENT PATTERNS, HEALTH CARE RESOURCE UTILIZATION AND COST ASSOCIATED WITH THE MANAGEMENT OF RELAPSED/REFRACTORY INDOLENT NON-HODGKIN LYMPHOMA. *Value Health*. 2017;20(5):A113-4.

123. Herold M, Sacchi S, Hieke K. The cost of treating relapsed indolent non-Hodgkin's lymphoma in an international setting: retrospective analysis of resource use. *Haematologica*. Jul 2002;87(7):719-29; discussion 729.

124. Frei A, Delmore G, Hitz F, Schwenkglenks M, Szucs T. Cost of 90Y-Ibritumomab Tiuxetan Radioimmunotherapy Versus Cost of Standard Regimens for the Treatment of Relapsed or Refractory Indolent Non-Hodgkin's Lymphoma in Switzerland. *Blood*. 11/16 2007;110:3320-3320. doi:10.1182/blood.V110.11.3320.3320

125. Gaballa S, Levy MY, Keith M, et al. Health care cost impact associated with adverse events (AEs) among treatments in third-line+ (3L+) relapsed/refractory follicular lymphoma (R/R FL). *Journal of Clinical Oncology*. 2021;39(15_suppl):e18836-e18836. doi:10.1200/JCO.2021.39.15_suppl.e18836

126. Meyer N, Song X, Gregory SA, Ramsey SD. Treatment patterns, mortality, and health care utilization and costs in indolent non-Hodgkin lymphoma patients with relapsed or refractory disease. *Journal of Clinical Oncology*. 2014;32(15_suppl):e19557-e19557. doi:10.1200/jco.2014.32.15_suppl.e19557

127. Ohno S, Shoji A, Hatake K, Oya N, Igarashi A. Cost-effectiveness analysis of treatment regimens with obinutuzumab plus chemotherapy in Japan for untreated follicular lymphoma patients. *J Med Econ*. Oct 2020;23(10):1130-1141. doi:10.1080/13696998.2020.1791890

128. NHS. National Cost Collection for the NHS. Accessed 25 March 2022, https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/

129. Roche. Gazyvaro (obinutuzumab) Summary of Product Characteristics. https://www.ema.europa.eu/en/documents/product-information/gazyvaro-epar-product-information_en.pdf

130. Cheson BD, Brugger W, Damaj G, et al. Optimal use of bendamustine in hematologic disorders: Treatment recommendations from an international consensus panel - an update. *Leuk Lymphoma*. 2016;57(4):766-782. doi:10.3109/10428194.2015.1099647

131. National Institute for Health and Care Excellence. British National Formulary. Rituximab. Accessed 25 March 2022, https://bnf.nice.org.uk/medicinal-forms/rituximab.html

132. GOV UK. Drugs and pharmaceutical electronic market information tool (eMIT). Accessed 25 march 2022, https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

133. National Institute for Health and Care Excellence. British National Formulary. Obinutuzumab. Accessed 25 March 2022, https://bnf.nice.org.uk/medicinal-forms/obinutuzumab.html

134. National Institute for Health and Care Excellence. British National Formulary. Lenalidomide. Accessed 25 March 2022, https://bnf.nice.org.uk/medicinal-forms/lenalidomide.html

135. National Institute for Health and Care Excellence. British National Formulary. Vincristine Sulfate. Accessed 25 March 2022, https://bnf.nice.org.uk/medicinal-forms/vincristine-sulfate.html

136. Addicott R, Dewer S. Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Curie Delivering Choice Programme in Lincolnshire. King's Fund. 10 February 2022. https://www.kingsfund.org.uk/sites/default/files/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf

© Roche (2022). All rights reserved

137. National Institute for Health and Care Excellence. British National Formulary. Tocilizumab. Accessed 25 March 2022, https://bnf.nice.org.uk/medicinal-forms/tocilizumab.html

© Roche (2022). All rights reserved

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

Clarification questions

June 2022

File name	Version	Contains confidential information	Date
ID3931_Mosun for RRFL_Clarification Question Responses_ACIC_REDACTED	1.0	No	23 June 2022

Please find the Company's responses to the EAG's clarification questions below.

In addition to these responses, an updated analysis from the [redacted] data cut of the GO29781 study has also been submitted, which provides an additional five months follow up. This supplementary report summarises the updated clinical efficacy data and provides new cost-effectiveness results too. Given the more robust survival estimates available with this data, the company has revised the base case in terms of the distributions applied to the survival curve extrapolations (as summarised in Table 1), while also applying the proportional hazards assumption for OB.

Table 1: Summary of changes to base case in updated model

	August 2021 model base case	**[redacted] model base case**
Mosun PFS distribution	Log-normal	Log-normal
Mosun OS distribution	Exponential	**Log-normal**
RB PFS distribution	Log-normal	**Log-logistic**
RB OS distribution	Weibull	**Exponential**
R^2 PFS distribution	Log-normal	Log-normal
R^2 OS distribution	Weibull	Weibull
OB PFS distribution	Gompertz	**Log-normal**
OB OS distribution	Log-normal	Log-normal
PH assumption – OB	No	**Yes**

PFS, progression-free survival; PH, proportional hazards; R^2, rituximab-lenalidomide; RB, rituximab-bendamustine; OB, obinutuzumab-bendamustine; OS, overall survival
Bold indicates changes from original submitted base case.

Please note, the results from the updated analysis should form the basis for decision making, therefore where appropriate, the responses to the questions below are provided from both the revised base case from the updated data cut as well as the original base case from the August 2021 data cut for completeness.

Section A: Clarification on effectiveness data

A1. Please provide a Kaplan-Meier plot (including risk table) of Time on Treatment from GO29781.

Kaplan Meier plots for time-to-off treatment are provided below. Note, since all patients were off treatment at the time of the August 2021 data cut, there is no difference between this plot and that for the [redacted] data cut.

Figure 1: Time-to-off treatment KM plot (Aug 2021 datacut)

[redacted]

Figure 2: Time-to-off treatment KM plot ([redacted] datacut)

[redacted]

A2. Please provide Kaplan-Meier plots for Overall Survival (OS) and Progression Free Survival (PFS) of the 19 patients in GO29781 who were not refractory to rituximab.

Figure 3: PFS KM plot – non-rituximab refractory patients (Aug 2021 datacut)

Figure 4: OS KM plot – non-rituximab refractory patients (Aug 2021 datacut)

Figure 5: PFS KM plot – non-rituximab refractory patients (datacut)

Figure 6: OS KM plot – non-rituximab refractory patients (datacut)

A3. Please provide p-values for the Cox models fitted to compare the PFS and OS of mosunetuzumab and its comparators before and after the matching had been performed.

P-values for the Cox models before and after matching (for both the August 2021 and data cuts) are provided below. Please note, any difference in bootstrap analyses from previously presented analyses is due to small differences in the random seed generator and does not have any impact on the conclusions.

Mosun vs R^2 – PFS

Table 2: Mosun vs R^2 – PFS (base-case scenario)

Method	August 2021			
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model				
HR (95% CI) from weighted Cox model				
Bootstrap median HR (95% percentile CI)				
Bootstrap median HR (95% BCa CI)				

Table 3: Mosun vs R^2 – PFS (include all covariates)

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]	[redacted]	[redacted]

Mosun vs R^2 – OS

Table 4: Mosun vs R^2 – OS (base-case scenario)

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]	[redacted]	[redacted]

Table 5: Mosun vs R^2 – OS (include all covariates)

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]	[redacted]	[redacted]

Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]	[redacted]	[redacted]

Mosun vs RB – PFS

Table 6: Mosun vs RB – PFS

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
Unadjusted	[redacted]	[redacted]	[redacted]	[redacted]
Optimal pair matching plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Inverse probability of treatment weighting plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Regression adjustment and matching using Austin 2020	[redacted]	[redacted]	[redacted]	[redacted]

Mosun vs RB – OS

Table 7: Mosun vs RB – OS

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
Unadjusted	[redacted]	[redacted]	[redacted]	[redacted]
Optimal pair matching plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Inverse probability of treatment weighting plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Regression adjustment and matching using Austin 2020	[redacted]	[redacted]	[redacted]	[redacted]

Mosun vs OB – PFS

Table 8: Mosun vs OB – PFS

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value

Unadjusted	[redacted]	[redacted]	[redacted]	[redacted]
Full matching plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Inverse probability of treatment weighting	[redacted]	[redacted]	[redacted]	[redacted]

Mosun vs OB – OS

Table 9: Mosun vs OB – OS

Method	August 2021		[redacted]	
	Hazard ratio (95% CI)	p-value	Hazard ratio (95% CI)	p-value
Unadjusted	[redacted]	[redacted]	[redacted]	[redacted]
Full matching plus covariate adjustment	[redacted]	[redacted]	[redacted]	[redacted]
Inverse probability of treatment weighting	[redacted]	[redacted]	[redacted]	[redacted]

A4. Please provide comprehensive information on the covariate definitions that were adjusted for the indirect comparison, as described in section B.2.9.1.4.

Prognostic factors and effect modifiers were classified as either high priority, low priority, or deprioritised according to clinical feedback.

If key covariates were reported using different definitions between GO29781 and the comparator trials, attempts to readjust the covariate definitions in GO29781 were made, where feasible. In instances where this was not possible the covariates might have been considered for testing in sensitivity analyses.

Of the 15 variables identified as relevant prognostic factors in R/R FL by clinical experts, the following generally have standard definitions in this indication and do not typically require any readjustment/redefinition other than potential unit of measurement conversions (for continuous variables) or pooling of categories:

- Age
- ECOG PS
- FLIPI
- Ann Arbor Stage
- Number of prior lymphoma therapies
- Prior ASCT
- Size of the largest lymph node lesion (longest diameter)

- Bone marrow involvement
- Presence of B symptoms
- Time since completion of last therapy

The potential remaining covariates that may have required redefinition in the ITCs submitted for this appraisal are summarised in the table below. Overall, the variable definitions were quite well aligned across comparators in the data sources used for the ITCs, the main exception being bulky disease, which was reported differently for the 2L+ FL subgroup of patients in AUGMENT in the R-Len EMA EPAR, compared to how it was defined in GO29781. As such, this variable was redefined.

Table 10: Definition of prognostic factors

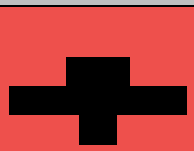
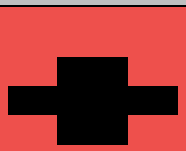
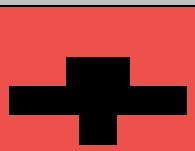
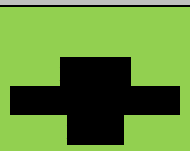
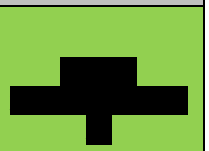
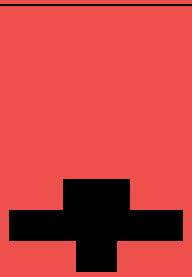
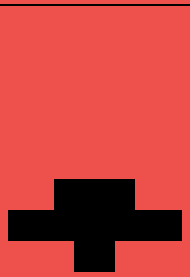
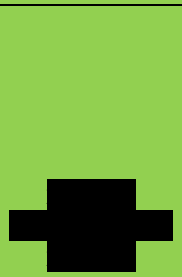
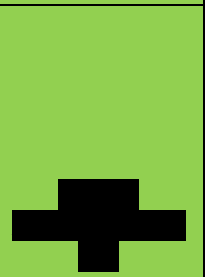
Prognostic factor	Mosunetuzumab data G029781 (n=90)	Comparators		
		OB GADOLIN (n=77)	R^2 AUGMENT (n=147)	BR CONTRALTO + GO29365 (n=46)
Double refractoriness to both an anti-CD20 containing regiment and an alkylating agent, n (%)	48 (53.33) (Failure to respond to previous treatment with both anti-CD20 mAb containing therapy and an alkylating agent containing regimen [alone or in combination], or progression within 6 months)	50 (65) (Failure to respond to previous treatment with both a rituximab containing therapy and an alkylating agent containing regimen [alone or in combination], or progression within 6 months)	0 (0) [assumed based on inclusion criteria]	21 (45.65) (Failure to respond to previous treatment with both a rituximab containing therapy and an alkylating agent containing regimen [alone or in combination], or progression within 6 months)
Refractory to rituximab, n (%)	71 (78.89) (Failure to respond to previous treatment with anti-CD20 mAb containing therapy or progression within 6 months)	77 (100) [based on inclusion criteria] (Failure to respond to previous treatment with rituximab or progression within 6 months)	0 (0) [based on inclusion criteria]	32 (69.6) (Failure to respond to previous treatment with rituximab or progression within 6 months)
Refractory to last prior therapy, n (%)	62 (68.89) (Failure to respond to previous treatment or progression within 6 months)	67 (89.6) (Failure to respond to previous treatment or progression within 6 months)	26 (18) Not reported (assumed to be the same as the one used in GO29781, as it seems to be the definition used across all published studies in FL)	22 (47.83) (Failure to respond to previous treatment or progression within 6 months)

Early relapse status (progression of disease within 24 months), n (%)	47 (52.22) (Progression of disease within 24 months from start of 1L therapy)	42 (54.6) (Progression of disease within 24 months from start of 1L therapy))	56 (38.1) (Relapse/progression of disease within 24 months of initial treatment)	22 (47.83) (Progression of disease within 24 months from start of 1L therapy)
Bulky disease, n (%)	>6 cm: 31 (34.44) >7 cm: 16 (17.78) >10 cm: 2 (2.22)	•Not applicable (deprioritized in favour of size of largest lymph node lesion [longest dimension])	39 (26.5) **(One lesion ≥7 cm or three lesions with ≥3 cm in longest dimension)**	Not applicable (deprioritized in favour of size of largest lymph node lesion [longest dimension])
High LDH, N (%)	35 (38.89) > upper limit of normal (ULN)	28 (36.36) > upper limit of normal (ULN)	39 (26.7) > upper limit of normal (ULN)	15 (32.61) > upper limit of normal (ULN)

A5. In Table 1 of ITC report_AIC.pdf, please explain the difference between the final two rows.

The final row of table 1 in the ITC report was included by mistake. Please find the amended table below.

Table 11: Summary of base-case analyses and MAIC results for unadjusted and adjusted models (bootstrap median HR/OR [95% percentile])

Comparator (Source of data & any limitations)	Mosunetuzumab cohort details and ESS after matching (absolute numbers)	Covariates used in base case analysis		Base-case MAIC results for the comparison of mosunetuzumab versus comparator				
			Method for estimating HR/OR	OS (HR)	PFS (HR)	ORR (OR)	CR (OR)	Discontinuation due to AEs (OR)
Rituximab plus lenalidomide AUGMENT (N=147) 2L+ mixed lymphoma (FL subgroup- 47% received 3L+, 100% of patients not refractory to prior rituximab)	All analyses are based on the full 3L+ FL pivotal cohort from GO29781 (n=90) ESS = ~32.9	• Age (mean) • FLIPI ≥3 (Yes) (prop) • Ann Arbor Stage III–IV (Yes) (prop) • Refractory to last line (prop) • High LDH (Yes) (prop) • Low Hgb level (Yes) (prop) [imputed for AUGMENT] • Bone marrow involvement (Yes) (prop) • POD24 (Yes) (prop) • Bulky disease (Yes) (prop)	HR/OR (95% CI) from unadjusted Cox/ logistic regression model					
			Bootstrap median HR/OR (95% percentile CI) from weighted Cox/logistic regression model					

Abbreviations: 2L+, second-line therapy or beyond; 3L+, third-line therapy or beyond; AEs, adverse events; ASCT, autologous stem cell transplant; CI, confidence interval; CR, complete response; ECOG PS, Eastern Cooperative Oncology Group Performance Status; ESS, effective sample size; FL, follicular lymphoma; FLIPI, follicular lymphoma international prognostic index; Hgb, haemoglobin; HR, hazard ratio; LDH, lactate dehydrogenase; MAIC, matching-adjusted indirect comparison; OR, odds ratio; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; POD24, progression of disease within 24 months; RR, relapsed refractory.

Statistically significant results shown in bold. Point estimates favouring mosunetuzumab indicated in green and those favouring comparator indicated in red.

A6. Please provide details on the patients in GO29781 who were thought to have pseudoprogression and continued mosunetuzumab (e.g. number of patients, how long beyond progression they received mosunetuzumab therapy, what the eventual outcomes were).

The suspicion of pseudoprogression of disease (defined as an increase in tumour size due to the influx of immune cells that may lead to a false judgement of progression) was assessed by investigators and was not captured via a dedicated data field in the electronic case report form (eCRF) . However, upon review of investigators' comment to the response assessment captured in the eCRF, one patient in the pivotal B11 FL expansion cohort was thought to have pseudoprogression and continued to receive mosunetuzumab.

At the first tumour assessment (Study Day 37; after receiving 2 cycles of mosunetuzumab), this subject had one indicator lesion (indicator lesion#2 mesenteric lymph node) which increased > 50% at the first tumour assessment. At the second tumour assessment (Study Day 78; after receiving 4 cycles of mosunetuzumab), indicator lesion#2 remained increased in size compared to baseline. The investigator assessed the response at both timepoints to be "not evaluable", with the comments that "tumour assessment 1 and tumour assessment 2 can be considered as pseudoprogression" and that the patient will undergo a biopsy to confirm "whether or not they have progressed". The patient continued treatment with mosunetuzumab, subsequently achieving a complete response at the third tumour assessment (Study Day 162) which was maintained at all subsequent tumour assessment visits.

A7. Please clarify how the 90 follicular lymphoma (FL) patients included in this submission have been identified compared to the patient characteristics reported in Budde 2022 (e.g. Note text below Table 1 of Budde reports: transformed follicular lymphoma (n = 26), follicular lymphoma grade 3B (n = 1), follicular lymphoma (grade 1-3A; n = 65), but these do not sum to 90).

Please justify why each of these groups of patients (and others if applicable) from Budde 2022 was wholly or partially included/excluded in the 90 patients presented in this submission.

The Budde 2022 paper reported data from the dose escalation portion of the study (1), which excluded the 90 FL patients included in this submission.

The data reported in the company submission is from the pivotal cohort of 90 FL patients who had relapsed after or failed to respond to at least two prior lines of systemic therapy AND received single-agent mosunetuzumab at the recommended phase II dose during the

dose expansion phase of the trial, which was presented at ASH 2021 (2). The 230 patients reported in the Budde 2022 Journal of Clinical Oncology paper were not treated at the recommended phase II dose and is therefore a separate patient population to that reported in the company submission.

A8. Please provide justification with supporting data from the trial as to why patients who discontinued the study and subsequently experienced disease progression or died should be censored at their last follow-up rather than have their event time included in the analysis.

Patients have the right to voluntarily withdraw from the study at any time for any reason. Patients who discontinued from the study were deemed to be lost to follow-up, and therefore were no longer followed up according to the schedule of assessment defined in the study protocol. After study discontinuation, data collection will not be able to continue for patients' tumour response status. This is because information on disease progression (PD) was not collected following study discontinuation since the methodology and criteria of PD assessment after study discontinuation may differ from that defined in the protocol, and the PD information would be considered real world evidence instead of study collected data.

If the patient withdraws from the study, the site's staff may use a public information source (e.g., county records to obtain information about survival status only) per Study GO29781 protocol. Therefore, at the time of the pre-specified primary analysis (15 March 2021), patients who discontinued from the study were followed up for death information only via public information source, and this information was only used to update the secondary efficacy endpoint of overall survival (OS).

As PD information after study discontinuation is not available, using death after study discontinuation as an event for endpoints such as PFS or DOR, would overestimate these endpoints (i.e. less conservative), and therefore patients were censored at the last follow-up prior to study discontinuation. For OS, patients who died after study discontinuation were included as events. The censoring rules for the GO29781 study defined in the statistical analysis plan are provided below.

Table 12: Censoring rules for GO29781

Situation	Endpoint	Date of event/censoring	Outcome
Death before first PD while on study	DOCR, DOR, PFS	Death	Event

Death between adequate assessment visits	DOCR, DOR, PFS	Death	Event
PD documented	DOCR, DOR, PFS	Earliest assessment of progression	Event
PD after more than 1 consecutively missed scheduled visits	DOCR, DOR, PFS	Earliest assessment of progression	Event
Death after more than 1 consecutively missed scheduled visits	DOCR, DOR, PFS	Death	Event
PD or death after the start of NALT	DOCR, DOR, PFS	Last adequate assessment of no progression prior to the start of NALT	Censored
Start of NALT	DOCR, DOR, PFS	Last adequate assessment of no progression prior to the start of NALT	Censored
No death, nor PD prior to CCOD	DOCR, DOR, PFS	Last adequate assessment of no progression	Censored
Study discontinuation prior to death or PD	**DOCR, DOR, PFS**	**Last adequate assessment of no progression**	**Censored**
Death	**OS**	**Death**	**Event**
No death prior to CCOD	OS	Last known alive date or CCOD, whichever is earlier	Censored
Death prior to start of NALT	TTNT	Death	Event
Start of NALT	TTNT	Start of NALT	Event
No death, nor started NALT prior to CCOD	TTNT	Last known alive date or CCOD, whichever is earlier	Censored

CCOD, clinical cutoff date; DOCR, duration of complete response; DOR, duration of response; NALT, new anti-lymphoma therapy; OS, overall survival; PD, progressive disease; TTNT, time to next treatment

A9. Please can the company reperform PFS, CR duration and DOR time-to-event analyses of GO29781 without censoring patients who discontinued the study and subsequently experienced disease progression, and instead using their event times.

Please incorporate these analyses into the economic model.

As mentioned in the answer to A8, patients who discontinued from the study were deemed to be lost to follow-up, therefore PD was not collected following study discontinuation.

There were only three patients that were censored after discontinuing from the study (Table 13). Of these three patients;

- The censoring data was equal to overall survival so no adjustment is required
- One is censored at the last tumour assessment, which has a relative difference of 7 days to OS assessment and therefore the impact would be negligible
- One is censored at the start of treatment and did not receive subsequent therapy

As such, if these patients were to have experienced disease progression, the impact on the analysis in the economic model will be null, and no other information in terms of other dates to censor the patient are available. Furthermore, the relative weights of these patients are minimal (less than 2% per patient), as can be shown below (Table 14).

Table 13: Summary of censored patients who discontinued the study

Subject ID	Parameter	Day of censored evaluation (from study start)
3	Duration of complete response	189
3	Duration of response	189
3	Earliest contributing event to IRF PFS	288
3	Overall survival	288
7	Duration of complete response	N/A
7	Duration of response	N/A
7	Earliest contributing event to IRF PFS	1
7	Overall survival	177
12	Duration of complete response	N/A
12	Duration of response	N/A
12	Earliest contributing event to IRF PFS	100

12	Overall survival	107

Table 14: Weight of censored patients who discontinued study

Subject ID	Weight	Total_weight	Percentage_weight	Comparison
3	0.621262	32.91837	1.89%	R^2
7	0.12158	32.91837	0.37%	R^2
12	0.040828	32.91837	0.12%	R^2
3	0.598646	67.36136	0.89%	OB
12	1.155755	67.36136	1.72%	OB
3	1.094812	80.78967	1.36%	RB
7	0.822584	80.78967	1.02%	RB
12	0.867846	80.78967	1.07%	RB

A10. Please clarify whether patients who discontinued study drug prior to CR have been distinguished from patients who discontinued from the study prior to CR. Please provide more information on the number of patients in these groups and how their responses were included in the CR/ORR analyses.

Patients who discontinued study drug prior to CR are distinguished from those who discontinued the study prior to CR in terms of analysing CR rate. The strategy for how these patients were included in the analysis is summarised below.

- Available assessments after the discontinuation of study drug were used to determine CR status in patients who discontinued the study drug (including COVID-related reasons) prior to achieving CR.
- Patients who discontinued from the study (including COVID-related reasons) prior to achieving CR were included as non-complete responders

The overall number of patients who discontinued study drug and who discontinued from the study (including those who achieved a CR), along with reasons for discontinuation are summarised in Table 15.

Table 15: Reasons for discontinuation in R/R FL expansion cohort

n, (%)	R/R FL cohort n=90
Discontinued from initial treatment	36 (40.0)
Reasons	
Progressive disease	25 (27.8)
Adverse event	4 (4.4)
Physician decision	4 (4.4)
Withdrawal by subject	1 (1.1)
Use of another anti-cancer therapy	2 (2.2)
Discontinued from study at any time, n (%)	12 (13.3)
Reasons	
Death due to progression of disease	6 (6.7)
Withdrawal by subject	4 (4.4)
Death	1 (1.1)
Death due to adverse event	1 (1.1)

A11. Why was low Hgb level imputed for AUGMENT rather than being excluded from the matching analysis? Please repeat the matching analysis excluding this variable.

The comparative effectiveness of mosunetuzumab versus rituximab-lenalidomide (R^2) was assessed via an unanchored MAIC methodology. In order to provide unbiased estimates of relative treatment effects, unanchored MAICs require all known prognostic factors and effect modifiers to be controlled for. A low Hgb level was flagged as a highly relevant prognostic factor that should be controlled for by clinical experts consulted by Roche. Even though a value for this prognostic factor was not available from AUGMENT related data sources, the company attempted to control for it using proxy values to show that the impact on the conclusions of the analyses would be negligible, rather than omitting it completely.

Baseline characteristics and results of the company base-case MAIC excluding low Hgb levels as a confounding factor are reported below, [redacted].

Table 16: Summary of baseline characteristics before and after matching

Variable	Mosunetuzumab unweighted (n=90)	Mosunetuzumab-weighted (ESS=35.32) [39.25%] (low Hgb removed)	Rituximab plus lenalidomide (n=147)
Age (mean)	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (prop)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III–IV (Yes) (prop)	[redacted]	[redacted]	[redacted]

Refractory to last line (prop)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (prop)	[redacted]	[redacted]	[redacted]
Bone marrow involvement (Yes) (prop)	[redacted]	[redacted]	[redacted]
POD24 (Yes) (prop)	[redacted]	[redacted]	[redacted]
Bulky disease (Yes) (prop)	[redacted]	[redacted]	[redacted]

Figure 7: Histogram of MAIC weights

[redacted]

Figure 8: PFS (per IRF assessment) in the mosunetuzumab vs R^2 MAIC (excluded low Hgb as confounding factor) – August 2021[redacted]

Figure 9: OS in the mosunetuzumab vs R^2 MAIC (excluded low Hgb as confounding factor) – August 2021[redacted]

Table 17: PFS hazard ratios – August 2021

Method	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]

Table 18: OS hazard ratios – August 2021

Method	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]

Figure 10: PFS (per IRF assessment) in the mosunetuzumab vs R^2 MAIC (excluded low Hgb as confounding factor) – [redacted]

Figure 11: OS in the mosunetuzumab vs R^2 MAIC (excluded low Hgb as confounding factor) – [redacted]

Table 19: PFS hazard ratios – [redacted]

Method	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]

Table 20: OS hazard ratios – [redacted]

Method	Hazard ratio (95% CI)	p-value
HR (95% CI) from unadjusted Cox model	[redacted]	[redacted]
HR (95% CI) from weighted Cox model	[redacted]	[redacted]
Bootstrap median HR (95% percentile CI)	[redacted]	[redacted]
Bootstrap median HR (95% BCa CI)	[redacted]	[redacted]

A12. Please amend the report and economic model to include results from the weighted MAIC results, rather than the bootstrapped analyses. As described in the ITC report_AIC.pdf section 3.1.5, the bootstrapped results provide confidence intervals rather than effect estimates. If any results from the bootstrapped analyses are still presented, these should be clearly labelled as such.

The ITC sheet in the economic models (August 2021 and [redacted] models) have been amended as requested to include the weighted cox results rather than the bootstrap analysis. Note, this change will only affect the results from the MAIC (i.e. only the comparison versus R^2) and only if the proportional hazard assumption is active. As this assumption does not hold for the R^2 comparison, there is no impact on the deterministic results because the extrapolations were obtained independently on each arm using the weights from the matching, rather than the hazard ratios.

Results from the probabilistic sensitivity analysis where the log normal HR is assumed rather than the bootstrapped HR is presented below (Table 23 and Table 24). These results are largely unaffected from the previous analysis. For the [redacted] data cut, mosunetuzumab is associated with QALY and LY gains versus and would be considered cost-effective at a £20,000 WTP threshold versus R^2.

Table 21: Weighted MAIC probabilistic results for mosun vs R^2 – August 2021 data and original base case settings (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R^2 (base case ICER: £16,103)							
Mosun	[redacted]	10.567	[redacted]				
R^2	[redacted]	9.816	[redacted]	[redacted]	0.751	[redacted]	£15,207

Table 22: Weighted MAIC probabilistic results for mosun vs R^2 - [redacted] data cut and revised base case settings (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R^2 (base case ICER: £8,822)							
Mosun	[redacted]	11.089	[redacted]				
R^2	[redacted]	10.095	[redacted]	[redacted]	0.994	[redacted]	£11,855

A13. Please provide further details, including the number of patients and timing (in relation to study entry and/or confirmation of complete remission) and possibly dose and dosing frequency for those who initially achieved a complete response, but who were re-treated with single agent mosunetuzumab due to subsequent development of progress disease.

Three patients achieved a complete response but were subsequently retreated with mosunetuzumab upon disease progression. A summary of these patients is provided below.

The response criteria for the study was from Cheson 2007 (3), which states that if the bone marrow was involved by lymphoma prior to treatment, the infiltrate must have cleared (i.e. negative) on repeat bone marrow biopsy not more than 42 days after the assessment of a complete response (CR). For patient 3, the first day of response was 27, however, because his/her earliest negative bone marrow biopsy was day 486, responses prior to day 444 (486-42) were downgraded to PR from CR. As a result, the first CR was only confirmed on Day 456.

Table 23: Patients retreated with mosunetuzumab

ID	First CR response day (post study entry)	Number of original cycles	Dose intensity (original), %	Retreatment day (post study entry)	Number of cycles of retreatment
1	95	8	96.1	718	2
2	189	8	87.6	315	8
3	456	8	100.7	771	4

CR, complete response

It is important to acknowledge that retreatment with mosunetuzumab will not be recommended in the label. As regulatory approval for mosunetuzumab will not include retreatment upon progression, it should not be considered in the current appraisal.

A14. Please provide further explanation with regard to the reasons behind the low completion rates (typically <25%) of questionnaires for patient reported outcomes beyond Cycle 8, as described on page 40 of the company submission.

Patient-reported outcomes (PRO) were summarised by scheduled visits (i.e. compliance rates, change from baseline, and responder analyses) focused on those patients who were still receiving initial treatment of mosunetuzumab at the time of PRO questionnaire administration.

Since 18% (16/90) of the patients received more than 8 cycles of treatment, the denominator for calculation of the completion rates beyond 8 cycles was based on these patients, where

patients who completed/discontinued from treatment were excluded from the analysis. Despite this low number of patients, the completion rate was considered high, ranging from 73% to 92% in assessments beyond cycle 8.

Consequently, the presentation of the results of compliance rates, change from baseline, and responder analyses are focused up to Cycle 8 of initial treatment with mosunetuzumab.

A15. Please explain the rationale for limiting assessment of adverse events only to those with onset on the day of or after first administration of mosunetuzumab. AE date for comparator studies were collected and reported for the study duration, can the company provide this data for mosunetuzumab (from GO29781)?

According to the GO29781 study protocol, after informed consent has been obtained but prior to initiation of study drug, only serious adverse events caused by a protocol-mandated intervention (e.g., invasive procedures such as biopsies, discontinuation of medications) should be reported as adverse events (AEs). AEs at screening/baseline prior to the first administration of mosunetuzumab were not collected as per study protocol. However, concurrent deficiencies or impairments at baseline were collected as general medical history and baseline conditions.

The AE reporting window of Study GO29781 is defined as follows: after the initiation of the study drug, all AEs with an onset date between the first dose of study drug and 90 days after the last administration of study drug or initiation of another anti-cancer treatment, whichever occurred first, were also reported. After the AE reporting period (i.e., within 90 days after the last dose of study drug or initiation of another anti-cancer treatment, whichever occurred first), only SAEs assessed to be related to prior study treatment were reported. The assessment of treatment-emergent AEs is based on all AEs that were reported based on the criteria above, and considered appropriate to assess the safety and tolerability profile of mosunetuzumab. To further clarify, the wording "**on the day of or after first administration of mosunetuzumab**" is defined to exclude events that occurred prior to the first dose of mosunetuzumab, rather than to exclude events that occurred after the second or subsequent administration of mosunetuzumab.

The method of presenting AEs from the GO29781 is consistent with comparator studies, which also presented treatment-emergent AEs, as this is the typical approach to presenting safety data from clinical studies.

A16. Please explain why the number of deaths, which is an objective event, differed between IRF assessment and investigator assessment of PFS (company submission, pages 50-51).

The difference in the number of deaths between IRF- and investigator-assessed PFS is due to difference in assessments between these two entities at the timing of the PD assessment by IRF for one patient. This patient was assessed as PD by IRF on day 91, however, the investigator's assessment was stable disease (SD), and remained as SD for subsequent investigator assessments prior to the patient's death. As a result, the IRF-assessed PFS would have the earliest contributing event as PD, while the investigator assessed PFS would have the earliest contributing event as death, which results in one more death as the earliest contributing event in the investigator-assessed PFS compared with the IRF-assessed PFS.

A17. For the EORTC QLQ-C30, please can you (1) clarify if only 2 subscales (physical functioning, fatigue) were completed by participants, or if the whole questionnaire was completed but only these 2 subscales were analysed, and (2) provide a rationale for why data from only these 2 subscales are reported in your submission?

We can clarify that the complete EORTC QLQ-C30 questionnaire was completed by patients but only the physical functioning and fatigue subscales were analysed. This is because these were deemed to be most relevant subscales in understanding the health-related quality of life status of relapsed/refractory NHL patients (4, 5). For instance, fatigue remains an area of unmet concern for survivors of NHL, while these patients have also demonstrated worse physical health functioning compared with disease-free individuals (5). Therefore, it was deemed appropriate that these subscales were selected as the outcomes of interest.

Section B: Clarification on cost-effectiveness data

B1. Please provide cost-effectiveness scenarios where OS is considered equivalent for mosunetuzumab and its comparators, both through using mosunetuzumab OS for the comparator, and through pooling their OS data and extrapolating this appropriately.

Results from the two scenarios for assuming equivalence in OS are provided below, both for the August 2021 data cut and original base case, and the updated [redacted] data cut and revised base case. Discussion and interpretation of these scenarios is provided after the data analyses.

Scenario 1:Using mosunetuzumab OS for the comparator

A scenario in which the OS data for mosunetuzumab is used for the comparator is provided below. This was calculated by applying the OS used data for the weighted mosunetuzumab KM curve to the KM curve for the respective comparator.

Table 24: Scenario 1 - Mosunetuzumab OS is used – August 2021 data and original base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R² (base case ICER: £16,103)							
Mosun	[redacted]	11.07	[redacted]				
R²	[redacted]	11.07	[redacted]	[redacted]	0	[redacted]	£87,545
Mosun vs RB (base case ICER: £51,148)							
Mosun	[redacted]	9.92	[redacted]				
RB	[redacted]	9.92	[redacted]	[redacted]	0	[redacted]	£1,240,724
Mosun vs OB (base case ICER: £10,397 [cost saving])							
Mosun	[redacted]	6.05	[redacted]				
OB	[redacted]	6.05	[redacted]	[redacted]	0	[redacted]	£22,024

Table 25: Scenario 1 - Mosunetuzumab OS is used – [redacted] data and revised base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R² (base case ICER: £8,822)							
Mosun	[redacted]	12.18	[redacted]				
R²	[redacted]	12.18	[redacted]	[redacted]	0	0.13	£51,196
Mosun vs RB (base case ICER: £23,504)							
Mosun	[redacted]	11.63	[redacted]				
RB	[redacted]	11.63	[redacted]	[redacted]	0	0.10	£473,223
Mosun vs OB (base case ICER: £7,727)							
Mosun	[redacted]	9.91	[redacted]				
OB	[redacted]	9.91	[redacted]	[redacted]	0	-0.18	Dominated

Scenario 2: Pooling OS data

Overall survival was obtained pairwise by pooling the weighted mosunetuzumab population with the corresponding comparator. For example, RB OS data was pooled with the RB weighted mosunetuzumab OS data, and extrapolations were obtained from the pooled data without any adjustments by arm. In the model, when the user activates the pooled OS option in the model inputs sheet, both the weighted mosunetuzumab population and the comparator make reference to corresponding pooled OS parameters. Note, the user should be careful when using this option if PH is active, as it will apply the HR from the ITC sheet, which is different from 1. In that case, the user should either deactivate PH or to use a HR of 1. Furthermore, the extrapolation curve chosen for OS for each treatment should be common, as the data of origin is the same.

August 2021 data cut

Figure 12 to Figure 14 show the Kaplan-Meier data for OS for each adjusted mosunetuzumab, relative comparator and respective pooled data. These are shown for each comparator included in the originally submit economic model using the August 2021 data cut. Table 26 shows the cost-effectiveness results where these pooled OS arms informs survival.

Figure 12: Kaplan-Meier data mosunetuzumab, R^2 and pooled OS – August 2021 data

Figure 13: Kaplan-Meier data mosunetuzumab, OB and pooled OS– August 2021 data

Figure 14: Kaplan-Meier data mosunetuzumab, RB and pooled OS– August 2021 data

Table 26: Scenario 2 - pooled OS is used – August 2021 data and original base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R² (base case ICER: £16,103)							
Mosun	[redacted]	12.62	[redacted]				
R²	[redacted]	12.62	[redacted]	[redacted]	0	[redacted]	£85,638
Mosun vs RB (base case ICER: £51,148)							
Mosun	[redacted]	9.16	[redacted]				
RB	[redacted]	9.16	[redacted]	[redacted]	0	[redacted]	£1,241,203
Mosun vs OB (base case ICER: £10,397 [cost saving])							
Mosun	[redacted]	7.64	[redacted]				
OB	[redacted]	7.64	[redacted]	[redacted]	0	[redacted]	£22,024

[redacted] **data cut**

Figure 15 to Figure 17 show the Kaplan-Meier data for OS for each adjusted mosunetuzumab, relative comparator and respective pooled data. These are shown for each comparator included in the revised economic model and base case using the [redacted] data cut. Table 27 shows the cost-effectiveness results where these pooled OS arms informs survival.

Figure 15: Kaplan-Meier data mosunetuzumab, R^2 and pooled OS– [redacted] data

Figure 16: Kaplan-Meier data mosunetuzumab, OB and pooled OS– [redacted] data

[redacted]

Figure 17: Kaplan-Meier data mosunetuzumab, RB and pooled OS– [redacted] data

[redacted]

Table 27: Scenario 2 - pooled OS is used – [redacted] data and revised base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosun vs R^2 (base case ICER: £8,822)							
Mosun	[redacted]	12.52	[redacted]				
R^2	[redacted]	12.52	[redacted]	[redacted]	0	[redacted]	£51,091
Mosun vs RB (base case ICER: £23,504)							
Mosun	[redacted]	11.06	[redacted]				
RB	[redacted]	11.06	[redacted]	[redacted]	0	[redacted]	£236,486
Mosun vs OB (base case ICER: £7,727)							
Mosun	[redacted]	8.76	[redacted]				
OB	[redacted]	8.76	[redacted]	[redacted]	0	[redacted]	Dominated

Discussion and interpretation

The company has significant reservations regarding the request to assume equivalence in OS between mosunetuzumab and the comparators as this may give greater weight to unlikely scenarios when compared to the results from the ITC. There are several reasons why the company believes this is not the most appropriate approach to take.

The company acknowledges that it is difficult to demonstrate a survival benefit in an indolent disease such as follicular lymphoma, particularly when OS data from the study is immature. Moreover, five-year survival among patients progressing within two years of rituximab-based therapy has been demonstrated to be around 50% (6), therefore mature OS data cannot be expected from the primary analysis of a clinical trial in the relapsed/refractory setting. As such, clinical studies focus on surrogate endpoints such as response rate or PFS.

The proposed approach of assuming no overall survival benefit would imply not using the full body of the evidence at hand. In particular, clinical experts have expressed great enthusiasm regarding the complete response rate (60%) and PFS data (median PFS 17.9 months) for mosunetuzumab in the GO29781, stating these results exceed what they would typically see in high-risk third-line patients in clinical practice. However, setting the OS equal for each treatment arm disregards the surrogate effect of these short-term outcomes and places too great an emphasis on unlikely scenarios.

For example, when using the log-normality of the hazard ratios from the [redacted] data cut, the hazard ratio for OS vs RB is <1 around 80% of the time. Conversely, you would expect to see a hazard ratio between 0.9–1.1 vs RB just 7% of the time. This approximately aligns with the EAG proposal to assume equivalence in OS but the available evidence shows this to be relatively unlikely.

Consequently, the company has conducted an additional analysis by pooling individual patient data available in-house for RB (CONTRALTO + GO29365 trials), OB (GADOLIN trial), R-CHOP (EORTC trial) and R-monotherapy (GAUSS trial), comparing ORR and CR vs OS in 3L+ FL patients with ECOG 0–1. This analysis was performed to demonstrate the potential relationship between a short-term outcome in response with long-term survival, as these datasets have much longer follow up. A total of 211 patients were included in the analysis, while the response assessment criteria used were different across datasets, helping to ensure that the results seen are independent of the specific response criteria used across trials. The KM curves and the results from the cox regressions show that there is a strong relationship between response and longer-term survival. Furthermore, the cox regressions show a significant HR indicating a positive relationship between being a responder and improved OS.

Table 28: Cox regressions analysis demonstrating relationship between response rate and OS

Cox-regression of OS with:	Hazard ratio (95% CI) (p-value) (SE, MCSE) Responders vs No Responders
ORR as a covariate	[redacted]
CR as a covariate	[redacted]

CR, complete response; MCSE, Monte Carlo standard error; ORR, overall response rate; OS, overall survival; SE, standard error

Figure 18: KM plot for ORR relationship to OS

[redacted]

Figure 19: KM plot for CR relationship to OS

[redacted]

In addition to the above, it is also important to acknowledge the significant heterogeneity between the patient populations and whether it is appropriate to pool the populations between treatments, especially since the comparator trials are not true third-line plus studies, which has led to imbalances in the matched populations and a bias against mosunetuzumab in the ITC. This is an important consideration since response duration and survival shorten as patients with FL relapse after each line of therapy (7, 8). In fact, a recent systematic literature review and meta-analysis in relapsed/refractory FL after ≥2 prior lines of therapy confirmed these findings, where 24 month OS and PFS rates were greater in 3L+ patients compared to 4L+ patients (9). Of note and in contrast to the GO29781 study, this analysis also demonstrated that very few R/R FL patients who received 2 or more prior therapies achieved CR (12.2%), while approximately one-third of those patients die within 2 years (9).

Consequently, pooling OS data for mosunetuzumab and the comparator would have the effect of increasing the heterogeneity in this population due to mixing of two different treatments from separate populations. Indeed, a recent study which assessed the bias of current methodology in estimating treatment benefits in heterogeneous data sets, demonstrated that fitting a single exponential model to immature follow-up of a heterogeneous population consisting of two components will overestimate the hazard rate and therefore underestimate the mean survival time (10).

In summary, the company considers it inappropriate to assume equivalence in OS between mosunetuzumab and the comparators since this would disregard the influence of short-term outcomes on overall survival and would result in greater uncertainty in determining cost-effectiveness estimates given the increased heterogeneity in the pooled patient populations.

It would be more appropriate to use alternative methods such as ITCs (where feasible) to base decision making on so these important and relevant factors are not masked.

B2. Please provide comprehensive information on the covariate selection process undertaken to select the appropriate mixed effects utility model to inform post-baseline utilities and also provide the full final list of covariates included in the regression model.

The utilities used are post-baseline utilities. All regressions include the following as covariates:

- Centralised baseline utilities
- Random effects at the patient level
- Relevant covariate of interest (e.g. PFS/PPS, on treatment/off treatment)

Some important variables such as the previous number of systemic therapies were explored but were discarded in the end due to non-significance. The approach taken looked for simplicity as the utility would be applied to all comparators equally. Given the pairwise characteristics of the comparisons (the patient characteristics are different depending on the comparison), covariate adjustment should potentially also be done by each pairwise comparison to reflect the characteristics of that comparison, which would make the model unnecessarily complex for little added precision.

B3. Please could you specify the value for the age-matched general population that is being referred to in the sentence below: "At baseline, 83 observations of 90 were available and a mean of 0.767 (SE 0.17) was measured indicating a reduced utility for patients compared to an age-matched general population" (CS pg. 134).

General population utility in the cost-effectiveness model is age adjusted and estimated with the Brazier coefficients (Ara & Brazier 2010) as described in Section B3.4.2.1 of the CS (Table 36) and shown in the Utility sheet in the cost-effectiveness model. Specifically, the age-matched general population utility that is referred to in the CS on page 134 refers to the age-matched general population utility as calculated with the Brazier coefficients for a population with an average age of 60 so as to match the average age from the GO29781 population. This age-matched general population utility is calculated to be 0.8195 and indicates that those in the GO29781 experienced a reduced utility compared to the age-matched general population.

B4. Please specify the cost-inflation index used to make inflation adjustments as PSSRU 2021 presents a range of indices.

Adjustments using inflation indexes were used to estimate prices where current cost data was not available. Section B.3.5.5 in the CS details that the costs of a full blood count, a patient history/physical exam, full profile (urea and electrolytes, liver function tests, calcium), serum IgG, immunoglobin A, immunoglobin M and electrophoresis, and lactate dehydrogenase teste were originally reported as 2005/06 costs in TA243. As such, they were first inflated to 2008/09 costs using the index in Figure 20 and then inflated to current prices using the index in Figure 21, which shows the NHS cost inflation index sourced from page 145 of the Personal Social Services Research Unit (PSSRU) (2021).

Figure 20. NHS cost inflation index (PSSRU 2016)

Year	Hospital & community health services (HCHS)		
	Pay & prices index (1987/8=100)	Annual % increases	
		Prices[2]	Pay[1]
2005/06	240.9	1.9	4.7
2006/07	249.8	3.0	4.1
2007/08	257.0	1.8	3.5
2008/09	267.0	5.2	3.0
2009/10	268.6	-1.3	1.8
2010/11	276.7	2.8	3.1
2011/12	282.5	4.1	0.9
2012/13	287.3	3.1	0.9
2013/14	290.5	1.8	0.7
2014/15	293.1	1.7	0.3
2015/16	297.0	2.7	0.3

Figure 21: NHS cost inflation index (PSSRU 2021)

HCHS/NHS inflators all sectors			
Annual % increases on previous year			
Year	HCHS prices	HCHS pay	HCHS Pay & Prices
2009/2010	-1.30	1.80	0.60
2010/2011	2.80	3.10	3.00
2011/2012	4.10	0.90	2.10
2012/2013	3.10	0.90	1.70
2013/2014	1.80	0.70	1.10
2014/2015	1.70	0.30	0.90
2015/2016	2.70	0.30	1.30
Annual % increases on previous year			
	NHSCII prices	NHSCII pay	NHSCII Pay & Prices
2015/2016	0.45	0.30	0.35
2016/2017	2.16	2.10	2.12
2017/2018	1.07	1.22	1.16
2018/2019	2.43	2.24	2.31
2019/2020	1.62	2.53	2.21
2020/2021	0.22	4.93	3.08

B5. Please provide the following additional information for the Mosunetuzumab arm and comparator groups:

(i) Total number and percentage of patients with adverse events regardless of frequency (Table 33 CS) and

(ii) Percentage (not just numbers) of patients with the adverse events considered in the model (Table 33 CS).

Table 33 from the company submission has been amended below to include the total number of patients in the safety population for each study, the total number and percentage of patients with adverse events regardless of frequency, and the percentage of patients with AEs considered in the model.

Table 29. Adverse events considered in the model (August 2021 data cut)

	Mosunetuzumab n=90	**RB n=50**	**OB n=194**	**R^2 n=176**
Total no. of patients with at least one AE, n (%)	90 (100)	50 (100)	191 (98)	174 (99)
No. of Grade 3–5 TRAE in each arm, n (%)				
Alanine aminotransferase increased	4 (4)	0	0	3 (2)
Anaemia	3 (3)	1 (2)	15 (8)	8 (5)
Aspartate aminotransferase increased	3 (3)	0	0	0
Cytokine release syndrome	2 (2)	0	0	0
Hypokalaemia	0	1 (2)	0	0
Hypophosphatemia	11 (12)	0	0	0
Lymphopenia	3 (3)	0	0	0
Neutropenia	13 (14)	14 (28)	64 (33)	88 (50)
Neutrophil count decreased	7 (8)	0	0	0
Rash erythematous	2 (2)	0	0	0
Thrombocytopenia	0	3 (6)	21 (11)	4 (2)
Tumour flare	2 (2)	0	0	0
Upper respiratory tract infection	2 (2)	0	4 (2)	0
Diarrhoea	0	0	2 (1)	5 (3)
Febrile neutropenia	0	3 (6)	0	0
Vomiting	0	0	4 (2)	0
Infections	0	0	0	4 (2)
Leukopenia	0	2 (4)	0	12 (7)
Cutaneous reaction	0	0	0	10 (6)

Abbreviations: OB, obinutuzumab and bendamustine; RB, rituximab and bendamustine; R2, rituximab and lenalidomide; TRAE, treatment-related adverse events

For the [redacted] data cut, one additional patient experienced a cytokine release syndrome event, taking the number (%) of patients with this event to 3 (3%). There were no further changes with the updated data.

B6. For Figures 22, 25, 28 ,31, 34, 37, please provide number of patients at risk (e.g., at month 0, 5, 10, 15, 20 etc).

The respective curves for the August 2021 as presented in the company submission have been updated to include the table of patients at risk. Please see sections 2.1 to 2.4 in the addendum report for the updated [redacted] data cut for the respective figures for this analysis.

Figure 22. PFS Kaplan Meier for mosunetuzumab (adjusted) and R^2 (unadjusted)

[redacted]

Figure 23. OS Kaplan Meier for mosunetuzumab (adjusted) and R^2 (unadjusted)

[redacted]

Figure 24. PFS Kaplan Meier for mosunetuzumab (adjusted) and RB (unadjusted)

[redacted]

Figure 25. OS Kaplan Meier for mosunetuzumab (adjusted) and RB (unadjusted)

[redacted]

Figure 26. PFS Kaplan Meier for mosunetuzumab (adjusted) and OB (unadjusted)

[redacted]

Figure 27. OS Kaplan Meier for mosunetuzumab (adjusted) and OB (unadjusted)

[redacted]

B7. Please can you provide the results of the regression model for the 65 evaluable patients reported on page 54 (CS) indicating the mean utility values at completion (regardless of reason)?

The 65 evaluable patients reported on page 54 of the company submission refers to the patients who provided data to calculate the change from baseline in the EQ-5D visual analogue scale. These data were derived directly via the case report forms from patients with both baseline values and data reported at the final visit or treatment completion.

Please find the below results for the regressions for the utilities using EQ-5D from the trial along with the regression coding.

August 2021 data cut

Table 30: Patients with utilities in baseline only or more than one observations (August 2021)

	Mean	n
Baseline only	0.6220	3
More than 1 observation	0.7538	81

Table 31:Utility by visit number (August 2021)

Visit number	Utility measurements	Mean
CYCLE 1 DAY 1	84	0.7490361
CYCLE 2 DAY 1	79	0.8113418
CYCLE 4 DAY 1	77	0.8387013
CYCLE 6 DAY 1	66	0.8665231
CYCLE 8 DAY 1	61	0.8200000
CYCLE 10 DAY 1	13	0.7732308
CYCLE 12 DAY 1	11	0.8079091
CYCLE 14 DAY 1	11	0.7723636
CYCLE 16 DAY 1	8	0.8715000
INITIAL TREATMENT COMP OR DISC	69	0.7939420
INITIAL TREATMENT FOLLOW-UP	142	0.8059275
RE-TREATMENT 1 CYCLE 1 DAY 1	3	0.8123333
RE-TREATMENT 1 CYCLE 2 DAY 1	2	0.7790000
RE-TREATMENT 1 CYCLE 4 DAY 1	2	0.8370000
RE-TREATMENT 1 CYCLE 6 DAY 1	1	0.8790000
RE-TREATMENT 1 CYCLE 8 DAY 1	1	0.8370000
RE-TREATMENT 1 COMP OR DISC	1	0.8370000

Table 32: Utility by visit (August 2021)

Visit	Obs	Mean	Std Dev	Min	Max
CYCLE 1 DAY 1	83	0.7490361	0.1839239	0.251	1.000
CYCLE 2 DAY 1	79	0.8113418	0.1572735	0.260	1.000
CYCLE 4 DAY 1	77	0.8387013	0.1410169	0.449	1.000
CYCLE 6 DAY 1	65	0.8665231	0.1386576	0.353	1.000
CYCLE 8 DAY 1	61	0.8200000	0.1957177	0.133	1.000
CYCLE 10 DAY 1	13	0.7732308	0.1399310	0.622	1.000
CYCLE 12 DAY 1	11	0.8079091	0.1404368	0.639	1.000
CYCLE 14 DAY 1	11	0.7723636	0.1615446	0.605	1.000
CYCLE 16 DAY 1	8	0.8715000	0.1165369	0.721	1.000

INITIAL TREATMENT COMP OR DISC	69	0.7939420	0.2024906	0.048	1.000
INITIAL TREATMENT FOLLOW-UP	138	0.8059275	0.1906549	-0.160	1.000
RE-TREATMENT 1 COMP OR DISC	1	0.8370000	NA	0.837	0.837
RE-TREATMENT 1 CYCLE 1 DAY 1	3	0.8123333	0.0818372	0.721	0.879
RE-TREATMENT 1 CYCLE 2 DAY 1	2	0.7790000	0.0820244	0.721	0.837
RE-TREATMENT 1 CYCLE 4 DAY 1	2	0.8370000	0.0000000	0.837	0.837
RE-TREATMENT 1 CYCLE 6 DAY 1	1	0.8790000	NA	0.879	0.879
RE-TREATMENT 1 CYCLE 8 DAY 1	1	0.8370000	NA	0.837	0.837

Figure 28: PFS/PPS regressions

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: eq5duk ~ c_eq5duk + hs + (1 | usubjid)
##    Data: eq5dt_postbl
##
## REML criterion at convergence: -620.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8500 -0.3902  0.1213  0.4920  2.8839
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  usubjid  (Intercept) 0.00851  0.09225
##  Residual             0.01277  0.11300
## Number of obs: 502, groups:  usubjid, 80
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.80699    0.01222  66.037
## c_eq5duk     0.43487    0.06728   6.464
## hspps       -0.04063    0.01936  -2.098
## hsunknown   -0.03125    0.01974  -1.583
##
## Correlation of Fixed Effects:
##           (Intr) c_q5dk hspps
## c_eq5duk  -0.135
## hspps     -0.200  0.023
## hsunknown -0.127 -0.010  0.026
```

[redacted] ***data cut***

Table 33: Patients with utilities in baseline only or more than one observations ([redacted])

	Mean	n
Baseline only	0.7072411	3
More than 1 observation	0.7698254	80

Table 34:Utility by visit number ([redacted])

Visit number	Utility measurements	Mean
CYCLE 1 DAY 1	83	0.7675633
CYCLE 2 DAY 1	79	0.8157868
CYCLE 4 DAY 1	77	0.8390606
CYCLE 6 DAY 1	65	0.8599728
CYCLE 8 DAY 1	61	0.8236125
CYCLE 10 DAY 1	13	0.7680757
CYCLE 12 DAY 1	11	0.7961616
CYCLE 14 DAY 1	11	0.7448310
CYCLE 16 DAY 1	8	0.8750591
INITIAL TREATMENT COMP OR DISC	69	0.7868906
INITIAL TREATMENT FOLLOW-UP	180	0.8106248
RE-TREATMENT 1 CYCLE 1 DAY 1	3	0.8124341
RE-TREATMENT 1 CYCLE 2 DAY 1	2	0.7729668
RE-TREATMENT 1 CYCLE 4 DAY 1	2	0.8496534
RE-TREATMENT 1 CYCLE 6 DAY 1	2	0.8406023
RE-TREATMENT 1 CYCLE 8 DAY 1	2	0.8169294
RE-TREATMENT 1 COMP OR DISC	1	0.8436719

Table 35: Utility by visit ([redacted])

Visit	obs	mean	std.Dev	min	max
CYCLE 1 DAY 1	83	0.7675633	0.1718081	0.1096331	0.9885285
CYCLE 2 DAY 1	79	0.8157868	0.1450423	0.3175636	0.9885285
CYCLE 4 DAY 1	77	0.8390606	0.1405445	0.3581498	0.9885285
CYCLE 6 DAY 1	65	0.8599728	0.1384045	0.4315038	0.9885285
CYCLE 8 DAY 1	61	0.8236125	0.1715958	0.2965281	0.9885285
CYCLE 10 DAY 1	13	0.7680757	0.1417782	0.5937912	0.9885285
CYCLE 12 DAY 1	11	0.7961616	0.1544958	0.5857954	0.9885285
CYCLE 14 DAY 1	11	0.7448310	0.1883594	0.4584472	0.9885285
CYCLE 16 DAY 1	8	0.8750591	0.1094431	0.6998012	0.9885285
INITIAL TREATMENT COMP OR DISC	69	0.7868906	0.2128578	0.0565961	0.9885285
INITIAL TREATMENT FOLLOW-UP	180	0.8106248	0.1937086	-0.3826415	0.9885285
RE-TREATMENT 1 COMP OR DISC	1	0.8436719	NA	0.8436719	0.8436719
RE-TREATMENT 1 CYCLE 1 DAY 1	3	0.8124341	0.0983475	0.7022617	0.8913688
RE-TREATMENT 1 CYCLE 2 DAY 1	2	0.7729668	0.0999922	0.7022617	0.8436719
RE-TREATMENT 1 CYCLE 4 DAY 1	2	0.8496534	0.0084591	0.8436719	0.8556349
RE-TREATMENT 1 CYCLE 6 DAY 1	2	0.8406023	0.0712981	0.7901869	0.8910177
RE-TREATMENT 1 CYCLE 8 DAY 1	2	0.8169294	0.0378196	0.7901869	0.8436719

Figure 29: PFS/PPS regressions

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: eq5duk ~ c_eq5duk + hs + (1 | usubjid)
```

```
##    Data: eq5dt_postbl
##
## REML criterion at convergence: -754.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.7069 -0.4123  0.1257  0.4987  3.3422
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  usubjid  (Intercept) 0.007159 0.08461
##  Residual             0.011146 0.10557
## Number of obs: 543, groups:  usubjid, 80
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.804058   0.011246  71.500
## c_eq5duk     0.505575   0.065392   7.731
## hspps       -0.056775   0.017463  -3.251
## hsunknown   -0.002953   0.016259  -0.182
##
## Correlation of Fixed Effects:
##           (Intr) c_q5dk hspps
## c_eq5duk  -0.156
## hspps     -0.201  0.008
## hsunknown -0.132 -0.012  0.027
```

B8. Please clarify whether any methods were used to account for the missing baseline EQ-5D-5L data and justify any assumptions made regarding missingness.

There were no methods in place to account for missing baseline EQ-5D-5L data.

In the standard utility regressions that are used to inform the cost-effectiveness model for the [redacted] data cut, seven patients had no baseline EQ-5D-5L, three only had baseline measurements and no further data, and 80 had baseline and at least one other datapoint. As the analysis was done on post-baseline utilities, a complete case approach was taken, so the analysis was performed on the 80 patients who had full data. Given the relatively small number of patients with missing data, the potential impact on the regression outcomes is expected to be minimal.

B9: Please implement treatment effect waning where the hazard rates of PFS and OS for mosunetuzumab are equivalent to the comparators beyond the point at which there is no observed follow-up for mosunetuzumab patients for each of the matching scenarios.

As described in Section B.3.2.3.2 and in Table 24 of the CS, application of the treatment waning effect was considered to be inappropriate for all arms as precedence from previous submissions indicates that the application of a treatment waning effect is appropriate where

the majority of patients in the informing pivotal trial are still on treatment at the point where patients had reached the maximum permitted number of cycles (11-13). In contrast for the GO29781 study, only 5.6% of patients were still receiving mosunetuzumab in cycle 17, and by the time of the August 2021 data cut all patients were off treatment and as such applying a treatment waning effect in this setting would be inappropriate.

The PFS and OS curves from the current available data reflect what would be seen in clinical practice where patients receive a course of treatment and subsequently stop after either achieving complete response by cycle 8 or having received the maximum number of 17 cycles. Any post-treatment effect is already captured in these curves, therefore any additional adjustment of this would be done so without clinical justification.

In addition, there is no universally agreed method by which to apply a treatment waning effect nor a universally agreed time for which to apply a treatment waning effect after treatment cessation. In the absence of long term data in this indication, it is very difficult to estimate an appropriate time frame for how long to apply a treatment effect for and therefore this would be done so arbitrarily.

As such, the Company does not believe that it is appropriate to apply a treatment waning effect to mosunetuzumab in the cost-effectiveness model and request that any scenarios where this is represented should be reviewed with due consideration. However, in response to this request, a scenario is presented below where treatment waning is applied to the mosunetuzumab treatment arm. In lieu of any evidence, the treatment waning effect is applied for 60 months as this aligns with previous submissions in R/R FL (11). Results are shown for the originally submitted data cut (Table 36) and the more recent data cut (Table 37).

Table 36: Scenario including treatment waning effect – August 2021 data cut and original base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosun vs R^2 (base case ICER: £16,103)								
Mosun	[redacted]	9.960	[redacted]					
R^2	[redacted]	10.360	[redacted]	[redacted]	-0.401	[redacted]	Dominated	-£10,743
Mosun vs RB (base case ICER: £51,148)								
Mosun	[redacted]	9.263	[redacted]					
RB	[redacted]	8.563	[redacted]	[redacted]	0.700	[redacted]	£92,955	-£37,514
Mosun vs OB (base case ICER: £10,397 [cost saving])								
Mosun	[redacted]	8.300	[redacted]					
OB	[redacted]	7.731	[redacted]	[redacted]	0.569	[redacted]	£1,716	£6,353

Table 37: Scenario including treatment waning effect – [redacted] data cut and revised base case (with mosun PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosun vs R^2 (base case ICER: £8,822)								
Mosun	[redacted]	10.144	[redacted]					
R^2	[redacted]	10.360	[redacted]	[redacted]	-0.216	[redacted]	Dominated	-£8,255
Mosun vs RB (base case ICER: £23,504)								
Mosun	[redacted]	9.243	[redacted]					
RB	[redacted]	8.301	[redacted]	[redacted]	0.942	[redacted]	£67,340	-£34,200
Mosun vs OB (base case ICER: £7,727)								
Mosun	[redacted]	9.333	[redacted]					
OB	[redacted]	8.214	[redacted]	[redacted]	1.119	[redacted]	£8,157	£7,924

Section C: Textual clarification and additional points

C1. Please can you provide definitions for the following (as used in the study): partial response, complete response, objective response.

Please find the definitions below, which were based on the revised response criteria for malignant lymphoma (3) (Figure 30):

- Complete response (CR): defined as disappearance of all evidence of disease using standard criteria for NHL (3). The CR rate is the proportion of patients whose best overall response is a CR based upon IRF or investigator assessment
- Partial response (PR): defined as the regression of measurable disease and no new sites disease using standard criteria for NHL (3). The PR rate is the proportion of patients whose best overall response is a PR based upon IRF or investigator assessment
- Objective response (OR): defined as either a CR or PR using standard criteria for NHL (3). The OR rate (ORR) is the proportion of patients whose best overall response is a PR or CR based upon IRF or investigator assessment.

Figure 30: Response definitions for clinical trials from Cheson 2007

Response	Definition	Nodal Masses	Spleen, Liver	Bone Marrow
CR	Disappearance of all evidence of disease	(a) FDG-avid or PET positive prior to therapy; mass of any size permitted if PET negative (b) Variably FDG-avid or PET negative; regression to normal size on CT	Not palpable, nodules disappeared	Infiltrate cleared on repeat biopsy; if indeterminate by morphology, immunohistochemistry should be negative
PR	Regression of measuable disease and no new sites	≥ 50% decrease in SPD of up to 6 largest dominant masses; no increase in size of other nodes (a) FDG-avid or PET positive prior to therapy; one or more PET positive at previously involved site (b) Variably FDG-avid or PET negative; regression on CT	≥ 50% decrease in SPD of nodules (for single nodule in greatest transverse diameter); no increase in size of liver or spleen	Irrelevant if positive prior to therapy; cell type should be specified
SD	Failure to attain CR/PR or PD	(a) FDG-avid or PET positive prior to therapy; PET positive at prior sites of disease and no new sites on CT or PET (b) Variably FDG-avid or PET negative; no change in size of previous lesions on CT		
Relapsed disease or PD	Any new lesion or increase by ≥ 50% of previously involved sites from nadir	Appearance of a new lesion(s) > 1.5 cm in any axis, ≥ 50% increase in SPD of more than one node, or ≥ 50% increase in longest diameter of a previously identifed node > 1 cm in short axis Lesions PET positive if FDG-avid lymphoma or PET positive prior to therapy	> 50% increase from nadir in the SPD of any previous lesions	New or recurrent involvement

Abbreviations: CR, complete remission; FDG, [^{18}F]fluorodeoxyglucose; PET, positron emission tomography; CT, computed tomography; PR, partial remission; SPD, sum of the product of the diameters; SD, stable disease; PD, progressive disease.

C2. Confirm if the median Duration of CR per investigator in B.2.6.3.2.4 should be highlighted to indicate AIC.

Yes, the statement "[redacted]" should be highlighted as AIC in the company submission document. Updated submission files and redacted copies will be provided.

C3. Confirm if the median DOR in B.2.6.3.2.2 should be highlighted to indicate AIC.

Yes, the median value for DOR per investigator assessment should be marked as AIC. Updated submission files and redacted copies will be provided.

C4. Confirm whether rituximab IV 200 mg / pack size 1 or 100 mg / pack size 2 is correct. If rituximab IV 200 mg / pack size 1, please provide a source and update results if required.

The original source, BNF, states that “Rixathon 100mg/10ml concentrate for solution for infusion vial (Sandoz Ltd)” has a composition of 100mg and consists of 2 vials at a cost of £314.33, which translates to a cost per mg of £1.57. In the economic model and CS it stated that rituximab has a composition of (200mg) at a cost of £314.33. This discrepancy between the source and the economic model and CS is present as the model calculations for rituximab do not take into account the unit size of the drug; specifically, 200mg is available for £314.33 in the form of two 100mg/10ml vials.

It was assumed that 200 mg would therefore be available at a cost of £314.33 in the form of two 100 mg vials. In the model a 375 mg dose is required and the only options available are 100 mg (at a cost of £314.33 for 2 vials) and 500mg (at a cost of £785.84 for 1 vial or £1,571.67 for 2 vials). All choices result in the same price per mg to two decimal places and so the decision was on the assumption that smaller vials may be used. If this assumption is not true, there is anticipated to be no change to the cost-effectiveness results.

C5. Confirm BNF price was used for vincristine 5 mg. (Note: Appendix J states eMIT price used).

In the economic model and CS there are three dosage options for vincristine. These are 1 mg, 2 mg and 5 mg at a cost of £12.71, £6.48 and £329.50, respectively. The NICE Methods Manual indicates in Section 4.4.4 that the preferred source for costs is eMIT. When sourcing costs, eMIT was used and further sources (BNF) were only used if a cost was unavailable.

The first two dosage options, 1 and 2 mg, were sourced from eMIT but the 5mg dosage option is not available from eMIT so it was instead sourced from the BNF. The cost for vincristine is reported as £329.50 and a unit size of 5 vials. Given the model functionality of the model, it was assumed that vincristine had a composition of 5mg and a unit size of 1 vial and the cost is adjusted accordingly. As such, there is a typographical error in appendix J, where the cost for the 5mg dose should state it was sourced from the BNF.

C6. Confirm eMIT price was used for prednisolone 20 mg. (Note: Appendix J states BNF price used).

The approach to costing is as described in the response to C5 and in line with the recommendations made in the NICE Methods Manual. The cost of prednisolone 20 mg was sourced from eMIT, at a cost of £3.30 and a composition size of 28 tablets. There is a typographical error to state this was sourced from the BNF in appendix J.

References

1. Budde LE, Assouline S, Sehn LH, Schuster SJ, Yoon SS, Yoon DH, et al. Single-Agent Mosunetuzumab Shows Durable Complete Responses in Patients With Relapsed or Refractory B-Cell Lymphomas: Phase I Dose-Escalation Study. J Clin Oncol. 2022;40(5):481-91.
2. Budde EL, et al. 127 Mosunetuzumab Monotherapy Is an Effective and Well-Tolerated Treatment Option for Patients with Relapsed/Refractory (R/R) Follicular Lymphoma (FL) Who Have Received ≥2 Prior Lines of Therapy: Pivotal Results from a Phase I/II Study. Poster presented at 63rd ASH Annual Meeting and Exposition, December 11−14, 2021. 2021.
3. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-86.
4. Pettengell R, Donatti C, Hoskin P, Poynton C, Kettle PJ, Hancock B, et al. The impact of follicular lymphoma on health-related quality of life. Ann Oncol. 2008;19(3):570-6.
5. Smith SK, Zimmerman S, Williams CS, Zebrack BJ. Health status and quality of life among non-Hodgkin lymphoma survivors. Cancer. 2009;115(14):3312-23.
6. Casulo C, Byrtek M, Dawson KL, Zhou X, Farber CM, Flowers CR, et al. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. J Clin Oncol. 2015;33(23):2516-22.
7. Batlevi CL, Sha F, Alperovich A, Ni A, Smith K, Ying Z, et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. Blood Cancer J. 2020;10(7):74.
8. Rivas-Delgado A, Magnano L, Moreno-Velázquez M, García O, Nadeu F, Mozas P, et al. Response duration and survival shorten after each relapse in patients with follicular lymphoma treated in the rituximab era. Br J Haematol. 2019;184(5):753-9.
9. Kanters S, Kahl BS, Wiesinger A, Gurung B, Limbrick-Oldfield E, Sudhindra A, et al. Clinical outcomes in patients relapsed/refractory after ≥ 2 prior lines of therapy for follicular lymphoma: A systematic literature review and meta-analysis. J Clin Oncol. 2021;39(15_suppl):e19548-e.
10. Gallacher D, Kimani P, Stallard N. Biased Survival Predictions When Appraising Health Technologies in Heterogeneous Populations. PharmacoEconomics. 2022;40(1):109-20.
11. National Institute for Health and Care Excellence. Technology appraisal guidance [TA627]: Lenalidomide with rituximab for previously treated follicular lymphoma 07 April 2020 [Available from: https://www.nice.org.uk/guidance/ta627.
12. National Institute for Health and Care Excellence. Technology appraisal guidance [TA531]: Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer 18 July 2018 [Available from: https://www.nice.org.uk/guidance/ta531.
13. National Institute for Health and Care Excellence. Technology appraisal guidance [TA713]: Nivolumab for advanced non-squamous non-small-cell lung cancer after chemotherapy 07 July 2021 [Available from: https://www.nice.org.uk/guidance/ta713.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

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	Lymphoma Action
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Lymphoma Action is a national charity, established in 1986, registered in England and Wales and in Scotland. We provide high quality information, advice and support to people affected by lymphoma – the 5th most common cancer in the UK. We also provide education, training and support to healthcare practitioners caring for lymphoma patients. In addition, we engage in policy and lobbying work at government level and within the National Health Service with the aim of improving the patient journey and experience of people affected by lymphoma. We are the only charity in the UK dedicated to lymphoma. Our mission is to make sure no one faces lymphoma alone. Our work is made possible by the generosity, commitment, passion and enthusiasm of all those who support us. We have a policy for working with healthcare and pharmaceutical companies – those that provide products, drugs or services to patients on a commercial or profit-making basis. This includes that no more than 20% of our income can come from these companies and there is a cap of £50k per company. Acceptance of donations does not mean that we endorse their products and under no circumstances can these companies influence our strategic direction, activities or the content of the information and support we provide to people affected by lymphoma.
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	Roche - £22,000.00 - Digital Patient Services, Lymphoma Management, Nurses Training Bristol-Myers Squibb - £21,000.00 - Education and Training, Publications, Lymphoma Matters, Lymphoma Management, Trials Link

companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	none
5. How did you gather information about the experiences of patients and carers to include in your submission?	We sent a survey to our network of patients and carers asking about their experience of current treatment and their response to this new technology, with particular emphasis on quality of life. We received eight responses from patients with a relevant diagnosis, which we have used as the basis of this submission. We have also included information based on our prior experience with patients with this condition.

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?	Follicular lymphoma is usually slow-growing so symptoms develop gradually over time. Many people have few symptoms and some have none. The most common symptom of follicular lymphoma are lumps, but symptoms can include weight loss, fevers, night sweats, frequent infections, itching or fatigue. If follicular lymphoma affects bone marrow, people can develop neutropenia, anaemia and thrombocytopenia. However, follicular lymphoma can be variable and some people might have faster-growing follicular lymphoma that causes more symptoms. It is generally treated with the intention of keeping it under control, rather than curing it. People live with the condition for many years. Both the lymphoma and its treatment can significantly affect quality of life. Patients report being exhausted and are unable to do things they used to. One patient said *"I tire easily. I get aches and pains from scar tissue in my abdomen."* People also report struggling with concentration and memory which affects their working life, social life and ability to do the things they enjoy. Many people need to take time off work or studies, or even stop work completely. This can be very difficult financially. Some people who have previously been employed find it frustrating to rely on government benefits.

	The uncertainty of relapse and the need for repeated treatment is also physically and psychologically challenging for patients. People find it exhausting living with the constant fear of relapse. One patient said it is *"stressful because it can come back at any time. I am currently being monitored every 3 months and I am always worried around the time of my appointment."* Another patient said *"there is always the fear of relapse and I do mark each month that goes by. Every twinge makes me wonder if it's started up again."* Some patients experience frequent relapses requiring repeated courses of treatment that can have significant side effects. Even for patients who have long-lasting remission, lymphoma and its treatment can have lasting consequences. This can be unexpected. One patient said *"several months after my first treatment I have made best efforts to come to terms with living with an incurable cancer which will no doubt return and need subsequent treatments over the rest of my life."* Caring for someone with follicular lymphoma is challenging emotionally, practically and financially. Carers often provide transport to-and-from hospital appointments and treatment sessions, requiring time off work. They also provide emotional support, whilst trying to deal with an emotionally difficult situation themselves. One patient said *"my wife carries the anxiety of losing me especially now it has returned".* Another said *"my husband has been shielding with me and has not been able to resume his hobbies or social life either during the COVID pandemic."* The psychological impact of an 'incurable' disease affects carers as well as patients. One patient said "*there has certainly been an emotional impact on my family. To begin with I tried to play everything down with my mother but she wasn't stupid. My husband has borne the brunt of it because of COVID. My first three treatments were pre-COVID and my son and daughter always came home. It has had a big effect on my daughter."* Another patient said "*for my husband, the mental and emotional toll of seeing me suffer was great. His life was very much interrupted too – cancelled holidays, having to plan life around nearly 3 years of treatment, endless appointments etc."* Some carers worry that their loved one's lymphoma has relapsed whenever they are ill. One patient said its *"very concerning for my husband who constantly keeps a vigil on my health, my emotions and he worries about me daily."*

Current treatment of the condition in the NHS

7. What do patients or carers think of current	Some people do not need treatment initially and enter a period of active monitoring. Patients report finding this psychologically challenging and emotionally draining.

treatments and care available on the NHS?	Patients report finding treatments that require a long appointments or a lot of time in hospital difficult.
8. Is there an unmet need for patients with this condition?	One of the main concerns about current treatments is the lack of a durable response and the need for repeated courses of treatment over the years. One patient said *"a treatment like T Cell that can remain in your body to destroy the cancer cells when they reoccur as a onetime treatment sounds to me to be the wonder drug we all wish for."* People worry that there will not be effective treatment available if or when they experience a relapse. One patient said "*I do worry constantly that when I relapse, they will run out of treatment options for me."* Another said *"more treatment options are needed for patients who relapse or become refractory to the treatments currently available."* Patients also report feeling anxious about having to go through the ordeal of treatment again. One patient said a *"perfect treatment would be a cure with few side effects."* One patient said *"it is good to know that there is continuing research into new treatments and that the availability of such could be life changing."*

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	Patients report the availability of an effective treatment for people who have experienced relapse, and particularly for those who have not responded to existing treatments, is crucial. Treatments that prolong time in remission are seen as particularly important in an 'incurable' condition: *"It may allow patients with relapsed disease to have an additional treatment option giving a longer life-expectancy without debilitating and unpleasant side effects."* *"The fact that it targets 2 proteins of a B cell and T cell must be an improvement on 1 protein targeted by Rituximab and therefore must improve the chances of a lasting remission."* *"I think the effect that it has on T cells (as well as B cells) sounds wonderful, and hopefully very effective in getting rid of the tumour B cells. I hope it would have no adverse effect on the other beneficial immune effects of the T cells.*

	"I have read about drugs that encourage your own body to be able to kill off cancer cells. I would have thought this would lead to fewer side effects and less damage to healthy tissue." *"Harnessing the patient's own immune system to fight their lymphoma is, I believe, the way forward in fighting follicular lymphoma."* *"It sounds like it will be less harsh on the body as I have quite a few side effects that may be with me for life."* *"As it is an antibody treatment it sounds like it is less severe on the body so should had less side effects than the current treatments."*

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	Patients reported the main disadvantages to be related to side effects and how long it takes to receive the treatment: *"There might be different side-effects but overall I feel it would be very positive to be offered a new, cutting-edge treatment."* *"From my point of view it is slightly disadvantageous if it requires a drip administration, whereas Rituximab is a quick injection."* *"I assume that having the infusion would be similar to Obinutuzumab, which takes several hours."* *"However effective the treatment, I think there is still an impact on your family as they will be worried about you and it is hard to imagine any treatment without some side effects."* *"Presumably you would need to have a good level of T cells which some people with follicular lymphoma may not have. The effect it might have on the person being treated and their family and carers would depend on whether it is delivered on a day unit or requires an inpatient stay which for some families could be complicated. Pre-COVID some people on treatment for follicular were able to continue going to work."*

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	

Other issues

13. Are there any other issues that you would like the committee to consider?	

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission.	• Follicular lymphoma can have a significant impact on the quality of life of patients and their carers. • Current treatment options may not produce durable responses and patients are keen for treatments that give them longer remissions. Patients also find the side effects of current treatments difficult. • There is an unmet need for effective, well tolerated treatment that prolongs time in remission. • Patients feel this is an important treatment option for people who have not responded or have relapsed after 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.

ID3931: Mosunetuzumab for treating relapsed or refractory follicular lymphoma.

External Assessment Group (EAG) Report

Produced by	*Warwick Evidence*
Authors	*Daniel Gallacher, Assistant Professor, Warwick Evidence*
	Mandy Maredza, Senior Research Fellow, Warwick Evidence
	Henry Nwankwo, Research Fellow, Warwick Medical School
	Yen-Fu Chen, Associate Professor, Warwick Evidence
	Anna Brown, Information Specialist, Warwick Evidence
	Chris Stinton, Senior Research Fellow, Warwick Evidence
	Iman Ghosh, Research Associate, Warwick Evidence
	Amy Grove, Professor of HTA, Warwick Evidence
Correspondence to	*Amy Grove*
	A.L.Grove@warwick.ac.uk
Date completed	*Date completed (17/08/2022)*

Source of funding: *This report was commissioned by the NIHR Evidence Synthesis Programme as project number 13/55/90.*

Declared competing interests of the authors

None.

Acknowledgements

Dr Beth Harrison Consultant Haematologist, University Hospital Coventry, and Warwickshire.

Dr Dan Todkill Associate Clinical Professor, University of Warwick and Consultant Epidemiologist UK HSA.

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.

Copyright statement:

Copyright belongs to Warwick Evidence. Copyright is retained by Roche for Figures 1-3, 5-16, and 22.

This report should be referenced as follows:

Gallacher D, Maredza M, Nwankwo H, Chen Y-F, Brown A, Stinton C, Ghosh I, Grove A. ID3931: Mosunetuzumab for treating relapsed or refractory follicular lymphoma. EAG Report: A Single Technology Appraisal. Warwick Evidence, 2022.

Contributions of authors

Daniel Gallacher (Assistant Professor) critiqued survival analysis in the company submission. Mandy Maredza and Henry Nwankwo (Health Economists) critiqued the cost-effectiveness evidence and undertook EAG's modelling. Yen-Fu Chen (Associate Professor) and Chris Stinton (Senior Research Fellow) critiqued clinical effectiveness evidence. Iman Ghosh (Research Associate) supported the critique of the clinical effectiveness evidence. Anna Brown (Information Specialist) critiqued the company's searches and conducted additional EAG searches. Amy Grove coordinated the project and commented on draft versions of the report. All authors contributed to the writing and editing of the report.

Please note that: Sections highlighted in [redacted] are '[redacted]. Sections highlighted in [redacted]. Figures that are CIC have been bordered with blue. [redacted] is highlighted in pink.

Table of Contents

Table of Tables

Table of Issues

Table of Figures

Summary

This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) for mosunetuzumab within its marketing authorisation (MA) for treating relapsed or refractory (R/R) follicular lymphoma (FL).

All issues identified in Table 1 represent the EAG's view, not the opinion of NICE.

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are:

- Application and use of half-cycle correction (Section 4.2.2)
- Inclusion of rituximab plus bendamustine comparison (Section 6.1.2)
- Extrapolation of rituximab plus lenalidomide progression free survivable (PFS) beyond [redacted] (Section 6.1.1)
- Magnitude of benefit following parametric extrapolation of rituximab plus lenalidomide overall survival (OS) (Section 6.1.1)
- Choice of parametric extrapolation for rituximab plus bendamustine PFS (Section 6.1.2)
- Extrapolation of rituximab plus bendamustine OS, obinutuzumab plus bendamustine PFS, and obinutuzumab plus bendamustine OS (Sections 6.1.2 and 6.1.3).

Table 1. Summary of key issues

ID3931	Summary of key issue	Report sections
Issue 1	**Concerns over the suitability of the indirect comparisons performed and presented** There is a lack of consistency in the analyses presented surrounding the inclusion of covariates, and more appropriate analyses may produce meaningfully different results.	3.4.1, 3.4.2
Issue 2	**Inconsistent application of adjusted and unadjusted survival data in economic analyses** The company combine the use of unadjusted and adjusted data for the economic analyses to OB and RB. For the comparison to RB, the company present clinical results based on propensity score matching but present cost-effectiveness results based on propensity score weighting.	4.2.6.2, 4.2.6.3
Issue 3	**Unsupported degree of modelled benefit of mosunetuzumab over its comparators.** The modelled benefit of mosunetuzumab in the company's economic analyses far exceeds that degree of benefit suggested by the indirect comparisons where benefit is often unclear due to an apparent lack of difference or contradictory outcomes.	4.2.6
Issue 4	**Unnecessary half cycle correction applied in the model.** The model cycle length is already precise enough to capture the Mosunetuzumab treatment costs accurately. The unnecessary addition of the half cycle correction has a moderate increase on the cost-effectiveness estimates.	4.2.2
Issue 5	**Immature data to model post-progression utilities** Post-progression utilities are based on data collected up to cycle 8 only (i.e., approximately 24 weeks post-baseline). This short-term follow-up period, coupled with the small sample size upon which estimates are based is unlikely to accurately capture the benefits of Mosunetuzumab treatment in improving health-related quality of life. The impact of this immature utility data on the cost-effectiveness estimates is uncertain.	4.2.7.2
Other issues		
Issue 6	**Inclusion of RB as a comparator:** EAG consider that RB is not representative for the wider R+Chemo group. The company selected RB due to data availability for the ITC. The EAG clinical advisor notes differences between patients who receive RB and those who receive other R+Chemo regimens.	Table 4, 3.3.3
Issue 7	**Lack of suitable clinical effectiveness data for the comparison with R-CHOP** The lack of clinical effectiveness data for the comparison with R-CHOP (and lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP), represents a major gap in this assessment given that R-CHOP is a common treatment combination used in third line setting for R/R FL in the UK.[1]	3.3.4
Issue 8	**Generalisability of the patient cohort to the NHS** The cohort included in this assessment is 90 patients. Of those, only two patients are from the UK. This small number of UK patients may raise some concerns relating to the generalisability of the findings from the cohort. The lack of additional data generates uncertainty which cannot be resolved without additional evidence or real work data.	3.2.1, 3.2.2

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 1.6 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.

The company's submission (CS) of the comparative clinical effectiveness, safety, and cost effectiveness of mosunetuzumab was obtained exclusively from a specific cohort of phase II of an open-label, international clinical study GO29781.[2] The primary outcome for GO29781 was the proportion of patients whose best overall response was a complete response (CR) based on independent review facility (IRF) assessment.

Therefore, this EAG report focuses on the cohort of patients (n=90, 2 from the UK) with R/R FL who had at least two prior lines of systemic therapy and who were treated with single-agent mosunetuzumab. This cohort is directly related to the marketing authorisation (MA) obtained.

- We refer to participants and data related specifically to this cohort as GO29781, for brevity.

A matched adjusted indirect comparison (MAIC) was performed following a MAIC feasibility assessment (see Section 3.3). For the comparators outlined in the NICE Final Scope:

- **Rituximab plus lenalidomide (R^2)** (critique provided in Section 3.3.1), the company included three trials in feasibility assessment: AUGMENT,[3, 4] MAGNIFY[5-7] and Gupta 2021.[8, 9] Only AUGMENT was included in the MAIC.

- **Obinutuzumab plus bendamustine (OB)** (critique provided in Section 3.3.2) the company included the GADOLIN trial.[10] As this trial was sponsored by

Roche, the company has individual patient data (IPD) which allowed the indirect treatment comparison (ITC) to be performed using propensity score matching.

- **Rituximab plus bendamustine (RB)** (critique provided in Section 3.3.3) the company included data from two trials, GO29365 and CONTRALTO).[11] The company had access to IPD for these two trials, which allowed the use of propensity matching for ITC.

- **Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP)** (critique provided in Section 3.3.4) the company selected EORTC 20981, a phase 3, open-label, randomised controlled trial. The company considered "[redacted]" due to differences in patient inclusion criteria between GO29781 and EORTC 20981 (CS ITC report, page 66).

    - The EAG agree with this assessment, and therefore, analyses based on data from EORTC 20981 are not described in this EAG report.

    - The lack of suitable clinical effectiveness data for the comparison with R-CHOP (and consequently lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP), represents a major gap in this assessment (see Table 4) given that R-CHOP is a common treatment combination used in third line setting for R/R FL in the UK.[1]

1.1 *Overview of the EAG's key issues*

All issues identified in Table 2 represent the EAG's view, not the opinion of NICE.

Table 2. Summary of key issues

ID3931	Summary of key issue	Report sections
Issue 1	**Concerns over the suitability of the indirect comparisons performed and presented** There is a lack of consistency in the analyses presented surrounding the inclusion of covariates, and more appropriate analyses may produce meaningfully different results.	3.4.1, 3.4.2
Issue 2	**Inconsistent application of adjusted and unadjusted survival data in economic analyses** The company combine the use of unadjusted and adjusted data for the economic analyses to OB and RB. For the comparison to RB, the company present clinical results based on propensity score matching but present cost-effectiveness results based on propensity score weighting.	4.2.6.2, 4.2.6.3
Issue 3	**Unsupported degree of modelled benefit of mosunetuzumab over its comparators.** The modelled benefit of mosunetuzumab in the company's economic analyses far exceeds that degree of benefit suggested by the indirect comparisons where benefit is often unclear due to an apparent lack of difference or contradictory outcomes.	4.2.6
Issue 4	**Unnecessary half cycle correction applied in the model.** The model cycle length is already precise enough to capture the Mosunetuzumab treatment costs accurately. The unnecessary addition of the half cycle correction has a moderate increase on the cost-effectiveness estimates.	4.2.2
Issue 5	**Immature data to model post-progression utilities** Post-progression utilities are based on data collected up to cycle 8 only (i.e., approximately 24 weeks post-baseline). This short-term follow-up period, coupled with the small sample size upon which estimates are based is unlikely to accurately capture the benefits of Mosunetuzumab treatment in improving health-related quality of life. The impact of this immature utility data on the cost-effectiveness estimates is uncertain.	4.2.7.2
Other issues		
Issue 6	**Inclusion of RB as a comparator:** EAG consider that RB is not representative for the wider R+Chemo group. The company selected RB due to data availability for the ITC. The EAG clinical advisor notes differences between patients who receive RB and those who receive other R+Chemo regimens.	Table 4, 3.3.3
Issue 7	**Lack of suitable clinical effectiveness data for the comparison with R-CHOP** The lack of clinical effectiveness data for the comparison with R-CHOP (and lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP), represents a major gap in this assessment given that R-CHOP is a common treatment combination used in third line setting for R/R FL in the UK.[1]	3.3.4
Issue 8	**Generalisability of the patient cohort to the NHS** The cohort included in this assessment is 90 patients. Of those, only two patients are from the UK. This small number of UK patients may raise some concerns relating to the generalisability of the findings from the cohort. The lack of additional data generates uncertainty which cannot be resolved without additional evidence or real work data.	3.2.1, 3.2.2

The key differences between the company's preferred assumptions and the EAG's preferred assumptions include the application and use of half-cycle correction in the model (Section **Error! Reference source not found.**), inclusion of RB as a comparator (Section 6.1.2) and extrapolation of effective estimates from the MAIC, namely:

- Extrapolation of R^2 PFS beyond [redacted] (Section 6.1.1)
- Magnitude of benefit following parametric extrapolation of R^2 OS (Section 6.1.1)
- Choice of parametric extrapolation for RB PFS (Section 6.1.2)
- Extrapolation of RB OS, OB PFS, and OB OS (Sections 6.1.2 and 6.1.3).

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 incremental cost-effectiveness ratio (ICER) is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Utility values assigned to the different health states
- Parametric model fit to PFS and OS data.

Overall, the technology is modelled to affect costs by:

- Parametric model fit to PFS and OS data
- Choice of treatment following disease progression
- Frequency of adverse events
- Costs and duration of treatment administration.

The modelling assumptions that have the greatest effect on the ICER are:

- Utilities assigned to the different health states
- Choice of parametric model fit to PFS and OS data
- Assumption of survival benefit.

1.3 *The decision problem: summary of the EAG's key issues*

The company submission (CS) decision problem partially matches the NICE final scope. The population and comparators are narrower than the final scope.

The narrower population is in line with the MA, however, the EAG consider the comparators only partially appropriate. The key deviations are described in detail in Table 4 and Issue 6.

1.4 *The clinical effectiveness evidence: summary of the EAG's key issues*

- The company conducted a reasonable systematic literature review (SLR) to identify evidence for both 'standard of care treatments' and 'novel emerging therapies' for the management of relapsed or refractory (R/R) FL (Section 3.1).

- The EAG conclude that there was insufficient detail provided in the CS for the selection of studies included in the feasibility assessment for MAIC.

- Five of the 21 studies included in the company's feasibility assessment were sponsored by Roche (GO29781,[2, 12] GADOLIN,[10] CONTRALTO,[13] G029365,[14] EORTC 20981[15, 16] IPD from these studies were included in an indirect comparison using propensity score matching without further consideration in the MAIC feasibility assessment. EAG critique of studies included in the ITC is provided in Sections 3.2 and 3.3.

The EAG consider the estimates of relative effectiveness of mosunetuzumab compared with comparator interventions obtained from the ITCs to be highly uncertain and potentially biased due to issues related to sparsity of data and methodological limitations. Detailed EAG critique of the methods used to perform the MAIC is provided in Section 3.4.

Issue 1: Concerns over the suitability of the ITCs

Report section	3.4.1, 3.4.2
Description of issue and why the EAG has identified it as important	**Unsuitable ITCs** There is a lack of consistency in the analyses presented surrounding the inclusion of covariates, and more appropriate analyses may produce meaningfully different results, which may be more influential when extrapolations are performed.
What alternative approach has the EAG suggested?	The EAG has been unable to investigate the potential impact or ensure the reliability of the results presented.
What is the expected effect on the cost-effectiveness?	Uncertain
What additional evidence or analyses might help to resolve this key issue?	The EAG suggests performing additional MAIC and propensity score analyses (removing correlated interaction-based terms), maximising effective sample size, to increase the validity of the results already presented.

Issue 2: Inconsistency in selection of data to extrapolate

Report section	4.2.6.2, 4.2.6.3
Description of issue and why the EAG has identified it as important	**Extrapolation inconsistencies:** The company combine the use of unadjusted and adjusted data for the economic analyses to OB and RB. For the comparison to RB, the company present clinical results based on propensity score matching but present cost-effectiveness results based on propensity score weighting.
What alternative approach has the EAG suggested?	The EAG has been unable to perform a correction for this issue.
What is the expected effect on the cost-effectiveness estimates?	Uncertain
What additional evidence or analyses might help to resolve this key issue?	The EAG suggests implementing economic analyses for OB and RB that are based on data that is consistent with the clinical section of the report, and that is also consistent with company's external ITC Report. (i.e., IPTW)

Issue 3: Unsupported degree of mosunetuzumab benefit modelled by company

Report section	4.2.6
Description of issue and why the EAG has identified it as important	**Mosunetuzumab benefit modelled by company** The modelled benefit of mosunetuzumab in the company's economic analyses far exceeds that degree of benefit suggested by the indirect comparisons where benefit is often unclear due to an apparent lack of difference or contradictory outcomes.
What alternative approach has the EAG suggested?	The EAG has modelled using different parametric models and sets of assumptions that offer an improved consistency with the output of the indirect treatment comparisons.
What is the expected effect on the cost-effectiveness estimates?	Large – see EAG base case Section 4.2.6
What additional evidence or analyses might help to resolve this key issue?	Longer follow-up or ideally head-to-head trials would provide improved estimates of mosunetuzumab benefit.

1.5 *The cost-effectiveness evidence: summary of the EAG's key issues*

- The company's review of cost-effectiveness evidence was reasonable, but contained some errors detailed in Section 4.1.

- The company submitted a simple *de novo* cost-utility model using partitioned survival with a weekly cycle length and a 40-year time horizon. See Section 4.2.2 for our critique.

The EAG's concerns regarding the MAIC analysis and subsequent extrapolations generate uncertainty in the results of cost-effectiveness analysis (see Sections 4.2.6 and 5).

Issue 4: Unnecessary half cycle correction applied in the model

Report section	4.2.2
Description of issue and why the EAG has identified it as important	**Half cycle correction applied in the model which is not required** The model cycle length is already precise enough to capture the Mosunetuzumab treatment costs accurately. The EAG considers application of a half-cycle correction for Time To Off Treatment unnecessary.
What alternative approach has the EAG suggested?	The EAG has performed additional analyses removing half-cycle correction for mosunetuzumab drug costs and included this in the EAG's base case assumptions.
What is the expected effect on the cost-effectiveness estimates?	Moderate increase– see Table 43
What additional evidence or analyses might help to resolve this key issue?	No additional evidence needed

Issue 5: Immature data to model post-progression utilities

Report section	4.2.7.2
Description of issue and why the EAG has identified it as important	**Data used to model post-progression utilities** Post-progression utilities are based on data collected up to cycle 8 only (i.e., approximately 24 weeks post-baseline). This short-term follow-up period, coupled with the small sample size upon which estimates are based is unlikely to accurately capture the benefits of mosunetuzumab treatment in improving health-related quality of life.
What alternative approach has the EAG suggested?	The EAG has been unable to offer an alternative approach

What is the expected effect on the cost-effectiveness estimates?	Uncertain
What additional evidence or analyses might help to resolve this key issue?	Longer follow-up or use of external data sources with more mature data would provide improved estimates of mosunetuzumab benefit on health-related quality of life.

1.6 *Other key issues: summary of the EAG's view*

In addition to the key issues outlined in Section 1.1, the EAG note the following three 'other issues'. The EAG acknowledge that these issues could not be addressed without suitable additional real-world evidence (RWE) or additional trial data.

Issue 6: Representativeness of rituximab + bendamustine (RB) comparator

Report section	Table 4
Description of issue and why the EAG has identified it as important	**Representativeness of RB as a comparator.** Rituximab + bendamustine (RB) was selected by the company as representative for Rituximab+Chemotherapy (R+Chemo) based on data availability. The EAG clinical adviser suggest that RB is **not** a good representative for R+Chemo. There are differences between patients who receive RB and those who receive other R+Chemo regimens. The EAG clinical advisor suggests that RB may be used in 1st and 2nd line (notes this may be common international practice), yet it is unlikely that a patient who relapses will be given RB again at 3rd line.
What alternative approach has the EAG suggested?	None.
What is the expected effect on the cost-effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Additional impartial clinical expert opinion.

Issue 7: Lack of suitable clinical effectiveness data for the comparison with R-CHOP

Report section	3.3.4
Description of issue and why the EAG has identified it as important	**Lack of clinical effectiveness data for the comparison with R-CHOP:** Major differences in patient inclusion criteria between GO29781 and EORTC 20981 were identified which resulted in a small and imbalanced proportion of patients randomised to the R-CHOP arm remaining in the selected patient cohort ([redacted]). The company and EAG agree that the "[redacted]" (CS ITC report, page 66). The lack of clinical effectiveness data for the comparison with R-CHOP (and lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP), represents a major gap in this assessment given that R-CHOP is a common treatment combination used in third line setting for R/R FL in the UK.[1]
What alternative approach has the EAG suggested?	None.
What is the expected effect on the cost-effectiveness estimates?	Uncertain. No cost-effectiveness analysis were performed for R-CHOP
What additional evidence or analyses might help to resolve this key issue?	The Company state in the ITC (see CS Section B.2.9.) and economic model (see CS Section B.3.2.5.**Error! Reference source not found.**) rituximab in combination with chemotherapy was represented solely by the RB regimen. They suggest that an ITC against R-CHOP was attempted but, despite availability of patient-level data from the EORTC 20981 trial, the analysis proved not to be methodologically feasible. Although we agree with the company assessment, we consider alterative UK data could be obtained from real world evidence (RWE). In the absence of any suitable RWE or trial data, clinical effectiveness data for the comparison with R-CHOP remains uncertain even though it is a comparator included in the NICE scope and decision problem (Table 4).

Issue 8: Generalisability of the patient cohort to the NHS

Report section	3.2.1 and 3.2.2
Description of issue and why the EAG has identified it as important	**Generalisability of the patient cohort to the NHS:** The primary outcome for GO29781 was the proportion of patients whose best overall response was a complete response (CR). This cohort of 90 patients only contained **two** patients from the UK. This small number of UK patients may raise some concerns relating to the generalisability of the findings from the cohort.
What alternative approach has the EAG suggested?	None
What is the expected effect on the cost-effectiveness estimates?	Uncertain
What additional evidence or analyses might help to resolve this key issue?	Without RWE or longer trial follow up, the EAG does not consider additional evidence exists which would increase the size of the patient cohort.

1.7 *Summary of EAG's preferred assumptions and resulting ICER*

The EAG preferred assumptions for each comparison are as follows (see Section 6.1):

Mosunetuzumab versus R^2 comparison

1. Given the high number of censored observations and subsequently very small number of actual observations that inform the mosunetuzumab weighted extrapolated model beyond [redacted], the EAG sets progression free survival (PFS) for mosunetuzumab equal to R^2 from [redacted] whilst maintaining the log-normal extrapolation used by the company.

2. For the overall survival (OS), the EAG prefers a pooled overall survival with a log-normal extrapolation given the uncertainty around the long-term extrapolation of OS.

3. Remove half cycle correction for Time To Off-Treatment (TTOT) from the mosunetuzumab arm.

Mosunetuzumab versus RB comparison

4. The EAG prefers to revert to the log-normal distribution to extrapolate PFS in the RB arm as was the case in the original company submission (section 4.2.6.2.1) whilst maintaining log-normal extrapolation for mosunetuzumab.

5. The EAG prefers an exponential model, fitted to pooled OS data from both treatments (section **Error! Reference source not found.**).

6. Remove half cycle correction for TTOT from the mosunetuzumab arm.

Mosunetuzumab versus OB comparison

7. For mosunetuzumab PFS, the EAG maintains a log-normal extrapolation used by the company. Non-proportional hazard is assumed, and an exponential extrapolation is used for OB PFS.

8. For OS, non-proportional hazard is assumed, and an exponential extrapolation is used for mosunetuzumab

9. Half-cycle correction is removed for TTOT from the mosunetuzumab arm.

Table 3 outlines the EAG's preferred assumptions with changes to each assumption from the company's base case presented individually. We present numerical estimates of the resulting ICERs in a fully incremental analysis and indicate the change from the company's base case ICERs.

Table 3: Summary of EAG's preferred assumptions and ICER

Scenario	Incremental cost	Incremental QALYs	ICER (change from company base case)
Company's base case: vs R[2] vs RB vs OB	[redacted]	[redacted]	[redacted]£8,822 £23,504 £7,727
EAG preferred base case assumptions **vs R[2]** PFS equal after [redacted] Pooled lognormal extrapolation of OS No half cycle correction of mosunetuzumab TTOT **vs RB** log-normal extrapolation of RB PFS Pooled exponential extrapolation of OS No half cycle correction of mosunetuzumab TTOT **vs OB** non-proportional hazard assumption Exponential extrapolation of OB PFS Exponential extrapolation of OB OS No half cycle correction of mosunetuzumab TTOT	[redacted]	[redacted]	[redacted]Dominated[1] [redacted]£248,335 [redacted]Dominated[1]

[1]Mosunetuzumab dominated

Modelling errors identified and corrected by the EAG are described in Section 4.2
For further details of the exploratory and sensitivity analyses done by the EAG, see Section 6.

Abbreviations

AE	Adverse event
AIC	Akaike information criterion
ASCT	Autologous stem cell transplantation
ASTCT	American Society for Transplantation and Cellular Therapy
BIC	Bayesian information criterion
BMI	Body mass index
BNLI	British National Lymphoma Investigation
BR	Bendamustine with Rituximab
BSC	Best supportive care
BSH	British Society for Haematology
CAR-T	Chimeric antigen receptor T-cell
CEAC	Cost-Effectiveness Acceptability Curve
CHOP	Cyclophosphamide, Doxorubicin, Vincristine and Dexamethasone
CI	Confidence interval
CR	Complete response
CRS	Cytokine release syndrome
CRUK	Cancer Research UK
CS	Company submission
CT	Computed tomography
CVP	Cyclophosphamide, Vincristine and Prednisolone
DLBCL	Diffuse large B-cell lymphoma
DOR	Duration of response
EAG	External Assessment group
ECOG	Eastern Cooperative Oncology Group
ECOG PS	Eastern Cooperative Oncology Group Performance Score
eCRF	electronic Case Report Forms
EMA	European Medicines Agency
EORTC	European Organization for the Research and Treatment of Cancer
EORTC QLQ-C30	European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire
EPAR	European public assessment report
ESS	Effective sample size
FACT-Lym	Functional Assessment of Cancer Therapy – Lymphoma
ESMO	European Society for Medical Oncology
FDG-PET	Fluorodeoxyglucose-positron emission tomography
FL	Follicular lymphoma
FLIPI	Follicular Lymphoma International Prognostic Index
GELF	Groupe d'Etude des Lymphomes Folliculaires
Hb	Haemoglobin
HRQoL	Health-related quality of life
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
IPD	Individual patient data
IPTW	Inverse probability of treatment weighting
IRF	Independent review facility
ISRT	Involved-site radiotherapy

ITC	Indirect treatment comparison
IV	Intravenously
KM	Kaplan-Meier
LDH	Lactate dehydrogenase
LYs	Life Years
MA	Marketing Authorisation
MAIC	Matched-adjusted indirect comparisons
MCL	Mantle cell lymphoma
MZL	Marginal zone lymphoma
NE	Not evaluable
NHL	Non-Hodgkin lymphoma
NMB	Net-Monetary Benefit
OB	Obinutuzumab Bendamustine
ORR	Overall response rate
OR	Odds ratio
OS	Overall survival
OWSA	One-way sensitivity analysis
PD	Progressive disease
PET-CT	Positron emission tomography - computed tomography
PFS	Progression-free survival
PI3K	Phosphoinositide 3-kinase
POD24	Progression of disease within 24 months from front-line therapy
PR	Partial response
PRO	Patient reported outcome
PSS	Personal social services
QALYs	Quality-Adjusted Life Years
R	Rituximab
R+Chemo	Rituximab + Chemotherapy
R^2	Rituximab with lenalidomide
RB	Rituximab + bendamustine
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
RCT	Randomised controlled trials
R-CVP	Rituximab, Cyclophosphamide, Vincristine and Prednisolone
RL	Rituximab Lenalidomide
R/R	Relapsed / Refractory Follicular Lymphoma
RWE	Real World Evidence
SAE	Serious adverse event
SD	Stable disease
SEER	Surveillance, Epidemiology, and End Results
SLR	Systematic literature review
TA	Technology appraisal
tFL	Transformed follicular lymphoma
TTOT	Time To Off-Treatment
VAS	Visual analogue scale
VEN+BR	Venetoclax with bendamustine and rituximab
VEN+R	Venetoclax with rituximab

External Assessment Group Report

2 INTRODUCTION AND BACKGROUND

2.1 *Introduction*

Remit of the appraisal

To appraise the clinical and cost effectiveness of mosunetuzumab within its marketing authorisation for treating relapsed or refractory (R/R) follicular lymphoma (FL).

Condition, symptoms, and survival

Follicular lymphoma is a form of non-Hodgkin lymphoma (NHL), a cancer of the lymphatic system. FL is the most common type of low grade lymphoma. Of all people with NHL, 19% will have FL.[17] According to Cancer Research UK, approximately 2,600 people in the UK are diagnosed with FL each year.[17] (The company states that 2,476 people were diagnosed in England in 2019 [CS Document B page 18]).

FL mainly affects adults who are more than 60 years of age.[17] Population-based cancer registries were used to estimate prevalence of FL at 4.8 per 10,000 people (national cancer registries in the EU-27, US SEER registry [Slovenia and Italy] 14% of the EU-27 population.[18] The company provide age-standardised incidence rates per 100,000 persons of "*4.9 [95% CI: 4.6, 5.2] in men and 4.4 [95% CI: 4.2, 4.7] in women"* and age-standardised mortality rates from FL per 100,000 population for "*men [0.5, 95% CI: 0.4, 0.5] and women [0.4, 95% CI: 0.3, 0.5]".* (CS Document B page 19).

FL is usually characterised by a chronic course of disease and disease relapses. Commonly reported FL symptoms include painless swellings in the groin, neck and armpit which are enlarged lymph nodes. Other reported symptoms (known as B symptoms) include high temperature, night sweating, and weight loss.

Diagnosis, staging and grading

NHL generally is classified into two groups, low grade, and high grade, depending on speed of growth, and spread. FL is low grade (indolent disease) and subdivided into 1, 2, and 3A and 3B determined by the number of large, FL cells (centroblasts) that can be seen. Unlike grade 1-3A, grade 3B FL is faster growing and likely to be treated as a high-grade lymphoma, or what the company terms "*aggressive disease*". Grade 3B is outside of the scope of this appraisal (see Table 4).

Diagnosis of FL include lymph node biopsy and blood testing. Staging of FL is performed depending on how many groups of lymph nodes are affected, size of area, location in the body, and whether other organs outside of the lymphatic system are affected.[19] Staging is performed according to the Lugano classification (see Table 3 CS Document B page 18), based on fluorodeoxyglucose-positron emission tomography (FDG-PET). FL is graded via histological appearance of FL biopsy samples. "*Stage I–II FL is considered limited stage, while advanced stage disease comprises stages III-IV."*

Survival

Data from CRUK, suggest almost 90% survive FL for 5 years or more after diagnosis. Survival reduces as people progress to more severe disease and periods of remission can become shorter (in low risk groups, almost 100% survive for 5 years, in medium risk groups this reduces to 90% and in high risk groups, 75% survival at 5 years).[17] Data submitted to the EMA suggests that survival of FL patients is improving, with median OS estimates ranging between 12.5 and 14.8 years, which are consistent with data observed in Denmark, Sweden, and Spain.[18] In the CS, the company provide age-standardised net survival from FL in England: *94.8% (95% CI: 94.3%, 95.3%) at 1 year from diagnosis and 83.0% (95% CI: 81.2%, 84.8%) at 5 years from diagnosis.*

Transformation

Low grade NHL can transform into more aggressive high-grade lymphoma over time. Treatment for transformed NHL is the same as for high grade lymphoma (see Section 2.1). The company suggest that risk of transformation is approximately 28%

over 10 years with median survival of 1.2 years from transformation. (CS Document B page 14)

2.2 *Background*

Mosunetuzumab (Lunsumio®, Roche) is a medication for cancer used to treat adults with follicular lymphoma (FL) that does not respond to (refractory) or has come back (relapsed) after at least two previous treatments.[20]

Mechanism of action

According to the EMA submission, mosunetuzumab is a type of antibody treatment called a T-cell bispecific antibody. It can attach to two targets, a protein called CD20 expressed on B cells and a protein called CD3 on T cells. Some of the B-cells in patients with FL can become cancerous. Mosunetuzumab is expected to bring healthy T cells into close contact with cancerous B cells, thereby helping T cells to get rid of the cancerous B cells.[21]

Treatment overview

Treatments for FL includes radiotherapy for limited disease (stage 1 or 2), and active monitoring when people do not have any symptoms. Treatment for advanced disease (stage 3 or 4) which aim to control, rather than cure, disease includes first and second line treatment and active monitoring. Maintenance treatment in remission is rituximab every two months for up to two years.

First line treatment for symptomatic FL is rituximab monotherapy (a monoclonal antibody), which might delay chemotherapy. This may be followed by combination treatment 'chemoimmunotherapy' (chemotherapy with rituximab). Rituximab may be given with bendamustine, a combination of cyclophosphamide, doxorubicin, vincristine and dexamethasone (CHOP), a combination of cyclophosphamide, vincristine and prednisolone (CVP) or chlorambucil.[17] In the CS, the company refer to The British Society for Haematology (BSH)[22] and The European Society for Medical Oncology (ESMO) practice guidelines on FL,[23] to outline current management of FL in England.

The company subdivide first line treatment into limited stage and advanced stage disease. The CS states that for limited stage FL, involved-site radiotherapy (ISRT) is the international standard and should be offered to patients with limited stage FL when their tumour can be covered within a radiotherapy field. The recommended dose is 24 Gy in 12 daily fractions. (Taken from The British Society for Haematology Guideline).

For advanced stage FL, the company state that rituximab monotherapy may be considered as a treatment option but that it *"does not currently have a UK marketing authorisation for this indication and is also not commissioned by NHS England in this setting".*[24]

The company suggest that when people with FL experience symptoms they should be treated with an anti-CD20 antibody combined with chemotherapy. Noting further, that choice of treatment depends on preferences of the treating clinician and the characteristics of patients.

Relapsed disease

Treatment for FL relapse can include one of the following regimes, followed by maintenance treatment for up to 2 years;

- a combination of 3 or 4 chemotherapy drugs with rituximab
- R-CVP again if the person is in remission for a long time
- a single chemotherapy drug, with or without rituximab
- rituximab monotherapy
- lenalidomide with rituximab

The company consider only two treatments relevant for the treatment of relapsed or refractory (R/R) FL (see Table 4). Rituximab in combination with chemotherapy, followed by maintenance treatment with rituximab (TA137)[25] and rituximab with lenalidomide (R^2) (TA627),[26] recommended as an option for previously treated FL (grade 1 to 3A) in adults (CS Document B page 21-21).

According to NICE, clinical management for relapsed and refractory FL includes:

- NICE technology appraisal (TA) 137 recommends rituximab either alone or in combination with chemotherapy as a treatment option for people with relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma.[25]

- NICE TA 629 recommends obinutuzumab with bendamustine followed by obinutuzumab maintenance monotherapy as an option for treating FL that did not respond or progressed up to 6 months after treatment with rituximab or a rituximab-containing regimen.[27]

- NICE TA 627 recommends lenalidomide with rituximab as an option for previously treated FL (grade 1 to 3A) in adults.[26]

- Consolidation with autologous or allogenic stem cell transplantation can also be offered for people with FL, in second or subsequent remission (complete or partial), who meet the eligibility criteria.

Position of the technology in the pathway

The CS states that "*Mosunetuzumab is proposed for use within NHS England as an alternative to any third- or later-line therapy option and irrespective of transplantation status (i.e., as a bridge to ASCT, in patients relapsing post-[autologous stem cell transplant] ASCT, and in those unsuitable for ASCT).*" (CS Document B page 22).

The EAG could not determine a rationale for third line positioning rather than second line. The EAG clinical expert suggests that the company proposition that there is no accepted standard of care for third line and subsequent FL treatment and that clinical practice in England is highly variable is "probably true". However, the EAG clinical expert suggest that clinicians will be "using a lot of Rituximab / Lenalidomide as third line treatment option as this is now available". The EAG clinical advisor agrees that mosunetuzumab may be used instead of rituximab combined with various chemotherapy regimens, R^2, or obinutuzumab with bendamustine.

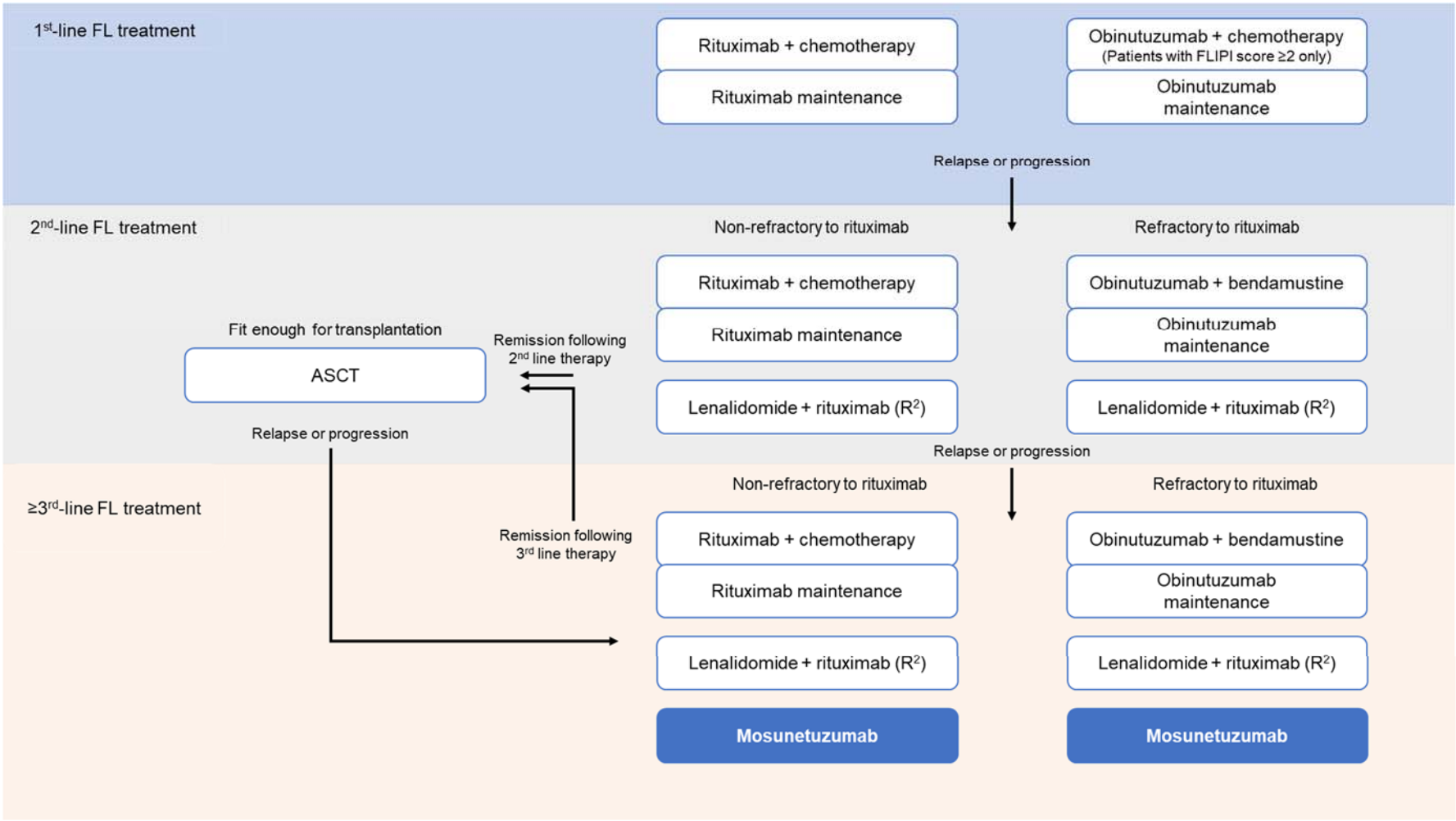


Figure 1. The company's positioning of mosunetuzumab (CS Document B Figure 2)

Abbreviations: ASCT, autologous stem cell transplantation; FL, follicular lymphoma; FLIPI, follicular lymphoma prognostic index.
Refractory to rituximab defined as disease progression on or within 6 months of rituximab treatment as per TA629
Adapted from TA604, TA627, and the British Society for Haematology guidelines on investigation and management of FL

Orphan medicine designation

Monsunetuzumab was designated as an orphan medicine for the treatment of FL in the European Union on 12 November 2021.[21] This means that the company will receive scientific and regulatory support from EMA to advance their medicine to the stage where they can apply for a MA.[18]

At the time of submission, mosunetuzumab did not have MA in the UK. MA was granted by the EMA 23/06/2022 (EMA/318043/2022)[20] in line with the proposed MA included in the CS "*treatment of adult patients with relapsed or refractory (R/R) follicular lymphoma (FL) who have received at least two prior systemic therapies"* (CS Document B page 12).

2.3 *Critique of company's definition of decision problem*

The company have adhered to some of the final scope issued by NICE, however, the population and comparators are narrower that the final scope.
The population is in line with the MA, however, the EAG consider the comparators only partially appropriate. The key deviations are described in Table 4.

Table 4: Summary of decision problem

	Final scope issued by NICE	***Decision problem addressed in the company submission***	***Rationale if different from the final NICE scope***	**EAG comment**
Population	Adults with relapsed or refractory follicular lymphoma	*Adults with relapsed or refractory follicular lymphoma who had received ≥2 prior lines of systemic therapy*	*Clinical data for mosunetuzumab is currently only available in FL patients who had received at least 2 prior systemic therapies. As such, this population is included in the submission, in line with the anticipated marketing authorisation.*	**Appropriate.** The population is narrower than in the NICE Final Scope. However, the population included in the clinical data and CS is in line with the MA "*adults with follicular lymphoma that does not respond to (refractory) or has come back (relapsed) after at least two previous treatments."*[20]
Intervention	Mosunetuzumab	*Mosunetuzumab*	*None*	As per final scope.
Comparator(s)	Established clinical management without Mosunetuzumab. Treatment choice will depend on previous treatments, and how effective those treatments were: • Obinutuzumab with bendamustine followed by obinutuzumab maintenance • Lenalidomide with rituximab • Rituximab in combination with chemotherapy • Best supportive care	• *Lenalidomide with rituximab (R^2)* • *Rituximab in combination with chemotherapy, represented by rituximab + bendamustine (RB)* • *Obinutuzumab with bendamustine followed by obinutuzumab maintenance (OB)*	*BSC was not included, as it could be considered a palliative approach in those FL patients who require treatment. As the age-standardised 5-year net survival rate for FL in England is estimated at 83.0%, a palliative treatment approach (BSC) is therefore unlikely to be of major relevance.*	**Partially appropriate.** R^2 and OB are incline with the NICE Final Scope. The company exclude BSC as they consider it to be a palliative approach. The EAG clinical advisor stated that "BSC equates to watch and wait" which they consider to be "a standard of care for follicular lymphoma." The EAG clinical advisor noted the subjectivity in assessment of BSC as described in the following: "The problem lies with the phrase "*who require treatment*". This is a very subjective assessment. It is often the case that patients who we are considering for treatment are not treated, because there is debate over whether they "*require treatment*" at this point in time. Many times, we do not start, and

	Final scope issued by NICE	*Decision problem addressed in the company submission*	*Rationale if different from the final NICE scope*	EAG comment
				patients are fine. The problem is that there is no definite trigger for treatment at any line in this disease." The EAG suggest that additional clinician expert views could be sought during decision making to address subjectivity around the definition of BSC for this patient population.
			*In the ITC (see Section B.2.9.) and economic model (see Section 3.2.5**Error! Reference source not found.**) rituximab in combination with chemotherapy was represented solely by the rituximab + bendamustine regimen.* *An ITC against R-CHOP was attempted but, despite availability of patient-level data from the EORTC 20981 trial, the analysis proved not to be methodologically feasible.*	The company selected rituximab in combination with chemotherapy to be represented by rituximab + bendamustine (RB), based on data availability for the ITC. The EAG clinical adviser suggest that RB is **not** a good representative for Rituximab+Chemotherapy (R+Chemo). Stating that RB has a particular set of side effects and risks, and these are different from other regimens. Our clinical advisor goes on to suggest that there are differences between patients who receive RB and those who receive other R+Chemo regimens. Including the following; • age • bone marrow resilience, • previous bendamustine use • whether or not stem cell collection is proposed • frailty • and infection risk.

	Final scope issued by NICE	*Decision problem addressed in the company submission*	*Rationale if different from the final NICE scope*	EAG comment
			OB is included for completeness but the company do not consider it to be a relevant comparator considering that clinical expert advice states that this regimen is not commonly used in the third-line setting and beyond, which is also corroborated by market share data which states that 5% of patients receive OB in the third-line setting.	Additionally, the EAG suggests that RB may be used in 1st and 2nd line (notes common international practice), yet it is unlikely that a patient who relapses will be given RB again at 3rd line. With regards to OB, the EAG clinical advisor states that OB is only available in UK for patients who are Rituximab refractory so inclusion at least 2nd line is appropriate.
Outcomes	• Overall survival • Progression free survival • Response rates • Adverse events of treatment • Health-related quality of life	• *Overall survival* • *Progression free survival* • *Response rates* • *Adverse effects of treatment* • *Health-related Quality of Life*	*n/a*	As per final scope.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for	*n/a*	*n/a*	n/a

	Final scope issued by NICE	*Decision problem addressed in the company submission*	*Rationale if different from the final NICE scope*	EAG comment
	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 and cost of biosimilar and generic products should be taken into account.			
Subgroups	None	*Subgroups defined by:* • *Demographics (gender, age range categories, race/ethnicity, ECOG PS),* • *Presence of bulky disease* • *Number of prior systemic therapies and refractory status to those prior treatments* • *FLIPI prognostic score*	*Available data from key subgroups of clinical relevance is presented in the submission.*	**Post hoc subgroups are presented**. Standard methodological bias introduced by post hoc analysis of subgroups needs to be considered. See Section 3.2.3 for EAG critique of subgroup analysis.

Abbreviations: BSC, best supportive care; ECOG PS, Eastern Cooperative Oncology Group Performance Score; FL, follicular lymphoma; FLIPI, Follicular Lymphoma International Prognostic Index; ITC, indirect treatment comparison; OB, obinutuzumab plus bendamustine; R^2, lenalidomide plus rituximab, RB, rituximab plus bendamustine; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone

3 CLINICAL EFFECTIVENESS

3.1 *Critique of the methods of review(s)*

The company conducted a systematic literature review (SLR) to identify evidence for both 'standard of care treatments' and 'novel emerging therapies' for the management of relapsed or refractory (R/R) FL. Detailed descriptions of the methods and findings of the review can be found in Appendix D of the CS, although a pre-defined protocol was neither mentioned nor supplied. Randomised controlled trials (RCTs), non-randomised clinical studies and real-world observational studies were all eligible.

3.1.1 Search strategies

A good range of sources were searched in the search. Including bibliographic databases as well as websites of HTA agencies, reference lists, clinical trial registries and conference proceedings. Unfortunately, search strategies used in the supplementary searches of conference proceedings, HTA agencies and clinical trials registries (CS Appendix D.1, page 20) are not reported. This means that these supplementary searches are neither transparent nor reproducible.

Although there are very few restrictions on study types to be included in the review (CS Appendix D, Table 7), study type filters are included in the Embase and MEDLINE search strategies (CS Appendix D, Tables 1 and 2). This inclusion risks missing some important studies, especially given that some terms for relevant study types are missing, including Emtree terms below 'randomized controlled trial/' in the hierarchy (as this term is not exploded), MeSH headings 'clinical trial/', 'comparative study/' and 'prospective studies/' and free text terms for systematic reviews, controlled trials/studies, and case series.

In CS Document B.2.1 (page 25), a population including "mixed lymphoma" is described in the relaxed eligibility criteria for the MAIC. However, the search strategies for the SLR (CS Appendix D.1) are not necessarily designed to capture all studies in a broader, mixed population. For example, the MeSH and Emtree terms for non-Hodgkin lymphoma are not used and '(relapse* or refact*) adj4 lymph*' is used in the MEDLINE, but not the Embase search strategy, and only searched for in article titles. Therefore, a relevant study in a mixed lymphoma population that did not

mention FL in the title, abstract, keywords or subject headings might not have been found.

Publication dates of 2021-current are used as a date limit in the December 2021 update searches (CS Appendix D, Tables 4-6). However, it is best practice to use date created (Embase) or entry date / create date / entrez date (MEDLINE) fields rather than publication dates, due to the possibility of substantial time lag between an article being published and its complete record being added to a database.

3.1.2 SLR methods critique

A total of 214 publications related to 172 studies were initially included in the review. As the inclusion criteria for the review were set much broader than the NICE scope for this appraisal, some of the identified studies evaluated treatments not listed in the scope, and/or included patients with mixed types of lymphoma and various lines of prior therapy.

The company adopted a hierarchical approach to classifying studies to focus on patient populations most relevant to this assessment (FL with at least 2 prior systemic therapy). The classification of the 172 identified studies based on this hierarchical approach was well-structured and clearly presented (see CS Appendix D, p.28-46). However, insufficient details were provided regarding the subsequent process of narrowing down the number of relevant studies, initially to 42 and subsequently to 21, which were considered in the feasibility assessment for MAIC.

The EAG noted the following areas in which the selection of studies for consideration in the MAIC feasibility assessment appears arbitrary or requires further justifications:

- It is not clear why some studies evaluating treatments not listed in the NICE scope were initially included in the MAIC feasibility assessment and then subsequently discarded.
- For studies which reported data on patient population directly relevant to this assessment (i.e., those shown in green boxes in Figure 4 and listed in Table 8, Table 9 and Table 12 of the CS appendices), it is not clear to the EAG why some of the studies which apparently evaluated treatments directly relevant to this appraisal were not selected for MAIC feasibility assessment (e.g. Trotman 2021,[28] Dai 2021,[29] Fuji 2020[30] from CS Appendices Table 8; Karadurmus

2021[31] from CS Appendices Table 12). Therefore, we undertook additional work on clinical effectiveness detailed in 3.5.

- For studies which included mixed populations (regarding lines of prior therapies and types of lymphoma) covering some patients directly relevant to this appraisal but did not report their data separately, the company considered the following criteria for selection of studies into the MAIC feasibility assessment:
    - (1) The proportion of FL patients >80% for studies with mixed types of lymphoma;
    - (2) median lines of prior therapies ≥2, i.e., ≥50% of patients included in these studies would have had at least 2 prior lines of therapy. However, various exceptions were then applied, rendering the final selection of studies into the MAIC feasibility assessment less systematic and transparent.

The EAG explored the impact of these potential issues by examining studies that may have included useful data but were not selected for MAIC feasibility assessment by the company. The findings are summarised in Section 3.5 of the EAG report.

The company's MAIC feasibility assessment considered 21 studies. Five of these studies directly relevant to this appraisal (GO29781,[2, 12] GADOLIN,[10] CONTRALTO,[13] G029365,[14] EORTC 20981[15, 16] were sponsored by Roche.

As individual patient data (IPD) were available for these studies, they were included in the company's indirect comparison using propensity score matching without further consideration in the MAIC feasibility assessment. From the remaining 16 studies, the company selected three potentially eligible studies for MAIC: AUGMENT,[3] MAGNIFY[32] and Gupta 2021.[8] The remaining 13 studies were excluded as they evaluated treatments not listed in the NICE scope.

All three studies potentially eligible for MAIC identified by the company evaluated the comparator of R^2. They will be described and appraised in Section 3.3.1 of the EAG report.

3.2 *Critique of trials of the technology of interest, the company's analysis and interpretation*

Evidence for mosunetuzumab included in the CS was obtained exclusively from a specific cohort of phase II of an open-label, international clinical study GO29781.[2] A brief description of the overall study design for GO29781 (which included an earlier dose expansion phase; patients with other non-Hodgkin's lymphoma and chronic lymphocytic leukaemia; and mosunetuzumab in combination with atezolizumab) can be found in CS Sections B.2.3.1.

The current assessment, CS and this EAG report focus on the cohort of patients with R/R FL who had at least two prior lines of systemic therapy and who were treated with single-agent mosunetuzumab, as this cohort is directly related to the MA obtained. We refer to participants and data related specifically to this cohort as GO29781 in the rest of this report for brevity.

3.2.1 Critique of methods for GO29781

Details of methods for GO29781 can be found in CS Sections B.2.3 and B.2.4.

The key inclusion criteria for GO29781 were:

- Adult patients with grade 1–3a FL who had relapsed after or failed to respond to at least two prior lines of systemic therapy.
- Eastern Cooperative Oncology Group (ECOG) performance status ≤1.
- Prior treatment with an anti-CD20-directed therapy and an alkylating agent.

Differences in the inclusion criteria between GO29781 and comparator studies with respect to the above items are key issues that need to be considered in the appraisal of indirect treatment comparison, which is detailed in Section 3.4 of this report. We highlight below some issues related to study methods that may have implications for the interpretation of findings from GO29781 and for the comparison with other studies.

Definition and assessment of treatment response

The primary outcome for the study was the proportion of patients whose best overall response was a complete response (CR) based on independent review facility (IRF) assessment. Data on treatment response assessed by individual investigator were

also presented in the CS. Treatment response was assessed by CT and PET-CT using the International Working Group (IWG 2007) criteria.[33] While this is well-accepted and widely used criteria, different assessment criteria were used in some of the comparator studies used in the company's indirect treatment comparisons (ITCs).

- Different assessment criteria has implications for the comparability of outcomes related to treatment response rates and duration of response between studies and increases the uncertainty regarding the validity of relevant ITCs.

Sample size

Power calculation for the study was based on a comparison with a CR rate of 14% among patients with R/R FL observed in a single-arm study of copanlisib.[34]

- The study may not be sufficiently powered to detect clinically important differences in survival, particularly for overall survival (OS).

Patient reported outcomes (PROs)

Patient reported outcomes (PROs) were obtained in GO29781 using EORTC QLQ-C30, FACT-Lym and EQ-5D-5L. The company response to EAG clarification question A14 stated that the data collection "*focused on those patients who were still receiving initial treatment of mosunetuzumab*".

The CS, therefore, only presented PROs up to Cycle 8 of the treatment (approximately 24 weeks post baseline), and the data presented appeared to be based on evaluable patients (patients who had a baseline and at least one post-baseline assessment of PRO scales), not based on intention-to-treat. Although the complete EORTC QLQ-C30 questionnaire was administered, only data related to the physical function and fatigue subscales were presented in the CS. The company explained that this was because these two subscales were considered to be most relevant for this patient population.

- The EAG had no access to, and was unable to evaluate the dataset used to generate HRQoL estimates for the model.
- The EAG cannot rule out potential reporting bias. These should be borne in mind when interpreting findings related to PROs from this study.

Retreatment

Patients treated with intravenous mosunetuzumab followed an 8-cycle regimen with each cycle lasting 3 weeks unless the patient experienced unacceptable toxicity or disease progression. No further treatment beyond 8 cycles was required for patients who achieved CR, although they were eligible for retreatment with mosunetuzumab for at least 8 additional cycles if they subsequently experienced disease progression.

Patients who achieved a partial response or maintained stable disease after receiving 8 cycles of therapy continued to receive mosunetuzumab for up to a total of 17 cycles unless disease progression or unacceptable toxicity occurred. Patients who achieved CR after 17 cycles of treatment were also eligible for retreatment if they subsequently experienced disease progression.

Data related to retreatment were not included and were not considered in the CS but were supplied by the company in their response to clarification questions. Three of the 90 patients were retreated with mosunetuzumab at approximately 1 year (one patient, 8 cycles) and 2 years (two patients, 2 cycles and 4 cycles respectively) after initial study entry. The company emphasised that retreatment following subsequent progression is not covered in regulatory approval and should not be considered in current appraisal.

- The EAG agrees but considers that as the time for longer-term follow-up accumulates, the retreatment could introduce bias in favour of mosunetuzumab if it has a positive effect on survival of the study cohort given that the costs of retreatments were not considered in the company's cost-effectiveness estimates.

Pseudoprogression

The study protocol also allowed treatment to be continued beyond radiographic progression if pseudoprogression was suspected. Pseudoprogression has been described as "*an increase in size of lesions, or the visualisation of new lesions, followed by a response, which might be durable"*[35] due to the influx of immune cells following the initiation of immunotherapy. It may lead to a false judgement of progression and may be difficult to distinguish from true disease progression. As the time on treatment data used in the company's model were limited to not exceeding PFS, the EAG is concerned that additional treatment cycles that might have been

administered (and corresponding costs incurred) beyond disease progress (that was confirmed later) due to initial suspicion of pseudoprogression might not have been accounted for in the company's cost-effectiveness estimates.

In their response to clarification questions, the company described one case of suspected pseudoprogression based on a review of investigators' comments recorded in the response assessment in the electronic case report forms (eCRF). The case subsequently achieved and maintained CR and therefore, would not have impacted on cost-effectiveness assessment.

- Under-reporting/recording of continued treatment post disease progression due to suspected pseudoprogression cannot be ruled out as the company acknowledged that it was not captured via a dedicated data field in the eCRF.

3.2.2 Characteristics of GO29781 study participants

The GO29781 cohort included 90 patients. Baseline characteristics of these patients are described in CS Table 11 (p.46) and will be juxtaposed with data from comparator studies in Section 3.3 of this report, and therefore, not duplicated here. An issue worth highlighting is that most patients in GO29781 were recruited from North America (n=40 for United States, n=13 for Canada) and Australia (n=17).

- Only two of the 90 patients were from the UK, and this may raise some concerns relating to the generalisability of the findings from the cohort.

3.2.3 Treatment outcomes of the GO29781 cohort

The planned primary analysis for GO29781 was based on a data-cut in March 2021. Findings from the study included in the CS were primarily based on a data-cut in August 2021. The company provided further findings based on a later data-cut in [redacted] with their response to EAG clarification questions. The EAG's critique mainly focuses on findings from the latest data-cut ([redacted]), but some comparison and contrast between data from different data-cuts are provided where relevant. Follow-up of patients for the study is planned to continue until [redacted].

Key effectiveness outcomes from GO29781 are shown in Table 5 below. Based on the August 2021 data-cut, a complete response rate of 60% and best overall response rate of 80% as assessed by independent review facility (IRF) using International Working Group (IWG) Cheson 2007 criteria [33] were achieved. The CS

also presented treatment responses as assessed by investigators. The EAG notes that these figures are very similar to the IRF assessment.

- There is very little difference in effectiveness findings between the August 2021 and [redacted] data-cuts.

Table 5: Key effectiveness outcomes from GO29781 (n=90)

Outcome	**Data-cut:** [redacted]	**Data-cut August 2021**
Duration of follow-up, months	"Additional 5 months"	18.3 (range 2.0 to 27.5)
Treatment response as assessed by independent review facility (IRF)		
Complete response (CR)	[redacted]	60% (49% to 70%)*
Partial response (PR)	[redacted]	20% (12% to 30%)
Stable disease (SD)	[redacted]	8% (3% to 15%)
Progressive disease (PD)	[redacted]	10% (5% to 18%)
Best overall response	[redacted]	80% (70% to 88%)
Duration of response (DOR), median, months	Not reported	22.8 (9.7 to NE)
Remained in CR or PR at 12 months	Not reported	62% (50% to 74%)
Remained in CR or PR at 18 months	Not reported	57%
Remained in CR at 12 months	Not reported	71% (58% to 85%)
Remained in CR at 18 months	Not reported	64%
Survival		
PFS, median, months	[redacted]	17.9 (10.9^ to NE) ^Shown as 10.1 on CS p. 88
PFS at 12 months	[redacted]	58% (47% to 68%)
PFS at 18 months	[redacted]	47% (34% to 60%)
OS, median, months	[redacted]	Not reached
OS at 12 months	[redacted]	93% (88% to 98%)
OS at 18 months	[redacted]	Not reported

Data source: CS Section B.2.6.3 and CS RRFL Updated results report 230622; the percentages have been rounded up to whole numbers as EAG considers the decimal place unnecessary given the relatively small sample size (n=90). Numbers shown in brackets are 95% confidence intervals unless otherwise stated.

*CR rate based on 15 March 2021 data-cut (planned primary analysis) was 58% (47% to 68%).

IRF: independent review facility; NE: not evaluable (due to insufficient events/follow-up); OS: overall survival; PFS: progression free survival

Patient-reported outcomes from GO29781 are summarised in Table 6 below. As mentioned earlier, the data were based on evaluable cases rather than intention-to-treat and focused on patients who were still on treatment. The data suggested that patient-reported outcome measures were largely maintained while patients were receiving treatment. A higher proportion of patients achieved clinically meaningful improvement for fatigue and lymphoma related symptoms than for physical function.

Table 6: Patient-reported outcomes (PROs) from GO29781

Outcome	Data-cut: [redacted]	Data-cut August 2021
Duration of follow-up, months	"Additional 5 months"	18.3 (range 2.0 to 27.5)
EORTC QLQ-C30, evaluable patients at baseline n=82		
Physical function		
Mean change* from baseline	Not reported	-0.3 (SD 19.7), n=68
Achieved* clinically meaningful improvement**	Not reported	12%, n=68
Fatigue		
Mean change* from baseline	Not reported	-1.1 (SD 28.4), n=68
Achieved* clinically meaningful improvement**	Not reported	46%, n=68
FACT-Lym, evaluable patients at baseline n=81		
Mean change* from baseline	Not reported	1.9 (SD 8.9), n=67
Achieved* clinically meaningful improvement***	Not reported	42%, n=67
EQ-5D-5L, evaluable patients at baseline (n=81 for utility, n=78 for VAS)		
Mean change from baseline		
Utility – self-care	Not reported	-0.01 (SD 0.7)
Utility – usual activities	Not reported	-0.2 (SD 1.3)
Utility – pain/discomfort	Not reported	-0.2 (SD 0.8)
VAS score	Not reported	4.2 (SD 22.0), n=65

Data source: CS Section B.2.6.4 and CS RRFL Updated results report 230622; the percentages have been rounded up to whole numbers as EAG considers the decimal place unnecessary given the relatively small sample size (n=90). Numbers shown in brackets are 95% confidence intervals unless otherwise stated.
* At completion or discontinuation of mosunetuzumab treatment.
** An improvement of ≥10 points on the scale
*** An improvement of ≥3 points on the scale

Serious and commonest adverse events reported in GO29781 are presented in Table 7. Cytokine release syndrome (CSR), also known as infusion reaction, was the commonest adverse event but most were grade 1-2 events (AE as per Lee 2014 criteria and as per ASTCT 2019 criteria).

[redacted] were the commonest adverse events of grade ≥3.

Table 7: Serious and commonest adverse events reported in GO29781

Adverse events (n=90)		
Outcome	**Data-cut:** [redacted]	**Data-cut August 2021**
Duration of follow-up, months	"Additional 5 months"	18.3 (range 2.0 to 27.5)
Death	[redacted]	9%
SAE	[redacted]	47%
AE of Grade 3-4	[redacted]	70%
Treatment related SAE	[redacted]	33%
AE leading to dose modification	[redacted]	Dose modification or interruption 38%
AE leading to dose interruption	[redacted]	
Treatment withdrawal due to AE or death	[redacted]	4%
Cytokine release syndrome (CRS) by Lee 2014 grade by ASTCT 2019 grade (Grading systems)	[redacted]	46% 44%
Fatigue	Not reported	37%
Headache	Not reported	31%
Neutropenia/neutrophil count decrease	Not reported	29%
Pyrexia	Not reported	29%
Hypophosphataemia	Not reported	27%

Data source: CS Section B.2.10 and CS RRFL Updated results report 230622; the percentages have been rounded up to whole numbers as EAG considers the decimal place unnecessary given the relatively small sample size (n=90). Numbers shown in brackets are 95% confidence intervals unless otherwise stated.

AE: adverse events; CRS: cytokine release syndrome; SAE: serious adverse events

Subgroup analysis

Findings from the subgroup analysis for CR and overall response rates based on IRF assessment are presented in CS Section B.2.7, Figure 12. Outcomes are consistent between subgroups defined by most patient characteristics and prognostic factors.

The EAG noted potential differences in the proportion of patients achieving CR (and to a lesser extent, in overall response rates) between subgroups defined by [redacted] as summarised in Table 8. No tests for interaction were presented in the CS.

- While these subgroup analyses are exploratory and are likely to be under-powered, they provide some indication that [redacted] might be an important effect modifier. To what extent these factors might influence relative treatment effects between mosunetuzumab and comparator therapies, and how this may be related to effects on OS are unknown.

Table 8: Selected findings from subgroup analyses of GO29781 showing possible subgroup effects

Subgroups	Complete response (CR)	Overall response rate (ORR)
Refractory to last prior therapy		
Yes (n=62)	52% (39% to 65%)	77% (65% to 87%)
No (n=28)	79% (59% to 92%)	86% (67% to 96%)
Refractory to any prior anti-CD20 therapy		
Yes (n=71)	55% (43% to 67%)	77% (66% to 87%)
No (n=19)	79% (54% to 94%)	89% (67% to 99%)
Received prior rituximab & lenalidomide		
Yes (n=8)	25% (3% to 65%)	75% (35% to 97%)
No (n=81)	63% (52% to 74%)	80% (70% to 88%)
Double refractory (to rituximab & alkylating therapy)		
Yes (n=48)	50% (35% to 65%)	71% (56% to 83%)
No (n=42)	71% (55% to 84%)	90% (77% to 97%)
[redacted]	[redacted]	[redacted]

3.3 *Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison*

3.3.1 Comparator: rituximab plus lenalidomide (R^2)

The company included three trials in their MAIC feasibility assessment for this comparator: AUGMENT,[3, 4] MAGNIFY[5-7] and Gupta 2021.[8, 9] Key characteristics of these studies are presented in CS Appendices Table 16. As IPD from these trials were not available, the feasibility of undertaking MAIC is heavily dependent on baseline patient characteristics being reported specifically for patients with FL. After the company's feasibility assessment, only the AUGMENT trial was included in its indirect treatment comparison (ITC).

Key features of the three trials and EAG's critique are described below.

AUGMENT

AUGMENT is a phase III, multicentre, double-blind, parallel group randomised controlled trial (RCT) comparing rituximab plus lenalidomide versus rituximab plus placebo in patients with relapsed/refractory indolent lymphoma, which includes FL (n=147) and marginal zone lymphoma (n=31). Only patients with FL who received rituximab plus lenalidomide were relevant to ITC. Outcome data for patients with FL were reported in the trial's publications,[3] but very limited baseline information specifically for patients with FL was reported. The company stated that "data from AUGMENT provide both baseline characteristics and results for RR FL patients" (CS Appendices p.79) and hence implied that the ITC was conducted using FL-specific baseline data, although no reference was cited for this statement and the EAG was unable to verify the source for the FL-specific baseline data from other references cited in the CS in relation to the AUGMENT trial.

Treatment for rituximab plus lenalidomide lasted for twelve 28-day cycles unless relapse or progress disease, withdrawal of consent or unacceptable toxicity occurred. The treatment regimens were as follows: oral lenalidomide 20 mg daily (10 mg for

creatinine clearance 30 to 59 mL/min) on days 1 to 21 plus intravenous rituximab 375 mg/m^2 days 1, 8, 15, and 22 of cycle 1 and day 1 of cycles 2 to 5 every 28 days.

The main trial publication for AUGMENT reported that patients were recruited from 97 centres in 15 countries [3] while 147 sites (located in USA, Brazil, Puerto Rico China, Japan, Israel, Turkey Russian Federation, and European countries including 4 sites in the UK) were listed on its clinical trial registry.[36] It is not clear whether or how many patients with FL were recruited from the UK.

The main methodological features and patient baseline characteristics for the AUGMENT trial is shown in Table 9 alongside those of GO29781, MAGNIFY and Gupta 2021. The inclusion criteria for AUGMENT reflect its target population (R/R FL patients who were non-refractory to rituximab and who would otherwise be suitable to be retreated with rituximab monotherapy). By contrast, 79% of patients in GO29781 were refractory to a prior anti-CD20. AUGMENT trial only required patients with at least one prior systemic therapy (compared with ≥2 for GO29781), and as a result 57% of patients in the R^2 arm had only one prior therapy (compared with none in GO29781). These major differences suggested that patients in GO29781 were likely to be more refractory and were recruited at a later stage in the treatment pathway compared with patients in the AUGMENT trial. As these factors could not be adjusted in the company's MAIC, the company suggested this introduced important bias against mosunetuzumab.

- The EAG broadly agrees with this but consider findings from MAIC to be highly uncertain given the major differences in the populations between the two trials (see Section 3.4).

MAGNIFY

MAGNIFY is a multicentre open-label, 2-part study started with an induction phase in which all participants recently 12 cycles of R^2. Patients were then randomised to receive either maintenance lenalidomide plus rituximab or rituximab alone for a further 18 cycles. The lenalidomide plus rituximab group may receive further cycles of lenalidomide alone until disease progression. The study

included patients with R/R FL grade 1-3a, marginal zone lymphoma (MZL) and mantle cell lymphoma (MCL). Outcomes related to R/R FL patients have been reported in conference abstracts,[7, 32, 37] but only limited information concerning baseline characteristics for a mixed patient cohort including R/R FL (n=318) and MZL (n=76) was presented and baseline characteristics specifically for R/R FL patients were not reported. The company further cited lack of OS data, difficulty in digitizing KM data (source of the KM curves not specified), inclusion of patients with ECOG PS 2 and with no prior systemic therapy as reasons for not performing MAIC using data from MAGNIFY trial (CS Appendice D, p.79-80). The trial registry of MAGNIFY listed 114 study locations predominantly in the USA. No UK centre appears to have been included.

While the EAG considers the exclusion of the study from MAIC to be reasonable given the circumstance, although data from the study could potentially be useful if further baseline and outcome data for RR F/L population are published / made available. It is also worth noting that the R^2 treatment regimens in AUGMENT and MAGNIFY were similar enough such that a pooled analysis of data from the two trials was undertaken and presented in lenalidomide's European Public Assessment Report (EPAR).[38]

- The EAG note that inclusion of data from MAGNIFY would also alleviate the lack of patients who were refractory to rituximab in AUGMENT, as MAGNIFY trial included these patients.

Gupta 2021

This is a single centre, single-arm study in patients with FL (n=30) and other types of indolent B-cell and mantle cell lymphomas (n=20) who were refractory to rituximab or a rituximab-containing regimen. Similar to MAGNIFY, baseline data specifically for patients with FL were not reported, which impedes the use of data from this study for MAIC. Nevertheless, given the median number of prior therapies (3, range 1-7), this study may be better aligned with GO29781 in terms of participants' treatment history.

Long-term data including OS and PFS up to 10 years of follow-up have been reported for this trial, and the EAG presents these data in Section 3.5 to facilitate the validation of survival curve extrapolation.

- Table 9 below highlight key differences in inclusion criteria, methods, and baseline patient characteristics between GO29781, AUGMENT, MAGNIFY and Gupta 2021 studies.

Table 9 Methods and patient characteristics between studies considered for MAIC between mosunetuzumab and rituximab plus lenalidomide (major differences between GO29781 and AUGMENT are highlighted in bold)

Methods and baseline characteristics	**Mosunetuzumab G029781[2]**	**Rituximab and lenalidomide**		
		MAGNIFY[7, 32, 37]	**AUGMENT[3, 4]**	Gupta 2021[8, 9]
Sample size (R/R FL)	90	318	147	30*
Trial registration	NCT02500407	NCT01996865	NCT01938001	NCT00783367
Masking	Open label	Double-blind	Double-blind	Open label
Duration of follow-up (months)	**Median 18.3 (range 2.0-27.5)** **August 2021 data-cut**	Median 40.6 (range 0.6 to 79.6)	**Median 28.3**	10.5 years
Number of previous therapies, median (range) and proportion of patients with a number of prior lines, n (%)	**3 (2 to 10)**	2 (0 to 8)	**1 (1 to 12)** **≥ 2 prior lines of therapy, 69 (47%)**	3 (1-7) [for all patients]
Double refractoriness to both an anti-CD20	**48 (53%)**	85 (22%)	**0 (0%) [assumed based on inclusion criteria]**	NR

containing regiment and an alkylating agent, n (%)				
Refractory to rituximab, n (%)	**71 (79%)** **(Refractory to anti-CD20 mAb therapy)**	140 (36%) (Progressive or stable disease or a response lasting < 6 months after last rituximab dose)	**0 (0%)**	24 (48%)
Refractory to last prior therapy, n (%)	**62 (69%)** **(Failure to respond to previous treatment or progression within 6 months)**	NR	**26 (18%)**	NR
Early relapse status (progression of disease within 24 months), n (%)	**47 (52%)** **(Progression of disease within 24 months from start of 1L therapy)**	133 (34%) (Progressing or relapsing within 24 months of initial diagnosis)	**56 (38%)** **(Relapse/progression of disease within 24 months of initial treatment)**	NR
Duration of prior response/time in previous remission/time since completion of last therapy, median (range)	203.5 (15 to 2,717) days since last therapy	NR	≤ 2 years: 77 (52.4) 2 years: 70 (47.6) (Time since last anti-lymphoma therapy)	NR
Prior ASCT, n (%)	19 (21%)	31 (8%)	NR	6 (12%)
Bulky disease, n (%)	>6 cm: 31 (34%) >7 cm: 16 (18%) >10 cm: 2 (2%)	>7 cm: 161 (41%) (>7 cm or three lesions >3 cm)	>7 cm: 39 (27%) (One lesion ≥7 cm or three lesions with ≥3 cm in longest dimension)	≥5 cm: 17 (34%)

Size of largest lymph node, median (range) and or mean (SD) mm	49.5 (18 to 153) 52 (23)	NR	NR	NR
FLIPI risk group (low (0-1), intermediate (2) , high (3-5)), n (%)	Low: 26 (29%) Intermediate: 24 (27%) High: 40 (44%)	NR	Low: 45 (31%) Intermediate: 46 (31%) High: 54 (37%)	NR
Age, median (range) and/or mean (SD)	60 (29 to 90) 60 (11.95)	66 (35 to 91)	62 (26 to 86) 61.63 (SD 11.31)	57.5 (35 to 85)
Disease stage, n (%) (Ann Arbor stage)	**I: 5 (6%)** **II: 16 (18%)** **III: 25 (28%)** **IV: 44 (49%)**	I–II: 66 (17%) III: 98 (25%) IV: 229 (58%)	**I: 13 (9%)** **II: 21 (14%)** **III: 69 (47%)** **IV: 44 (30%)**	NR
High LDH, N (%)	**35 (39%)**	NR	**34 (23%)**	14 (28%)
Bone marrow involvement	**4 (4%)**	123 (31%)	**20 (24%)** **Intermediate involvement: 2 (1%)**	14 (28%)
Haemoglobin level (low: <12 [or 12.5] g/d)	37 (41%)	NR	NR	NR
B symptoms, n (%)	13 (14%)	NR	12 (8%)	NR
ECOG PS, n (%)	0: 53 (59%) 1: 37 (41%)	0: 193 (49%) 1: 192 (49%) 2: 9 (2%)	0: 99 (67%) 1: 47 (32%) 2: 1 (2%)	0-1: 49 (98%) 2: 1 (2%)
Modified from CS Appendix D, Tables 16 & 17 *Baseline characteristics reported for Gupta 2021 included mixed types of indolent B-cell and mantle cell lymphomas (n=50), within which 30 patients had follicular lymphoma.				

3.3.2 Comparator: obinutuzumab plus bendamustine

The company selected the GADOLIN trial [10] for the comparison against obinutuzumab plus bendamustine (OB). As this trial was also sponsored by Roche, the company has IPD for this study, which allowed ITC to be carried out using propensity score matching. GADOLIN is a phase 3, multicentre, open-label, randomised trial comparing OB to bendamustine monotherapy in patients with indolent non-Hodgkin lymphoma refractory to rituximab. Only the OB arm of the trial was used in ITC.

The dosing regimen for the OB arm was:

Obinutuzumab 1000 mg (days 1, 8, and 15, cycle 1; day 1, cycles 2-6) plus bendamustine 90 mg/m² per day (days 1 and 2, cycles 1-6).

Patients who did not experience disease progression received obinutuzumab maintenance (1000 mg every 2 months) for up to 2 years.

The trial was conducted in 14 countries across Europe, Asia, and North and Central America, including five centres in the UK recruiting 29 patients, although it is not clear how many of these were patients with FL in the OB arm.

Baseline characteristics for patients with FL were not separately reported in the main publication for GADOLIN trial.[10]

Table 10 below compares the baseline characteristics of patients between GO29781 and OB arm of the GADOLIN trial (mixed population), and the characteristics of patients selected from the respective cohort after a “filtering” step removing patients with clearly different characteristics due to differences in the inclusion criteria in respective trial. The data show that the original cohort in GO29781 had more lines of prior therapies while a higher proportion of the original patients in the GADOLIN trial were refractory to the prior line of therapy and to both rituximab and alkylating therapies.

The filtering process appears to have improved the comparability in terms of refractory to therapies, but substantial difference remains with regard to prior line of therapy and the filtering seems to increase the difference in FLIPI high risk category between the two cohorts. It was therefore necessary for the company to use the inverse probability of treatment weighting (IPTW) method to further adjust for the remaining major differences. EAG's critique of the IPTW is presented in Section 3.4.2.2.

Table 10 Comparison of patient baseline characteristics between GO29781 and GADOLIN trial

Patient characteristics	**Mosunetuzumab – original GO29781 cohort (n=90)**	[redacted]	[redacted]	**OB – original trial arm including mixed population (n=194)[10]**
Age (mean)	60.0 (29-90)	[redacted]	[redacted]	63 (55-69)
FL grade 1 2 3a Unclassified	 NR NR NR NR	 NR NR NR NR	 NR NR NR NR	 36% (54/150)* 46% (69/150)* 16% (24/150)* 2% (3/150)*
ECOG PS (1 vs 0) (%)	 41% (37/90)	[redacted]	[redacted]	 NR
FLIPI (%) Low (0-1) Intermediate (2) High (≥3)	[redacted]	[redacted]	[redacted]	 28% (42/149)*

				32% (47/149)* 40% (60/149)*
Ann Arbor Stage (%) I II III IV	6% (5/90) 18% (16/90) 28% (25/90) 49% (44/90)	[redacted]NR NR [redacted]	[redacted]NR NR [redacted]	NR NR NR NR
Prior lines of anti-lymphoma therapies Median (range)	3.0 (2-10)	NR	NR	NR
Prior therapies (any) 1 2 ≥3 (%)	[redacted]	[redacted]	[redacted]	47% (92/194) 32% (62/194) 21% (40/194)
Prior rituximab-containing therapies 1	[redacted]NR NR NR	[redacted]	[redacted]	56% (108/194)

2 ≥3 (%)				32% (62/194) 12% (24/194)
Refractory to last line (Yes) (%)	69% (62/90)	[redacted]	[redacted]	92% (178/194)
Double refractory (yes) (%)	53% (48/90)	[redacted]	[redacted]	76% (147/194)
POD24 (Yes) (%)	52% (47/90)	[redacted]	[redacted]	NR
Prior ASCT (Yes) (%)	21% (19/90)	[redacted]	[redacted]	NR
Size of the largest node lesion [cm] (mean)	NR	[redacted]	[redacted]	NR
Low Hb (Yes) (%)	[redacted]	[redacted]	[redacted]	NR
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]	NR
Bone marrow involvement (Yes) (%)	[redacted]	[redacted]	[redacted]	32% (60/187)

Extranodal involvement (%)	NR	NR	NR	58% (107/183)
Bulky disease (>6 cm)	[redacted]	[redacted]	[redacted]	34% (66/194)
Presence of B symptoms (Yes) (%)	[redacted]	[redacted]	[redacted]	NR
Time since completion of last therapy)	>2 years [redacted]	>2 years [redacted]	>2 years [redacted]	Mean 4.0 (2.5-7.0)
Time from initial diagnosis to randomisation (years)	NR	NR	NR	4.2 (1.7-5.4)
Modified from CS Table 11, Table 14 *Patients with follicular lymphoma				

3.3.3 Comparator: rituximab plus bendamustine

The company included data from two trials, GO29365 and CONTRALTO, in the ITC against the comparator of rituximab plus bendamustine. Key features of these two trials and EAG's critique are presented below. The company had access to IPD for these two trials, which allowed the use of propensity matching for ITC.

GO29365

GO29365 is a multicentre, open-label, randomised controlled trial comparing polatuzumab vedotin, bendamustine, rituximab, and obinutuzumab in people with R/R FL or diffuse large b-cell lymphoma. Data from the FL participants have been published only in the form of a conference abstract.[14] This provides little information about the trial. In brief, 80 transplant-ineligible with relapsed/refractory FL were randomised 1:1 to polatuzumab vedotin (1.8mg/kg) with bendamustine and rituximab [bendamustine: 90mg/m^2 x 2 days; rituximab 375 mg/m^2) or bendamustine and rituximab (dose as above) for 6 cycles. At the time of publication, participants had been followed up for a median of 15 months.

No further details of the study methods are provided in the conference abstract. The trial inclusion/exclusion criteria (as report in the trial registration for GO29365) are shown in Table 11. The overlap between GO29365, CONTRALTO, and GO29781 trials is limited to

(1) participant age (≥18 years),

(2) the presence of at least one bi-dimensionally measurable lesion on imaging,

(3) adequate hematological function, and

(4) histologically confirm relapsed/refractory follicular lymphoma (grade 1, 2, or 3a).

However, only GO29781 explicitly stated that participants must be R/R to a minimum of 2 prior treatments, so it not possible to know how comparable the trials are on this inclusion criterion. Most inclusion and exclusion criteria are reported for only one of the trials. This leads to considerable uncertainty about the comparability of trial populations for ITC, although the company were able to use data restricted to patients with R/R FL who had at least two prior systemic therapy with ECOG 0-1 in the ITC due to availability of IPD.

The GO29365 trial registration lists 62 sites in 13 countries (USA, Australia, Canada, Czechia, France, Germany, Hungary, Italy, Republic of Korea, Netherlands, Spain, Turkey, and the UK).[39]

- It is unclear whether any participants from the UK were included in the trial. Therefore, the EAG are unable to confirm how representative these trials are to UK populations and NHS service provision.
- In general, the trial maps onto the NICE Final Scope. Though response rates and health-related quality of life are not reported in the conference abstract for trial GO29365.
- It is not clear whether the trial population maps to the population as described by the company (adults with relapsed or refractory follicular lymphoma who had received ≥2 prior lines of systemic therapy) as prior lines of therapy were not specified.

CONTRALTO

CONTRALTO was a multicentre, open-label phase 2 trial comprising a safety-run in followed by assignment to 3 treatment arms that compared venetoclax with rituximab (VEN + R), venetoclax with bendamustine and rituximab (VEN + BR), and bendamustine with rituximab (BR) in participants with relapsed/refractory follicular lymphoma. In brief, 154 participants were initially assigned to

either chemotherapy-free (Arm A: VEN + R) or chemotherapy-containing (Arm B: VEN + BR; Arm C: BR) cohorts according to investigator discretion. Participants in the chemotherapy-containing cohort were then randomised 1:1 to either VEN+BR or BR using stratified permuted block randomisation. Stratification was according to duration of response before therapy (≤12 months/>12 months) and disease burden (high/low, according to modified Groupe d'Etude des Lymphomes Folliculaires criteria). Drug regimens were as follows:

Arm A (VEN + R):

VEN - 800mg, oral, daily for one year

R - 375 mg/m^2 intravenous infusion on days 1, 8, 15, and 22 of cycle 1 and day 1 of cycles 4, 6, 8, 10, and 12. Twenty-eight day cycles.

Arm B (VEN +BR):

VEN - 800mg, oral, daily for one year

R - 375 mg/m2 intravenous infusion on days 1, 8, 15, and 22 of cycle 1 and day 1 of cycles 4, 6, 8, 10, and 1.2 Each cycle is 28 days

B - 90 mg/m^2 intravenous infusion on days 1 and 2 of each 28-day cycle, for 6 cycles.

Arm C (BR):

R - 375 mg/m^2 intravenous infusion on day 1 of each 28-day cycle

B - 90 mg/m^2 intravenous infusion on days 1 and 2 of each 28-day cycle, for 6 cycles.

The drug regimen was identical to that of trial GO29365.

The study inclusion and exclusion criteria are shown in Table 11. As stated in the corresponding section for trial GO29365, there is uncertainty about whether the trial participants are comparable.

The CONTRALTO trial registration lists 71 sites in 8 countries (USA, Australia, Belgium, Canada, France, Germany, Italy, and the UK).[11]

- It is unclear whether any participants from the UK were included in the trial. As stated before, generalisability to a UK population cannot be determined.
- In general, the trial maps onto the NICE Final Scope. Though health-related quality of life is not reported for the CONTRALTO trial.
- The extent to which the trial population maps to the population as described by the company (adults with relapsed or refractory follicular lymphoma who had received ≥2 prior lines of systemic therapy) is not clear; while the median number of trial treatments was 2, this ranged from 1 to 4.

Baseline characteristics of the three trials

Baseline characteristics from the three trials are presented in Table 12.

All three trials appear to be comparable in terms of median participant age. The only other variable for which data are available for all three trials is the proportion of participants refractory to last treatment. This was somewhat higher in GO29781 (68.9%) than G029365 (42%) or CONTRALTO (45.1%).

- Overall, there is little overlap in information presented for the three trials, preventing the EAG from making an accurate assessment of the comparability of trial participants prior to the company's initial matching (filtering) through aligning inclusion/exclusion criteria between the trials (which reduced the sample sizes for monsunetuzumab from 90 to [redacted] and for bendamustine plus rituximab from 92 to [redacted]). Substantial differences remained for several baseline characteristics between the two cohorts after the filtering process (see CS Table 16, p.75), and therefore the company attempted both propensity score matching and IPTW to adjust for the remaining differences. EAG's critique of these indirect comparisons is presentation in Section 3.4.2.1.

Table 11 Inclusion criteria of GO29781, G029365 and CONTRALTO trials

	GO29781	G029365	CONTRALTO
≥18 years	Yes	Yes	Yes
Histologically confirmed relapsed or refractory FL (Grades 1, 2, or 3a) or DLBCL	Yes	Yes	Yes
If prior bendamustine, DOR > 1 year (for pts with relapse disease after a prior regimen)	NR	Yes	Yes
≥one bi-dimensionally measurable lesion on imaging (>1.5 cm in its longest dimension)	Yes	Yes	Yes
Confirmed availability of archival or freshly collected tumor tissue	NR	Yes	Yes
Life expectancy of at least 24 weeks	No	Yes	NR
Life expectancy of at least 12 weeks	Yes	No	NR
ECOG = 0, 1, 2	Partial*	Yes	Yes
Adequate hematological function unless inadequate function is due to underlying disease	Yes	Yes	Yes
Adequate coagulation, renal, and hepatic function	NR	NR	Yes
≥1 prior FL therapy	Yes	NR	Yes
Relapsed/refractory to at least 2 prior treatments	Yes	NR	NR
Prior treatment with anti-CD20 therapy and an alkylating agent	Yes	NR	NR

* The ECOG inclusion criteria for GO29781 is stated as ≤1

Table 12 Baseline characteristics of GO29781, G029365 and CONTRALTO trials

	GO29781	G029365 (BR only)	CONTRALTO (BR only)
	n = 90	n = 41	n = 51
Transplant ineligible?	NA	Yes	NR
BR cycles	NA	6 28-day cycles	6 28-day cycles
BR dose	NA	B: 90mg/m^{2}x 2 days R: 375mg/m^2	B: 90 mg/m^2 IV on days 1 and 2 R: 375 mg/m^2 IV on day 1
Median age	60 years (29 – 90)	63 years	61 years (35 – 80)
≥ 65 yrs	28 (31.1%)	NR	22 (43.1%)
Male	55 (61.1%)	NR	30 (58.8%)
Time since diagnosis, median months	[redacted]	NR	NR
Lymph node ≥ 10 cm	NR	NR	7 (13.7%)
Ann Arbor stage	N = 90	NR	N = 51
I	5 (5.6%)	NR	4 (7.8%)
II	16 (17.8%)	NR	10 (19.6%)
III	25 (27.8%)	NR	7 (13.7%)
IV	44 (48.9%)	NR	30 (58.8%)
ECOG			N = 50
0	53 (58.9%)	NR	34 (68%)
1	37 (41.1%)	NR	16 (32%)
2	0	NR	0
BMI median	27.5 (17 – 45)	NR	NR
Bulky disease (>6 cm)	[redacted]	NR	NR
FLIPI			
0,1 (low)	[redacted]	NR	NR
2 (intermediate)		NR	NR
3 – 5 (high)		15 (37%)	NR
FL grade 3a	NR	NR	9/50 (18%)
Bone marrow infiltration		NR	N = 49
Yes	25 (27.8%)	NR	13 (26.5%)
No	NR	NR	35 (71.4%)
Unknown	NR	NR	1 (2%)
Extranodal involvement			
Yes	NR	NR	27 (52.9%)
No	NR	NR	24 (47.1%)
Prior therapies			
Min-max	NR	NR	1-4
Median	NR	2	2

Prior anti-lymphoma therapies			
Median	3 (2 – 10)	NR	NR
2	[redacted]	NR	NR
3		NR	NR
>3		NR	NR
Prior cancer therapies			
anti-CD20	90 (100%)	NR	NR
alkylating agent	90 (100%)	NR	NR
auto-SCT	19 (21.1%)	NR	NR
CAR-T	3 (3.3%)	NR	NR
PI3K	17 (18.9%)	NR	NR
Refractory			
to last treatment	62 (68.9%)	17 (42%)	23 (45.1%)
any prior treatment	[redacted]	NR	NR
any prior anti-CD20	71 (78.9%)	NR	NR
prior anti-CD20 + alkylating agent	48 (53.3%)	NR	NR
Refractory to rituximab	NR	NR	21 (41.2%)
Disease progression within 2 years of 1st therapy	47 (52.2%)	NR	NR
Duration of prior response			N = 50
≤ 12 months	NR	NR	26 (52%)
>12 months	NR	NR	24 (48%)
Disease burden (GELF)			N = 51
Low	NR	NR	17 (33.3%)
High	NR	NR	34 (83.7%)
BLC-2 IHC			N = 43
Negative (0-1)	NR	NR	7 (16.3%)
Positive (2-3)	NR	NR	36 (83.7%)
BLC-2 FISH			N = 33
Negative	NR	NR	3 (9.1%)
Positive	NR	NR	27 (81.8%)
Undetermined	NR	NR	3 (9.1%)
BLC-XL			N = 42
IHC score ≥2,3	NR	NR	14 (33%)
MCL1 IHC			N = 38
IHC score ≥2	NR	NR	2 (5%)

BR = bendamustine with rituximab, NA = not applicable, NR = not reported

3.3.4 Comparator: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP)

The company selected EORTC 20981 for providing data on rituximab (R) in combination with cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) as a comparator (R-CHOP). EORTC 20981 was a phase 3, open-label, randomised controlled trial that compared an initial 6 cycles of standard CHOP with R-CHOP in patients with CD20+ grade 1-3 R/R FL. Patients who achieved a CR or PR following the initial 6 cycles of either treatments underwent a second randomisation to rituximab maintenance therapy (intravenous 375 mg/m2 once every 3 months until relapse or for a maximum of two years) or no further treatment. Only patients initially randomised to R-CHOP arm (n=234) were relevant to the ITC.

However, it was unclear whether all patients who underwent second randomisation were included, or whether only patients who were subsequently allocated to rituximab maintenance therapy at second randomisation were included. As rituximab maintenance therapy was shown to be more effective than no maintenance therapy and this has become a standard practice, inclusion of patients receiving no maintenance therapy after second randomisation would have under-estimated the effectiveness of R-CHOP therapy.

The trial recruited patients from 130 centres across Canada, Australia/New Zealand, Europe (including UK), and South Africa. Major differences in patient inclusion criteria between GO29781 and EORTC 20981 included Ann Abor state at diagnosis (no limit vs stage III or IV) and prior therapy (≥2 prior lines including an anti-CD20-directed therapy and an alkylating agent vs ≤2 non-anthracycline-containing systemic chemotherapy), prior rituximab treatment (required vs not allowed) and EGCO PS (≤1 vs ≤2). These resulted in a small proportion of patients randomised to the R-CHOP arm remaining in the selected patient cohort ([redacted]) after initial matching/filtering, with important imbalance between the selected mosunetuzumab ([redacted]) and R-CHOP patients ([redacted]). Further attempts to adjust for imbalance using genetic matching or Inverse probability of treatment weighting (IPTW) methods described in an ITC report appended to CS resulted in even smaller effective sample sizes for both treatments and did not remove major imbalance in some important

prognostic factors. The company therefore considered "[redacted]" (CS ITC report, page 66).

- The EAG agree with this assessment, and therefore analyses based on data from EORTC 20981 are not described further in this EAG report. However, the lack of suitable clinical effectiveness data for the comparison with R-CHOP (and consequently lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP), represents a major gap in this assessment given that R-CHOP is a common treatment combination used in third line setting for R/R FL in the UK.[1]

3.4 *Critique of the indirect comparison and/or multiple treatment comparison*

Given the absence of any head-to-head trials, the company performed several indirect comparisons as it sought to obtain estimates of benefit of mosunetuzumab relative to the chosen comparators. These indirect comparisons aimed to produce comparable populations of patients between the different trials, reducing the potential impact of any differences in baseline characteristics. Whilst these approaches can reduce bias, their results are not as reliable as results coming from a phase III two arm trial. In this Section, the EAG first examine and critique the methods used by the company, and later examine the clinical outcomes estimated from the indirect comparisons.

- Overall, the EAG considered the estimates of relative effectiveness of mosunetuzumab compared with comparator interventions obtained from the ITCs to be highly uncertain and potentially biased due to issues related to sparsity of data and methodological limitations as acknowledged in the company's ITC report (page 10).

3.4.1 Matched Adjusted Indirect Comparison

3.4.1.1 Comparison vs Rituximab Lenalidomide (R^2)

For the comparison to R^2, an unanchored MAIC was performed as patient level data was unavailable for the R^2 population from the AUGMENT study. The variables included in the MAIC are shown in Table 13.

Table 13: Variables used in each of the indirect comparisons base-case analyses.

High Risk Variables as identified by the company	**Outcome**	**Used in MAIC vs R^2**	**Used in PSA vs Obinutuzumab Bendamustine**	**Used in PSA vs Rituximab Bendamustine**
No of prev therapies	3 vs >3 or median	No	Yes (≥3)	Yes (≥3)
Refractory to prev therapy	Progressed/relapsed vs No	Yes	Yes	Yes
Refractory to prior aCD20	Yes vs No	No	No	Yes
Early relapse status (POD24)	Yes vs No	Yes	Yes	Yes
Prior ASCT	Yes vs No	No	Yes	Yes
Size of largest lymph node lesion	Mean	No	Yes	Yes
Bulky disease	Yes vs No	Yes	No	No
FLIPI	<3 vs >=3	Yes	Yes	Yes
Age	Mean	Yes	Yes	Yes
Ann Arbor stage	1-2 vs 3-4	Yes	Yes	Yes
High lactate dehydrogenase	Yes vs No	Yes	Yes	Yes
Bone marrow involvement	Yes vs No	Yes	Yes	Yes
Low haemoglobin	Yes vs No	Yes (value imputed)	Yes	Yes
Low Priority Factors				
Duration of prior response or Time since last therapy	?	No	Yes (time since last therapy, mean)	Yes (time since last therapy, mean)
Presence of B symptoms	Yes vs No	No	Yes	No
ECOG PS	1 vs 0, or 2 vs <2	No	Yes (1 vs 0)	Yes (1 vs 0)
Other Factors				
Double refractory			Yes	Yes
Interaction of Age and No of prior therapies				Yes

Interaction of Age and Refractory to prev therapy				Yes
Interaction of Refractory to prev therapy and POD24				Yes

It is notable that the company has not been able to include several high priority variables in the MAIC, meaning there is the potential for imbalance and bias between these populations. A further concern is the company's imputation of 'Low Hgb level' for AUGMENT, using the population average from GO29781. It is unclear why the company chose this approach rather than excluding it from the analysis, as was done for other high priority variables that were not available for matching, such as number of previous therapies, size of largest lymph node, prior ASCT. The company have not provided any evidence to support this imputation, and so the EAG requested analyses that excluded low Hgb level from the matching analysis. This was provided and is critiqued in a later section.

The company were unable to match using the variable 'Refractory to Previous AntiCD20 therapy' as these patients who were refractory were excluded from the AUGMENT study, whilst ~80% of patients from GO29781 were refractory, and so including this as a matching variable would have decreased the effective sample size (ESS) considerably prior to any further matching being implemented. The EAG accept this rationale for excluding this variable. The company conclude that this biases the analysis in against mosunetuzumab as being refractory to rituximab is a high priority variable. The uncertainty surrounding the other unmatched variables means it is not possible to conclude which way the analysis may be biased, but that it is clear tremendous uncertainty remains over the balance of these two populations and this should be carefully considered when interpreting the results.

For variables with missing data, the ITC report does not state how these observations were handled but refers to R code that has not been made available to the EAG. Poor management of missing values can introduce bias into an analysis.

The company performed some exploratory analyses investigating the inclusion of the low-priority factors.

The only additional analysis the company do present some details for is one which includes all three low-priority variables, reducing the ESS to [redacted]. This scenario analysis was associated with a lower OS and similar PFS to the company's preferred analysis (Figures 1-4 of Company ITC report). An examination of the weights calculated in the company's base case MAIC does not show any further cause for concern.

- The EAG is unable to conclude that the company's base case MAIC analysis should be considered the most appropriate possible analysis given the higher ESS associated with one of the company's unreported exploratory analyses, and the unusual decision to impute for one of the missing covariates.
- A limitation of the analysis, and the later implementation of these results in the economic model, are that the MAIC is estimating the efficacy of mosunetuzumab in a population of the AUGMENT study, which may be less representative of the patients who may receive mosunetuzumab under this indication than the original population of GO29781.

3.4.1.1.1 Interpretation of base case results

This section explores the relative treatment effect of mosunetuzumab using the weighted data coming from the company's base case MAIC analysis. An overview of the results is presented in Table 14. Table 14 of the original company submission shows that the nine covariates included in the MAIC analysis are well matched at population level, however, one of these is the imputed value for 'Low Hgb level'. It is unclear why the company have been unable to report the post-matching values for three variables that were included in their priority list and in Table 14, but not in the MAIC: ECOG, Time since completion of last therapy, and presence of B symptoms all have "NA".

There remains the possibility that these variables are not well balanced between the weighted mosunetuzumab and R^2 populations and are biasing the comparisons. The company initially only updated PFS and OS outcomes for the new [redacted] data-

cut of GO29781, but after EAG request they provided updated response and discontinuation outcomes.

Table 14: Overview of results from MAIC analysis vs R^2

	Unadjusted estimate (95% CI)	**Weighted estimate (95% CI)**	**Weighted Bias Corrected Bootstrap 95% confidence interval, p value**
PFS Hazard Ratio	[redacted]	[redacted]	[redacted]
OS Hazard Ratio	[redacted]	[redacted]	[redacted]
Complete Response Odds Ratio	[redacted]	[redacted]	N/A
Overall Response Odds Ratio	[redacted]	[redacted]	N/A
Discontinuation due to AEs Odds Ratio	[redacted]	[redacted]	N/A

[redacted]

[redacted]. Figure 2 and Figure 3 show

[redacted].

[redacted]

Figure 2: PFS for R^2 and mosunetuzumab after applying the MAIC weights (taken from Company Addendum Figure 3)

Figure 3: OS for R^2 and mosunetuzumab after applying the MAIC weights (taken from Company Addendum Figure 6)

Visually assessing the response outcomes, the adjusted data show a

The company has not presented a comparison of the duration of response (DOR) of the therapies for any data-cut, as data for FL population of AUGMENT were unavailable. The EAG has provided a naïve comparison which overlays the Kaplan-Meier (KM) plots for the primary DOR results reported from each trial (Figure 4). Note that the AUGMENT trial included patients with marginal zone lymphoma who are included in this figure.

.

Figure 4: Overlayed Kaplan-Meier plots comparing duration of response from AUGMENT and GO29781

- The inconsistency between the response related outcomes and the PFS/OS outcomes may be caused by confounding by differences in the populations, perhaps from latent variables or those not included in the analysis.
- These results combined leave no clear conclusion as to whether mosunetuzumab offers any meaningful clinical benefit to patients compared to R^2.

3.4.1.1.2 EAG additional requested analysis

The EAG requested the company reperform the MAIC analysis excluding the 'Low Hgb level' as its value for the R^2 population was imputed but no other missing variables had their value imputed. Examination of the weights and covariates shows this MAIC analysis is well performed and has an ESS of 35.32 for mosunetuzumab. The lower maximum weight and larger ESS relative to the company-preferred analysis suggests excluding 'Low Hgb level' is beneficial to the analysis.

The effects of excluding the variable have an effect on the hazard ratio estimates for PFS and OS (Table 15), as well as on the KM plots. In both Figure 5 and Figure 6, the comparison of interest is the red mosunetuzumab arm on the left plot to the green line appearing on the right.

[redacted].

- No other outcomes from this comparison were reported, so the impact on response rates and safety is unclear.

Table 15: Comparison of results from weighted Cox models from MAIC analyses against R^2

	Including imputed low Hgb, preferred by company	Excluding low Hgb
PFS Hazard Ratio (95% Confidence Interval) p value	[redacted]	[redacted]
OS Hazard Ratio (95% Confidence Interval) p value	[redacted]	[redacted]

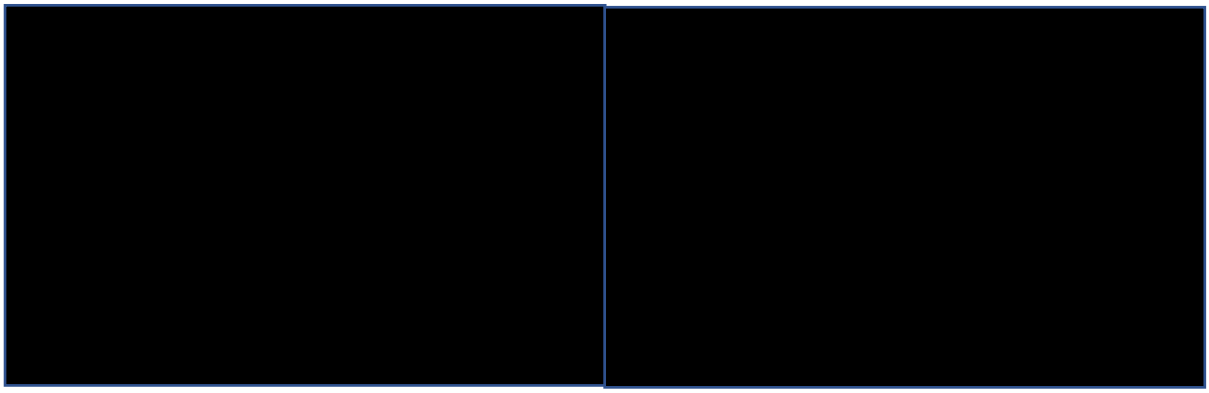


Figure 5: MAIC PFS showing impact of excluding imputed value for low Hgb from MAIC

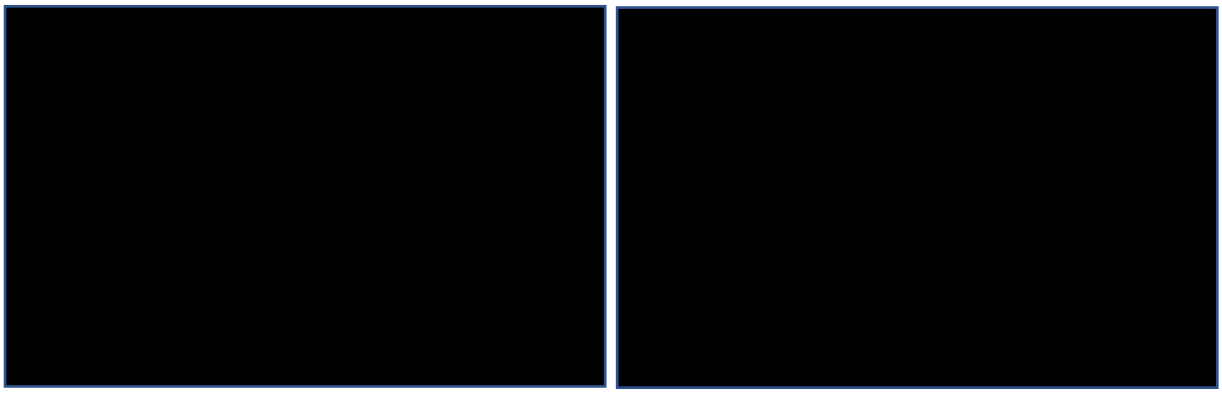


Figure 6: MAIC OS showing impact of excluding imputed value for low Hgb from MAIC

3.4.2 Propensity Score Comparisons

The comparison of mosunetuzumab to both obinutuzumab plus bendamustine, and rituximab plus bendamustine uses propensity score-based analyses, which the EAG will describe briefly. This approach is possible as the company had access to patient level data for all the relevant trials, and it is generally thought to be an improvement over a MAIC. Propensity scores can be used to estimate the probability of a patient being in one trial based on their baseline characteristics. These scores can then be used in a matching analysis or as weights to, in theory, obtain a population balanced on propensity score and so also the underlying baseline characteristics.

The company explored a range of matching approaches in an attempt to obtain an optimal balanced population. These were optimal pair matching, nearest neighbour matching, genetic matching, and full matching. The best performing matching method was then compared against inverse probability of treatment weighting, which weights participants based on their propensity score rather than applying any matching algorithm. Typically, consideration of the ESS and a comparison of the covariates of the resulting matched/weighted populations will suggest the optimal analysis approach.

- As with the MAIC analysis, the ITC report does not state how variables with missing data were managed but refers to R code that has not been made available to the EAG.
- The EAG note that poor management of missing values can introduce bias into an analysis and further add to the uncertainty in the company's cost-effectiveness estimates.

3.4.2.1 Comparison vs Rituximab plus Bendamustine

For the comparison to rituximab plus bendamustine, the company compared data from patients in the GO29781 trial to data from a combination of the CONTRALTO and GO29365 trials. Aligning the inclusion/exclusion criteria of the trials reduced the samples sizes to 81 and 46 patients respectively.

When describing the propensity score model specification algorithm, the company's ITC report states that all possible two-way covariate interactions were generated and

tested for inclusion. The report later says that three specific interactions were included following consultation with the company's internal medical experts.

- The EAG is concerned that these interactions terms did not warrant inclusion by the routine selection, and so did not meaningfully improve the comparison. It is therefore unclear why they have been specified for inclusion at this stage as important prognostic factors when they were not previously specified before the analysis was conducted.
- The decision to include them is potentially subjective and may influence the final matched population and resulting health outcomes.

The interaction terms included by the company are:

- Age and number of prior therapies
- Age and refractory to last line of therapy
- POD24 and refractory to last line of therapy

The EAG's clinical expert stated it is unlikely that that there is a scientific basis for these interactions.

In addition, the company included the 'Double Refractory' variable, which was identified as of "unclear priority". This variable also has the potential to be correlated with the other variables based on refractory status already included in the propensity score calculation, effectively prioritising them in the matching analysis. No justification for its inclusion is provided, and it may be detrimental to the analysis.

Optimal pair matching was reportedly the best performing matching method and was compared with IPTW. This is despite an individual in the matching analysis initially being given a weight of 26 which was later reduced to 10 to reduce the individuals influence on the analysis, however it is possible all matching analyses encountered similar problems. Large weights can give individuals undue influence on the analysis introducing bias and their presence are a weakness of the associated indirect comparison.

The ITC report recommends the IPTW analysis which has an ESS of [redacted] for mosunetuzumab and [redacted] for RB. This is smaller than the ESS for optimal pair

matching, which had an ESS of [redacted] for both interventions, but IPTW yielded superior covariate matching, hence IPTW was recommended.

The company, however, have instead chosen to implement optimal pair matching across their subsequent clinical and cost-effectiveness analyses. The reason for this inconsistency with the ITC report is not reported which is a cause of concern.

It is also unclear how the company has inflated the ESS for mosunetuzumab from [redacted] as shown in Figure 15 of the ITC report, however the EAG predicts this may be from stabilisation of the IPTW weights. If different weightings have been used for different outcomes, this is major limitation to the company's analyses.
Due to the remaining imbalance of patient characteristics from using the optimal pair matching, the company reports adjusting for certain imbalances in the subsequent analyses of health outcomes.

The EAG infers from the ITC report that the covariates adjusted for in these post-hoc analyses are: time since completion of last therapy, refractory to last line of therapy, ECOG PS, prior ASCT, size of the largest node lesion, Ann Arbor stage III/IV, and refractory to prior aCD20.

- The full output for the models showing these covariate effects was not provided, and the EAG cannot be certain that the final model used to estimate the treatment benefit is appropriate and sensible for decision making.
- In addition, there is uncertainty over the specification of the covariates included in the calculation of the propensity scores and company's choice of model, leaving large concern over the appropriateness of the analysis put forward by the company.

The EAG present the resulting covariate matching of the two approaches in Table 16, where it is clear that IPTW results in better matching, with the absolute standardised mean difference between fewer variables being above the 0.1 threshold. Hence, the company's text in their original submission which justifies the use of optimal pair matching: "*Optimal pair matching resulted in the greatest number of balanced covariates compared with other methods and was selected as the*

preferred adjustment method for this comparison", is factually inaccurate. Hence, for the remainder of the critique the EAG will present results from both analyses where possible.

Table 16: Comparison of the IPTW and Optimal Pair matching for the RB comparison

	IPTW		Optimal Pair	
Matching Variable	Mosunetuzumab	RB	Mosunetuzumab	RB
Age (mean)				
ECOG PS (1 vs 0)				
FLIPI (≥ 3)				
Ann Arbor (3-4)				
Prior Therapies (≥ 3)				
Refractory to last line				
Refractory to any prior aCD20				
Double refractory				
POD24				
Bone Marrow Involvement				
Prior ASCT				
Size of largest node lesion (mean)				
Low Hgb				
High LDH				
Time since completion of last therapy				
BOLD indicates absolute standardised mean difference between variables after matching is > 0.10				

- The EAG has concerns over the method of selection of covariates included in the calculation of propensity scores, and over the statistical models use to estimate effect size, which may mean that the results provided by the company fail to accurately represent the true relative effect of mosunetuzumab and RB.

3.4.2.1.1 Interpretation of base case results

This section explores the relative treatment effect of mosunetuzumab using the weighted data coming from the company's base case propensity score analysis. An overview of the results is presented in Table 17.

All outcomes have been updated for the [redacted] data-cut.

Both the adjusted and unadjusted comparisons all suggest that there is [redacted], showing the uncertainty but this is unsurprising given the large number of covariates included in the model with a relatively small sample size.

- The EAG recommends estimation of effect sizes excluding all covariates from the final model and relying on the populations as balanced by the propensity score matching/weighting.

Table 17: Overview of results for adjusted and unadjusted comparison of mosunetuzumab to BR

	Unadjusted estimate (95% CI)	**IPTW based estimate (95% CI) [post-hoc adjusted]**	**Optimal Pair Matching estimate (95% CI) [post-hoc adjusted]**
Sample Size	81 + 46	81 + 42	46 + 46
PFS Hazard Ratio	[redacted]	[redacted]	[redacted]
OS Hazard Ratio			
Complete Response Odds Ratio			
Overall Response Odds Ratio	[redacted]	[redacted]	[redacted]
Discontinuation due to AEs Odds Ratio	[redacted]	[redacted]	[redacted]

AE, adverse event; CI, confidence interval; IPTW, inverse probability treatment weighting; PFS progression-free survival; OS, overall survival

A comparison of Figure 7 to Figure 9 suggest [redacted]. The [redacted] appears to be heavily influenced by the post-hoc covariate adjustment, rather than the matching analysis. The hazard ratio for the propensity score based analyses without post-hoc covariate adjustment is likely to be similar to that of the unadjusted analysis. RB appears to

have a [redacted] in the optimal matching analysis, compared with the IPTW and unadjusted analyses.

Comparing Figure 10 to Figure 12 shows a similar story for PFS, [redacted], regardless of adjustment. They suggest the hazard ratio for the unadjusted analysis would be similar to that of the other PFS analyses if they did not include the post-hoc covariate adjustment.

Figure 7: Unmatched output for OS comparison of mosunetuzumab vs RB

Figure 8: IPTW output for OS comparison of mosunetuzumab vs RB

Figure 9: Optimal pair matching output for OS comparison of mosunetuzumab vs RB

Figure 10: Unmatched output for PFS comparison of mosunetuzumab vs RB

Figure 11: IPTW output for PFS comparison of mosunetuzumab vs RB

Figure 12: Optimal pair matching output for PFS comparison of mosunetuzumab vs RB

3.4.2.2 Comparison vs Obinutuzumab plus Bendamustine

For the comparison of mosunetuzumab to obinutuzumab plus bendamustine (OB), the company compared patients from trial GO29781 to those from the GADOLIN trial. Aligning the inclusion/exclusion criteria of the trials reduced the samples sizes to 71 and 77 patients respectively.

The covariate selection algorithm for this comparison was identical to the propensity score analysis against RB. The company again included the 'Double Refractory' variable, but this time their internal medical experts did not recommend including any covariate interaction terms (Table 13). The company also excluded the 'Refractory to prior aCD20 therapy' variable because only patients who were refractory were included in the population for this comparison.

On this occasion IPTW was recommend by the independent ITC report and was also the analysis presented by the company to estimate the clinical benefit of mosunetuzumab over OB.

- IPTW was associated with a higher ESS and better covariate balance than the optimal matching method, and so the EAG agree it appears the most suitable approach.

The IPTW weights gave an ESS of [redacted] patients for mosunetuzumab and [redacted] for OB. It is again unclear how the sample sizes for mosunetuzumab and OB are inflated to [redacted] and [redacted] respectively (Figure 15 and 16 of original company submission), but the EAG predicts this is from the stabilisation of the IPTW weights. If alternative weights have been used for different outcomes, this would be a major concern. Overall, the covariates included in the analysis are well-balanced and the distribution of weights are acceptable.

3.4.2.2.1 Interpretation of base case results

Table 18 shows the estimates of clinical benefit both before and after applying the IPTW weights. Due to the balance achieved by the IPTW, no post-hoc adjustment for variables was made in the model fitting. In both the unadjusted and IPTW analyses of PFS,

[redacted]. [redacted]. It is unclear why there [redacted]. The company state that this is because the GADOLIN trial allowed maintenance therapy of obinutuzumab for up for 2 years, however, [redacted].

- Given the limited explanation, it raises the question as to whether there is an error in the analysis meaning any following inference should be disregarded.
- The EAG considers there are possible differences in the prior and subsequent therapies received in the two trials, given that GADOLIN was initiated in 2010, which may explain the unusual results.

Figure 13 to Figure 16 show little impact of the application of the IPTW weights compared to the unadjusted survival functions across PFS and OS, however the difference in any subsequent extrapolations may be more important.

Table 18: Unadjusted and adjusted outcomes for the comparison of mosunetuzumab and OB

	Unadjusted estimate (95% CI)	**IPTW based estimate (95% CI)**
Sample Size	71 + 77	61.2 + 58.0
PFS Hazard Ratio	[redacted]	[redacted]
OS Hazard Ratio		
Complete Response Odds Ratio		
Overall Response Odds Ratio	[redacted]	[redacted]
Discontinuation due to AEs Odds Ratio	[redacted]	[redacted]

AE, adverse event; CI, confidence interval; IPTW, inverse probability treatment weighting; PFS progression-free survival; OS, overall survival

Figure 13: Unmatched output for OS comparison of mosunetuzumab vs OB

Figure 14: IPTW output for OS comparison of mosunetuzumab vs OB

Figure 15: Unmatched output for PFS comparison of mosunetuzumab vs OB

Figure 16: IPTW output for PFS comparison of mosunetuzumab vs OB

3.5 *Additional work on clinical effectiveness undertaken by the EAG*

As described in Section 3.1.2, the reasons for excluding some of the studies initially included the company's MAIC feasibility assessment were not clearly reported. The EAG examined these studies in further detail and the findings are briefly described below:

- Trotman 2021:[28] this was a conference abstract reporting a prospective, multicentre study from Australia on R^2 treatment among patients (n=17) with relapsed FL who remained PET-CT positive after reinduction rituximab-chemotherapy.

- Dai et al. 2021:[29] this was a conference abstract reporting a retrospective study using real world evidence from the United Stated on treatment regimens and clinical outcomes for patients with FL treated with 3rd line therapy.

- Fuji et al 2020:[30] this was a paper reporting a retrospective multicentre study from Japan on R/R FL patients who received 3rd line treatment.

While the EAG agree that it would not be feasible to carry out ITCs based on studies reported in these publications due to limited data on baseline patient characteristics and/or mixed treatments received by patients, treatment response and survival data reported in studies may still be valuable for model validation purpose given the

substantial uncertainties associated with the company's ITCs and very limited data on longer-term follow-up in GO29781 and comparator studies included in the ITCs.

The EAG has therefore, collated data from these studies alongside additional data identified through citation tracking of references presented in the CS and a focused search of survival data in R/R FL third line setting. These data are presented in Table 19.

Table 19 Additional treatment response and survival data from studies covering 3rd line treatment in patients with R/R FL

Study & location	Treatment	Prior lines of therapy, median (range)	Refractory to prior therapy	CR	ORR	Survival (median, 95% CI), months	6 m	1 yr	2 yrs	3 yrs	5 yrs	10 yrs	Note
Gupta 2021 [8, 9] USA N=30	R^2 (10 mg regimen)	3 (1-7)	Refractory to rituximab	30% (9/30)	57% (17/30)	PFS 16.5 OS NR	96% NR	74% NR	33% NR	7% NR	7%** 60%	** 40%	Patients were treated between 2008-2012
Trotman 2021 [28] Australia N=17	R^2 (10 mg regimen)	2 (2-11)	Remained PET-CT positive after reinduction of rituximab-chemotherapy	NR	NR	PFS 30.8 (5.7 to 37.6) OS 68.1 (9.6 to NE)	71% 88%	71% 76%	58% 70%	37% 63%	9% 52%	NR NR	Included stage IV and ECOG 2 patients (% not reported)
Dai et al. 2021 [29] USA N=687	Clinical practice* (74% rituximab-based regimens)	2	NR	NR	NR	PFS 12.5 (11.3 to 14.4) OS	NR NR	NR 83%	NR 75%	NR NR	16% 61%	NR NR	
Fuji et al. 2021 [30] Japan N=41	Clinical practice* (49% bendamustine-based regimen)	2	Received rituximab-based therapy as first-line	42% (16/38)	54% (22/41)	PFS 19.3 (5.4 to 34.1) OS 56.5 (21.8 to NE)	65% 92%	54% 87%	39% (23-56%) 66% (48%-79%)	28% 66%	NR 48%	NR NR	

*Treatments included rituximab, bendamustine and rituximab, phosphatidylinositol 3-kinase (PI3K) inhibitors (copanlisib, duvelisib, idelalisib), lenalidomide and rituximab (R^2), tazemetostat, and stem cell transplant.**Both 5-year and 10-year PFS was reported as 13% in Gupta et al 2021.[8] However this is incompatible with earlier data from the same study published in Chong et al. 2015,[9] which reported PFS at 5 years as 7%. NE: not estimable, NR: not reported

3.6 *Conclusions of the clinical effectiveness section*

In summary, the company conducted a reasonable SLR to identify evidence for both 'standard of care treatments' and 'novel emerging therapies' for the management of relapsed or refractory (R/R) FL (Section 3.1).

The EAG considered the estimates of relative effectiveness of mosunetuzumab compared with comparator interventions obtained from the ITCs to be highly uncertain (see Section 3.3) due to issues related to sparsity of data and methodological limitations as acknowledged in the company's ITC report (page 10), in particular:

- Misalignments across GO29781 and comparator studies in terms of inclusion/exclusion criteria.
- Small sample sizes and short duration of follow-up
- Unavailability of information on some important prognostic factors
- Important residual imbalances after adjusting for differences in the available factors.

The EAG have strong concerns over implementation of MAIC analysis (see Section 3.4) including:

- inconsistent management of missing values for MAIC analysis vs R^2 (company have imputed one value but not others)
- the selection of covariates in preferred MAIC analysis (ESS has not been maximised)
- and a failure to report comprehensive post-matching summary of baseline characteristics.

We have concerns over implementation of propensity score analysis in comparison to RB (see Section 3.4.2) including:

- unclear rationale for inclusion of terms in propensity score modelling in comparison to RB (terms were added by consultancy company at request of Roche team).

We have concerns over output of propensity score analysis in comparison to RB (see Section 3.4.2.1) including:

- going against recommended propensity score analysis for RB comparison using optimal pairs instead of IPTW,
- and inadequate reporting of models used to estimate treatment effect.

4 COST EFFECTIVENESS

4.1 *EAG comment on company's review of cost-effectiveness evidence*

The CS provides detailed reports of three systematic literature reviews (SLRs), aimed at identifying economic evaluations (Appendix F) health state utility values and health-related quality of life data (Appendix G) and studies describing costs and resource use (Appendix H), all relating to the follicular lymphoma population.

4.1.1 Search strategies

An appropriate selection of bibliographic databases was searched, supplemented by hand-searching of reference lists, conference proceedings and websites of international HTA bodies. The searches for the costs and resource use SLR were limited to studies published from 2011 onwards (CS Appendix H.3), though the review eligibility criteria (CS Appendix H.4, Table 42) state that publications since 2001 will be included.

Further websites, databases and search engines including Cost-effectiveness Analysis Registry, EconPapers/RePEc, EuroQoL website, INAHTA and Google Scholar were also searched. However, search strategies used in these additional sources are not reported (despite some of them supporting relatively complex search strings), meaning the supplementary searches are neither transparent nor reproducible.

In the reported search strategies (CS Appendices F.3, G.3 and H.3), sensitive search strings for follicular lymphoma are combined with sensitive filters for the different study types of interest. The EAG therefore considers the searches unlikely to have missed any relevant studies.

There appears to be an error in the EBM Reviews January 2022 update search for the health-related quality of life SLR (Appendix G.3, page 149), where some search strings have been included in square brackets and therefore retrieve 0 results. However, as EBM Reviews is not a major source of recent economic evidence, it is unlikely to have affected the review outcome.

4.1.2 Inclusion/exclusion criteria used in the study selection

The inclusion and exclusion criteria for the review of cost-effectiveness evidence, health state utility values, and costs and resource are presented in Table 29 of Appendix F.4, Table 36 of Appendix G.4, and Table 42 of Appendix H.4.

- The EAG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies.

4.1.3 Included/ excluded studies in the cost-effectiveness review

A total of 32 publications, reporting 19 published analyses and 13 HTAs were included in the cost-effectiveness review. Details of these studies are provided in Table 21 and Table 22 of the company submission. Two studies were excluded in the cost effectiveness SLR (Appendix F) (Hughes et al., 2006;[40] Wake et al., 2002[41]), at full text review stage because they were "unobtainable" (see Table 35, Company Appendix F.6). However, these studies were easily retrievable, with one being a HTA review that is freely accessible online.

- The EAG does not feel that exclusion of these studies biases the results as one study on rituximab is quite dated and it is unlikely that the information extracted from the study would have informed the current appraisal.

The search for health state utility values resulted in six included studies, of which four were presented as conference abstracts and two were presented as full texts (Table 37; Appendix G.6). The CS stated that the search for cost/resource use data resulted in 13 publications (5 presented as full-text and 8 as conference abstracts). However, this contradicts information presented in Appendix H.5 (see also Figure 8: PRISMA flow diagram in Appendices). In the appendix the company stated that in total, 17 publications reporting cost/resource use data were identified. Of these, six were presented as full publications and 11 as conference abstracts only.

- The EAG note an error in reporting in Document B which is inconsistent to Sections H.5 H.6 of the CS appendix. In the CS, the company provide detail about the 13 studies in B.3.5.1. Information about additional identified in the updated search (Jan 2022) is missing from this description.

4.2 *Summary and critique of the company's submitted economic evaluation by the EAG*

The eligibility criteria were suitable for the SLR performed. The SLR search strategies were comprehensive enough despite some limitations highlighted above. However, it was not clear to the EAG how information obtained from the SLR for health state utility values was used to inform the de novo analysis. The company states in appendix G.6 that "*three studies fully met the requirements of the NICE reference case*" and proceeds to state that "*these studies are likely to be considered most appropriate for informing economic evaluations in an English/UK setting*".

However, only utility values from Cognet et al.,[42] were used in some way to determine the impact of alternative health state utility sources on ICER (Table 19; Company addendum). The other source of utilities explored in the scenario analysis was Wild et al.,[43] the company does not explain or justify its choice of Wild et al despite previously stating that Wild et al. study was excluded in their SLR on the following basis: (i) utility values from this study are unverifiable (not included in the study abstract itself) and; (ii) their use has been criticised in previous appraisals by EAGs (Appendix G.6, page 164).

4.2.1 NICE reference case checklist

The EAG assessment against the NICE reference case checklist is presented in Table 20

Table 20: 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	Yes
Perspective on costs	NHS and PSS	Yes
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes

Element of health technology assessment	Reference case	EAG comment on company's submission
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes
Synthesis of evidence on health effects	Based on systematic review	Yes
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.	Yes
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes. The standard UK tariff (based on EQ-5D-3L) crosswalk index utilities used
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
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	Yes
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes

PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.

4.2.2 Model structure

The company constructed a *de novo* cost-utility model using partitioned survival with a weekly cycle length and a 40-year time horizon. The model defines three health states: progression free survival (PFS), progressed disease (PD) and death (absorbing health state) (Figure 17) All patients entered the model in the PFS state and remained there until disease progression or death.

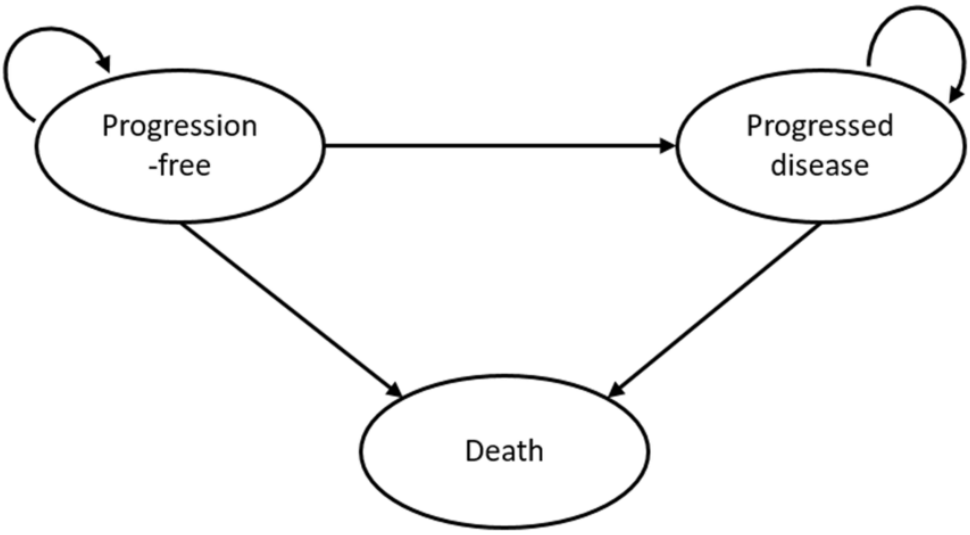


Figure 17 Model structure

Source: Figure 21 - Company Submission

The partitioned survival method uses “area under the curve” approach, where the number of patients in each health state at a given time point is taken directly from survival curves fitted to the clinical data. The PFS curves indicate, for each time point, the proportion of patients who have not progressed or died whilst the OS curves indicate the proportion of patients who are alive at a given time point. The proportion of patients in the PD state was calculated as the difference between the proportion of living patients (OS health state) and the proportion of patients who were both living and pre-progression (PFS health state). The OS and PFS curves were determined by fitting parametric models to the GO29781 data. For the comparator arms, OS and PFS distribution were determined by fitting parametric models to reconstructed KM curves from relevant comparator studies.

Time to off-treatment (TTOT) in the PFS state for the mosunetuzumab arm was determined by KM data on treatment duration, limited to not exceed PFS. For comparators, TTOT in the PFS was set equal to the chosen parametric distribution for PFS and capped at treatment-specific maximum number of cycles. No treatment waning effect was applied in the model, justified on the basis that majority of patients taking mosunetuzumab had completed their treatment within the observed period and any treatment waning effect would benefit mosunetuzumab over comparators. A half-cycle correction was applied to all health states distributions including the TTOT

data, from week 1 (second cycle) rather than week 0 (first cycle). The model was programmed in Microsoft Excel.

EAG Comments

- Even though the model is simple, with three health states, it captures two important clinical endpoints of PFS and OS and is consistent with models built in this disease area and used in previous technology appraisals.
    - PSM seemed appropriate for the modelling the decision problem.
- The weekly cycle length was short enough to capture changes over the relevant time interval.
- The 40-year time horizon was long enough to capture important differences in costs or clinical outcomes as very few patients are modelled as surviving beyond 35 years.
- While the EAG understands the decision by the company not to apply a treatment waning effect in their original submission, the data is very immature, and impact of applying treatment waning effect should be considered in decision-making. In response to clarification question B9 (Table 37), the company explored the impact of treatment waning by applying waning effects for 60 months in line with previous appraisals. Applying the treatment waning effect considerably worsened the ICER for the R^2 and RB comparisons. There was a marginal increase in ICER for the OB comparison.
    - Uncertainties regarding the need of a treatment waning effect are best resolved by long-term outcomes on OS and PFS.
- Half-cycle correction should generally be applied in situations where the timing of the event is not known. Mosunetuzumab was administered at the known timepoints hence a half cycle correction applied to TTOT is considered unnecessary (see Section 1.1).

4.2.3 Population

The patient population considered in the model is in line with the MA: adult patients with relapsed or refractory (R/R) follicular lymphoma (FL) who have received at least two prior systemic therapies (see Section 2.3).

As described in Section 3.2.1 and 3.2.2, the submission mainly relies on one single arm study, the GO29781 trial, a Phase I/II, multi-centre, open-label dose escalation and expansion study. The GO29781 study provided data on the use of, and clinical efficacy, safety, and time on treatment of mosunetuzumab for treatment of adult patients in the intended population.

Baseline patient parameters for the modelled populations were derived from the baseline characteristics of pivotal cohort patients enrolled in GO29781 (i.e., mean age: 60.01 years; baseline body weight: 81.40 kg; baseline height: 169.92 cm; baseline BSA:1.96 m2; and proportion of males in the cohort: 60.0%. (Table 23 CS, v2.0).

- The EAG note that the CS does not explicitly state whether these patient characteristics are reasonably similar to the UK treatment population to provide a valid comparison (see other issue Section 1.6).

For the comparator arms, efficacy estimates for R^2 were drawn from the AUGMENT trial. Efficacy estimates for OB were based on the GADOLIN trial and estimates for RB were based on CONTRALTO and GO29365 trials.

A detailed EAG critique of these trials and comparisons is provided in Section 3.2.

4.2.4 Interventions and comparators

The description of comparators in the NICE scope is as follows: Obinutuzumab with bendamustine followed by obinutuzumab maintenance (OB); lenalidomide with rituximab (R^2); Rituximab in combination with chemotherapy and best supportive care (see Table 4: Summary of decision problem).

The company's base case compares mosunetuzumab with RB, OB and R^2, partly reflecting the description of comparators in the NICE scope. In the CS, best supportive care was not included as a comparator and rituximab with bendamustine

was used as a substitute for rituximab with chemotherapy (see other issue Section 1.6).

Advice from our clinical experts suggest that Rituximab with Bendamustine can be an appropriate substitute for the patient population served by rituximab with chemotherapy. However, rituximab plus bendamustine has a particular set of side effects and risks which are different from other regimens. Furthermore, patients who receive RB are likely to differ from those who receive other R+Chemotherapy regimens in terms of age, bone marrow resilience and whether or not they have had bendamustine before (see Table 4: Summary of decision problem).

EAG Comments:

- The base-case analysis incorporates a comparator (RB) in which the EAG is uncertain is representative on NHS practice following clinical expert advice (see Table 4.

4.2.5 Perspective, time horizon and discounting

The perspective is as per NICE reference case, with benefits from a patient perspective and costs from an NHS and personal social services (PSS) perspective. In the base-case, costs and benefits were discounted at an annual rate of 3.5% in line with NICE reference case used and discount applied is in line with NICE reference case. The time horizon is 40 years which is sufficient to capture the extrapolated OS curves given model cohort age.

4.2.6 Treatment effectiveness and extrapolation

4.2.6.1 Mosunetuzumab Vs Rituximab with lenalidomide (R^2) (MAIC)

This section explores the company's methods of extrapolating the data generated via their MAIC analysis which applied weighting to the population of GO29781 to be comparable to the AUGMENT trial population of who received R^2. The EAG's comments and concerns regarding the MAIC analysis presented by the company are described in detail in Section 3.4.1.

4.2.6.1.1 Progression free survival (PFS)

For PFS, the company fitted parametric models to data from the relevant trial follow-up. The mosunetuzumab data were less mature, with an ESS of [redacted] and maximum follow-up of 28 months, compared to a sample size of 147 and 48 months of follow-up for R^2. In their addendum, the company were consistent with their choice of extrapolation from their original submission and maintained the selection of a log-normal extrapolation for both mosunetuzumab and R^2. This was supported by hazard and cumulative hazard plots which showed a decreasing hazard trend over time. An overview of the company's preferred models is shown in Table 21. The company's choice of models is supported by consideration of fit to data and information criterion. This choice of model results in a large PFS benefit for mosunetuzumab as shown in Figure 18.

Table 21: Company's preferred extrapolations for comparison to R^2

Outcome	Treatment	Company Original Model (Discounted Life-years)	Company New Model (Discounted Life-years)
PFS	R^2	Log-normal ([redacted] years)	Log-normal ([redacted] years)
PFS	Mosunetuzumab	Log-normal ([redacted] years)	Log-normal ([redacted] years)
OS	R^2	Weibull ([redacted] years)	Weibull ([redacted] years)
OS	Mosunetuzumab	Exponential ([redacted] years)	Log-normal ([redacted] years)

However, for most of the observed period, that is, where we know the status of at least one patient, [redacted].

[redacted].

[redacted].
[redacted]

- The EAG recommend modelling the PFS of mosunetuzumab to be equivalent to that of R^2 beyond [redacted].
- This approach would account for the uncertainty around the MAIC analysis and for the uncertainty around the long-term efficacy of mosunetuzumab. [redacted]
- This approach utilises the larger and more mature data from AUGMENT, and avoids extrapolation of data with limited follow-up and small sample, which is ill-advised and unlikely to be reliable.[44] Both R^2 and mosunetuzumab are taken for up to a year, so it is reasonable to assume that their relative efficacies will not differ greatly over time, despite the uncertainty around the comparison of response rates and durations.

Figure 18: The PFS extrapolations as preferred by the company

4.2.6.1.2 Overall survival

For OS, parametric models were fitted to the relevant post-MAIC follow-up. These were selected based on consideration of the fit to data, information criterion and hazard rate shape. The hazard rates for both arms show a decreasing trend towards in the latter half of their observed period, however, due to the small number of events the true underlying hazard behaviour may not be well represented. The company selected the log-normal for the OS extrapolation of mosunetuzumab, considering the reducing hazard rate trend and low AIC and BIC. The OS hazard rate for this and all OS extrapolations in this appraisal was capped so it could not fall below the hazard rate general population mortality.

For R^2 the company has provided inconsistent information in their original addendum as to which is their preferred extrapolation. In table 9 of their addendum, the company report the Weibull model is used, however on page 9 of the same document the company states: "*Of the models fit to R^2 OS, the generalised gamma model was considered to predict implausibly optimistic survival and so was not considered. The observed hazard suggests that it is expected to initially increase and then decrease making the log models, or potentially the Gompertz model the most reasonable. Given the fit statistics, the log-normal model was chosen for the base case*."

The cost-effectiveness results presented by the company use the Weibull extrapolation, and this was confirmed as the company's preferred choice. As can be seen by comparing Figure 19 and Figure 20 the difference between the Weibull and log-normal extrapolations is large, and is consequential on the cost-effectiveness outcomes. The EAG sees no justification for selecting the Weibull model as it directly contradicts the CS text, modelling an increasing hazard rate over time, and neither is it associated with the lowest AIC/BIC.

Due to the small number of events and similarity of the arms, the EAG requested that the company implement an analysis that combines the OS data from both arms. The EAG meant for the post-matching/weighting populations to be pooled, but unfortunately the company appeared to misinterpret the request and pooled the populations without any adjustments.

- Pooled post-matching/weighting populations remains the EAG's preferred approach for modelling OS.

The company did not provide any related goodness of fit information, so the EAG opt to use a log-normal extrapolation of the pooled OS data for their base-case and explore a log-logistic extrapolation as a scenario, as both could be considered equally plausible when fitted separately to either arm.

The company outlined some reservations to pooling the OS data, and the EAG will consider these below:

- The company mention that assuming equal OS across comparator groups in this appraisal may be inconsistent with observed PFS differences, however, this is not relevant to the comparison to R^2 as the PFS data were markedly similar to those of mosunetuzumab.

- The company's other main point is the issue of pooling potentially heterogeneous populations, citing a recent paper,[45] however this paper defined heterogeneity as each population's survival coming from different distributions. In this case, there is no evidence provided to show that the survival times of the populations of mosunetuzumab and R^2 might come from different distributions and so the EAG is unconcerned about pooling for this comparison.

- **EAG Comments:**

In summary, there is no evidence to support the magnitude of benefit of mosunetuzumab that is associated with the company's choice of models for PFS and OS.

In contrast, the EAG's preferred assumptions are entirely consistent with the results of the MAIC analysis.

Figure 19: OS extrapolations preferred by company if selecting Weibull for R^2.

Figure 20: OS extrapolations preferred by company if selecting log-normal for R^2.

Figure 21: EAG preferred OS extrapolation pooling OS data for mosunetuzumab and R^2.

4.2.6.1.3 Time to off treatment

Time to off treatment is the outcome used to estimate the time patients receive active therapy. For R^2, the company assume that the TTOT is equal to PFS but is capped at a maximum of 1 year for lenalidomide (12 cycles), and 19 weeks for rituximab (5 cycles). This is consistent with the stopping rule for the MA of R^2, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic analysis. This may lead to overestimation of the proportion of patients remaining on R^2 treatment, potentially overestimating the costs associated with this treatment.

For mosunetuzumab, the company model TTOT as it was observed in GO29781, which would include discontinuations prior to disease progression for reasons such as withdrawal from the trial. Whilst the extrapolation of the PFS data which acts as an upper bound to TTOT, the TTOT data itself has not been adjusted by the MAIC weighting. The EAG requested the company weight the TTOT data, and the corresponding analysis showed the weights had little impact.
The TTOT approach implemented within the economic model is described as 'KM with Exponential tail', however no information on parametric extrapolations has been

provided to the EAG and it is unclear whether any parametric modelling has been used.

Due to the stopping rule of mosunetuzumab, where patients could stop at 6 months if they had achieved a CR, there is a clear difference between the modelled TTOT of mosunetuzumab and R^2. In practice, if fewer patients achieved CR than was observed in the trial and more patients received treatment for the full 12-month period, then the costs associated with mosunetuzumab therapy would increase considerably. Figure 22 shows the difference between the TTOT modelling of the two treatments, with mosunetuzumab having an average treatment duration of [redacted] months compared to [redacted] months for R^2. The difference in the modelling of TTOT may be introducing bias in favour of mosunetuzumab.

- The EAG will explore a scenario where the TTOT of mosunetuzumab is increased, by increasing the proportion of patients receiving mosunetuzumab beyond 6 months.
- The EAG will also explore removing the half-cycle correction for TTOT as mosunetuzumab was administered at specific time points. Furthermore, the model cycle length is already precise enough to capture the treatment costs accurately (see 1.1 Key Issues).

Figure 22: Comparison of time to off treatment modelling of mosunetuzumab and R^2.

4.2.6.2 Mosunetuzumab vs Rituximab Bendamustine (RB)

This section explores the company's methods of incorporating data from their trials (GO29781, CONTRALTO and G029365) into the economic model. The magnitude of difference in this section generally appears smaller compared to the results presented in the ITC critique, and this is because the post-matching covariates adjustments that were included in the calculation of hazard and odds ratios are not applied here.

In their original submission, the company use unadjusted data for the extrapolations of RB. This would be consistent with the decision either to model using unadjusted data for both arms, or to apply the optimal pair matching results for both arms. However, a comparison of the PFS KM plots in the original company submission across figures 17 and 28 showed inconsistency of the mosunetuzumab data used.

- The data used in the modelling of mosunetuzumab did not use the unadjusted or optimal pair matching adjusted data (Figure 23). Instead, the EAG noted that it was identically to the IPTW adjusted mosunetuzumab data, which is not presented in the main clinical section of the company's report.

- These new data are presented without any introduction or justification for the inconsistency with the results presented in the clinical section.

Figure 23: Comparison of data used in clinical and cost-effectiveness sections, showing inconsistency of mosunetuzumab data (red) used in clinical and cost effectiveness sections with it not matching unadjusted (left) or optimal pair matched (right) data

In the addendum submission, the company again use the unadjusted data from the 46 patients RB patients. For mosunetuzumab, the company maintain their use of the IPTW adjusted data. This inconsistency is performed without introduction or justification which the EAG can verify.

- Using output from different matching sources means the matching is unlikely to be optimised meaning any biases present at baseline are likely to be carried forward into the fitting and extrapolation of survival models.
- The EAG is concerned over the lack of transparency, and the lack of justification to support the company's unusual and varying selection of data to extrapolate. It is our view that the analyses are selected inappropriately and inconsistently. Therefore, the output likely represents an exaggeration of the treatment benefit.

In contrast to the company's approach, the EAG favours extrapolation of data that are consistent with the clinical evidence (as described in Section 3). The EAG

recommends that the time-to-event analyses be reperformed on datasets which incorporate the IPTW weights, and also separately where the optimal matching has been applied.

- The EAG now critiques the survival extrapolation from the company's base case addendum analysis, despite having concerns about its suitability and whether the findings are appropriate for consideration by the committee.

4.2.6.2.1 Progression Free Survival

Table 22 shows the life-year estimates from the company's new approach and from their original submission. The mean time spent in the progression-free health state for RB is now slightly worse in the company's new modelling. The change for RB is attributable solely to the choice of parametric model. Given that the data used are unchanged it is unclear to the EAG why the company has changed its choice of parametric curve. However, the log-logistic does have the lowest AIC by a very small margin.

- The EAG chooses to revert back to the log-normal model as it is associated with an equivalently low AIC and BIC and is also consistent with the choice of parametric model for every other PFS extrapolation considered in this appraisal.

For mosunetuzumab, the average time spent in pre-progression increases in the addendum compared to the original submission, however, it is not possible to determine whether this can be mostly attributed to an improved performance of mosunetuzumab from the additional follow-up or the use of unadjusted data. The company are consistent with their selection of the log-normal extrapolation, which was selected due to its low AIC and the similarity of its hazard function with the observed behaviour.

- The EAG will explore a scenario where the exponential distribution is used for PFS for both arms, as it had the lowest BIC, and fits well to the KM survival function.
    - Both arms show a trend of a decreasing hazard rate over time, but this trend is captured by an approximation, using a kernel or smoother

function, which may be oversensitive or misleading, and the potential optimal fit of the exponential function should not be ignored.

Table 22: Comparing life-year estimates from companies old and new approach to modelling

	PFS life-years	PFS life-years (new data using old parametric modelling)	OS life-years	OS life-years (new data using old parametric modelling)
RB (old) RB (new)	[redacted] (log-normal) [redacted] (log-logistic)	[redacted](log-normal)	[redacted](Weibull) [redacted](exponential)	[redacted](Weibull)
Mosunetuzumab (old) Mosunetuzumab (new)	[redacted] (log-normal) [redacted] (log-normal)	(parametric model is same)	[redacted](exponential) [redacted](log-normal)	[redacted](exponential)

4.2.6.2.2 Overall Survival

Table 22 also shows the total life-year estimates from the company's preferred assumptions in their addendum submission and compares them to the company's previous base case. The mean life-years for RB has decreased slightly as the company has switched from a Weibull model to an exponential model, despite there being no change to the data used. The exponential model is associated with the lowest BIC, however there is little to distinguish between all models regarding their information criteria.

For mosunetuzumab the company have selected a log-normal model, switching from their previous choice of an exponential model. The log-normal model had the second lowest AIC and BIC, behind the exponential model. As with PFS, it is unclear where to attribute the additional benefit: the extended follow-up, the switch to unadjusted data or change of parametric model.

- The EAG note that implementing the exponential model in the addendum economic model produces very similar life-year estimate to the original submission, suggesting the majority of the benefit has come from the change of parametric model (Table 22).

The EAG will present scenarios that select the same parametric form for both treatments, either a log-normal or exponential model, as there is insufficient evidence to support the selection of different parametric forms.

The EAG will also consider pooling the OS data for both arms, given the low quality of reporting by the company and lack of reliable evidence supporting an OS benefit for mosunetuzumab. Whilst the indirect comparison provided suggests a potential difference in [redacted], it also suggests that [redacted]. These contradictory outcomes suggest it is plausible to consider that there may be no OS benefit of mosunetuzumab over RB.

- Considering the issues identified above, the EAG recommends using exponential model fitted to pooled OS data from both treatments. This is because there is extremely limited OS data available meaning that the complexity associated with a survival model with two or more parameters is very unlikely to be well-reflected in the data or well-captured by a model.[45]

4.2.6.2.3 TTOT

For RB, the company assume that the TTOT is equal to PFS but is capped at a maximum of 17 weeks (6 treatment cycles) for both rituximab and bendamustine. This is consistent with the stopping rule for the MA of RB, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic analysis.

- The EAG suggest that this may lead to overestimation of the proportion of patients remaining on RB treatment, potentially overestimating the costs associated with this regimen.

For mosunetuzumab, the company again model TTOT as it was observed in GO29781. This approach is subject to the same limitations which are already outlined in the critique of the R^2 TTOT modelling.

- The EAG will again explore a scenario where the TTOT of mosunetuzumab is increased by increasing the proportion of patients receiving mosunetuzumab beyond 6 months.
- The EAG will also explore removing the half-cycle correction for TTOT as the model cycle length is already precise enough to capture the treatment costs accurately.

4.2.6.2.4 Company's exploratory analysis for mosunetuzumab vs RB

In the addendum, the company presents an exploratory analysis which implements a regression analysis after the propensity score matching, to try and remove the effect of the remaining imbalanced covariates between the mosunetuzumab and RB populations.

- The EAG does not object the principle of the approach, however it has concerns regarding the company's implementation.

The company cite a paper by Austin et al,[46] supporting the methodology behind the exploratory analysis. The aim of this approach is to obtain a better estimate of treatment effect by adjusting for covariate imbalance that remains after implementation of propensity score matching, and this study does indeed demonstrate an improvement in effect estimation by performing a regression adjustment implementing nearest neighbour matching. There are, however, clear differences between the simulation of Austin et al. and the company's analysis.

In their simulation, Austin et al. use a sample size of 1,000 subjects per simulation whose event times are known and have a total of 10 independent variables that influence a subject's event time. These 10 variables are used in the generation of propensity scores used in the matching, and in the post-matching regression analysis. In contrast, we note that the company's analysis varies the covariates used at each stage, and does not include every confounding variable. The sample size of

the company's analysis is smaller, with 46 patients in each arm, many of which are censored. The EAG consider that this results in analysis which is underpowered to suitably detect and adjust for the covariate effects.

Recall that the company's propensity score analysis for RB included 18 covariates, three of which were interaction terms. From examining the results of the optimal pairs matching, the company identify seven covariates as poorly balanced as their absolute standardised mean difference between the groups exceeded their 0.1 threshold. These were: ECOG, Ann Arbor Stage, Refractory to last line, Refractory to prior aCD20, prior ASCT, size of largest node lesion, time since completion of last therapy. Another (POD24) was only just below the company's threshold of 0.1 for the absolute standardised mean difference.

- The EAG suggest that these would all be included in the post-matching regression analysis, alongside all other covariates that may affect the outcome. Unfortunately, due to the sample size, it is unlikely that a model could successfully and reliably estimate effect sizes for each of these simultaneously.

Looking at the company's analysis, it is unclear what the starting set of candidate covariates was, however those included in the final models are shown in Table 23. The covariate effects included in the models vary for each outcome, meaning each is adjusted differently. They range from two to five covariates, with none including more than two of poorly matched variables.

- In summary, the EAG consider that the analysis implemented by the company is not comparable with that of the study by Austin et al.
- It is not clear whether this approach improves or worsens the accuracy of the effect size estimates.

Table 23: Covariates included in company's propensity score matching and post-matching adjustment in the exploratory analysis

	Original RB propensity score analysis	Analysis on mosunetuzumab PFS data	Analysis on RB PFS data	Analysis on mosunetuzumab OS data	Analysis on RB OS data
Poorly Matched Covariate					
ECOG	Yes	-	-	-	Yes
Ann Arbor Stage	Yes	-	-	-	
Refractory to last line	Yes	-	-	-	Yes
Refractory to prior aCD20	Yes	-	Yes	-	-
Prior ASCT	Yes	-	-	-	-
Size of largest node lesion	Yes	Yes	-	Yes	-
Time since completion of last therapy	Yes	Yes	-	-	-
POD24	Yes	-	Yes	-	-
Well matched Covariate					
High LDH	Yes	Yes	-	-	Yes
FLIPI	Yes	-	-	Yes	-
Age	Yes	-	-	-	Yes
Low Hgb	Yes	-	-	-	Yes
No of Prev Therapies	Yes	-	-	-	-
Bone marrow involvement	Yes	-	-	-	-
Double refractory	Yes	-	-	-	-
Interaction terms x 3	Yes	-	-	-	-
Total Number of covariates	18	3	2	2	5

A further EAG concern is the company's decision to censor event times for patients if they were predicted to fall beyond the limit of follow-up. We consider that this introduces bias into the analysis, especially if patients are censored shortly before an

event time occurs. If event times were excessively large, it could be an indication that the model adjustment is inappropriate and not fit for purpose. This analysis also assumes that the survival times follow a Weibull distribution, which was deemed unsuitable by the company to represent a good fit to the data for some of the outcomes.

- The EAG has concerns over the implementation of this analysis and concludes that there is no evidence to show that this exploratory analysis offers any improvement in accounting for the imbalance in baseline characteristics (see Section 1.1). The results are unlikely to be useful to the decision-making process.

4.2.6.3 Mosunetuzumab vs Obinutuzumab Bendamustine (OB)

This section explores the company's methods of incorporating time-to-event data from their trials (GO29781, GADOLIN) into the economic model. As in the RB comparison, for PFS and OS the company use unadjusted data for the control arm, and IPTW adjusted data for mosunetuzumab.

- The rationale for this remains unclear and the EAG considers it inappropriate. Instead, it is our view that the company should implement the IPTW adjusted data for both arms.

Nevertheless, the EAG proceeds to critique the subsequent extrapolations provided by the company.

4.2.6.3.1 Progression Free Survival

In their original submission, the company noted that there was no evidence that proportional hazards assumption was violated but chose to model each arm independently to be consistent with other comparators.

In the addendum, the company has taken a different approach and reports they have implemented a proportional hazards approach. Typically, the EAG consider that this means fitting models to both arms of data, including a parameter estimate of treatment effect which could be considered a sensible approach. However, in this instance, the company have fitted parametric models to the mosunetuzumab data as

before, and to these then apply a hazard ratio as calculated in the clinical effectiveness section.

- The EAG consider that this approach is flawed as it does not make use of the extended follow-up from the GADOLIN trial when fitting the parametric model. It also does not consider the goodness of fit to the OB data as it is not included in the model fitting.

The EAG demonstrates the inappropriateness of this approach in Figure 24 where the extrapolation deviates from the observed data very early on, overestimating PFS for OB. The effects of estimate of progression free life years are shown in Table 24. Focusing on the mosunetuzumab PFS extrapolation, the company maintained a log-normal extrapolation across their original and addendum submissions. This was associated with the lowest BIC and second lowest AIC.

- The EAG is happy with this selection as the model extrapolations are plausible, and there is little to choose between many of the candidate curves.

In selecting a preferred extrapolation for OB PFS, the EAG opts to not use the proportional approach implemented by the company, and instead consider the parametric models fitted to the OB data. The Gompertz model as previously selected by the company is associated with the lowest AIC, whilst the exponential model has the lowest BIC.

- These curves are similar, but the EAG selects the exponential as the PFS hazard rate for OB is noticeably constant (Figure 17, company addendum).

The Gompertz PFS model is explored in a scenario analysis.

Figure 24: Comparing EAG's and company's (red) preferred extrapolation of PFS for OB with observed follow-up from GADOLIN.

Table 24: Comparison of PFS and OS life-year estimates from company's original and addendum submissions

	PFS life-years	**PFS life-years (new data using old parametric modelling)**	**OS life-years**	**OS life-years (new data using old parametric modelling)**
OB (old) OB (new)	(Gompertz) (log-normal PH)	(Gompertz)	(log-normal) (log normal PH)	(log-normal)
Mosunetuzumab (old) Mosunetuzumab (new)	(log-normal) (log-normal)	(log-normal)	(exponential) (log-normal)	(exponential)

4.2.6.3.2 Overall Survival

The company took a similar approach for OS as they did for the PFS extrapolation of this comparison. For mosunetuzumab, the company changed from their original preferred exponential extrapolation to a log-normal model in their addendum

submission. The original justification selected the exponential as it had the lowest AIC and BIC. This was also the case for the data-cut of the addendum submission but the company switched to a log-normal distribution as the hazard plot suggests a decreasing trend. Whilst there does appear a decreasing trend, the EAG deems the hazard plot unreliable as this trend is informed by roughly three death events, with few patients remaining at risk. Earlier in the mosunetuzumab follow-up, where the data are more reliable, there appears to be an increasing hazard rate.

- As the true underlying hazard rate is unlikely to be well represented by the observed data, the EAG select an exponential model as it models a constant hazard rate as was the best fitting model according to the goodness of fit statistics.

For OB, in the addendum the company applied the hazard ratio from the clinical effectiveness section to their preferred extrapolation for mosunetuzumab. Again, this fails to make use of the full follow-up of GADOLIN and results in a deviation from the observed data (Figure 25).

- The EAG prefers to use a log-normal curve fitted to the GADOLIN data, as used by the company in their original submission. This model has the second lowest AIC and BIC, behind the exponential, but it captures well the decreasing hazard rate that is evident across the complete observed period of the GADOLIN follow-up.
- The EAG-preferred OS models result in very similar mean survival times for mosunetuzumab and OB (Table 24), which is in-line with the clinical experts consulted by the company who stated there would be little difference in OS between the two treatments.
- The EAG will explore a scenario where the exponential distribution for OS of OB is explored, and another where OS is pooled for both arms and extrapolated using a log-normal model.

Figure 25: Comparing EAG's and company's (red) preferred extrapolation of OS for OB with observed follow-up from GADOLIN.

4.2.6.3.3 Time to off treatment

For OB, the company assume that the TTOT is equal to PFS but is capped at a maximum of 104 weeks (6 treatment cycles + maintenance therapy every 2 months up to 2 years) for obinutuzumab and 23 weeks (6 cycles) for bendamustine. This is consistent with the stopping rule for the MA of RB, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic analysis.

- The EAG note that, as with the other comparisons this may lead to overestimation of the proportion of patients remaining on OB treatment, potentially overestimating the costs associated with this regimen.

For mosunetuzumab, the company model TTOT as it was observed in GO29781. This approach is subject to the same limitations which are already outlined in the critique of the R^2 TTOT modelling.

- The EAG will again explore a scenario where the TTOT of mosunetuzumab is increased by increasing the proportion of patients receiving mosunetuzumab beyond 6 months.
- The EAG will also explore removing the half-cycle correction for TTOT as the model cycle length is already precise enough to capture the treatment costs accurately (as mentioned in Section 1.1).

4.2.7 Health related quality of life

The utility values were calculated based on EQ-5D-5L data (mapped onto EQ-5D-3L valuation set) collected in GO29781. Linear mixed regression models were used to estimate utilities for the different heath states (PFS and PPS). Each regression included the following as covariates:

- Centralised baseline utilities
- Random effects at the patient level
- Relevant covariate of interest (e.g., PFS/PPS, on treatment/off treatment).

In their response to clarification (Question B2) the company stated that other important variables such as the previous number of systemic therapies were explored but were discarded in the end due to non-significance.

The company also argued that covariate adjustment would have added complexity to the model with little added precision as the pairwise characteristics of the comparisons, would demand that covariate adjustment be done by each pairwise comparison.

The company applied age-related disutility to the estimates using the brazier age-adjusted health state utility value coefficient.

4.2.7.1 Health-related quality of life data identified in the review

According to the CS, the SLR identified a total of 6 studies which reported health state utility values (HSUVs) associated with patients with FL/indolent non-Hodgkin's lymphoma (iNHL) in the 2L+ treatment setting Out of these, the company considered the utility values of the studies of Haukaas et al., Cognet et al., and Wang et al., as most relevant.[42, 47, 48]

4.2.7.2 Health state utility values

The utility values resulting from the mixed effects model were used to inform the health states in the model for mosunetuzumab and all comparators and utility values from the studies by Cognet et al. and Wild et al. were tested in scenario analyses (Table 25). The CS does not provide justification for selecting utilities from Wild et al in its scenario analyses and not utilities from the other two studies identified through the SLR.

- The EAG believes the values from Wang et al. were not adaptable to the mosunetuzumab model structure and using utility values from Haukass et al 2018 on a rituximab-refractory patient population, would result in estimates that were similar to those used by the company (both on and off treatment values).

Table 25. Base case utility values and scenario utility values

Scenario	State	Utility values	Standard Error
Base case	PFS	0.804	0.01
	PPS	0.75	0.02
Scenario (Wild et al 2006)	PFS	0.81	0.02
	PPS	0.62	0.06
Scenario (Cognet et al 2015)	PFS	0.71	0.30
	PPS	0.51	0.35
Source: Table 37 (Company Submission; v2.0)			

EAG Comments:

- The HRQoL estimates used in the model present a key area of uncertainty. The company presented patient reported outcomes up to Cycle 8 of treatment (approximately 24 weeks post baseline) and the data presented appear to be based on patients who had a baseline and at least one post-baseline assessment of PRO scales), not based on intention-to-treat. At clarification (Question B8), the company stated that a total of 7 patients did not have a baseline value and 3 patients only had baseline but no other data.
    - The EAG was concerned about the potential bias that missing data introduced into the analysis.

- The company justified the lack of accounting for missing data on the basis that they undertook a complete case approach to analysing the utility data and the number of patients with missing data was relatively small to have a significant impact on regression outcomes.
    - The EAG is unconvinced by this rationale. Firstly, the definition of a complete case used by the company is rather conservative as they considered a patient with data at only 2 timepoints as compliant (leaving room for a lot of missingness). Secondly, though there is no threshold for what is considered acceptable missingness, there is some consensus that 5% missingness is unlikely to have an impact on analysis.[49]
    - In the CS (using their conservative definition of compliance), 11% of patients had missing data. The EAG suggest that the cost-effectiveness results should, therefore, be interpreted with these factors in mind: immaturity of data used to inform post-baseline utilities, small sample sizes, and the potential bias introduced by missing data.

- The EAG is concerned with the inconsistencies in the way in which source of data used for PPS utilities are reported.

    - In the CS (table 35) and elsewhere in the report, GO29781 trial is reported as the source for PPS utility estimates used in the economic model. However, the information presented on table 24 (CS) in which the company states "*Trial data was used where possible, but data were immature for PPS and so estimates were sourced from published evidence*" seems to suggest that the company used external data sources to inform utility estimates for the PPS state.

- The EAG agree that external data sources using more mature data would be more appropriate to inform PPS utilities as opposed to GO29781 data.

- The company argues that covariate adjustment would have added complexity with little added precision. However, without presenting utility data in which covariate adjustment was done for any of the pairwise comparisons it is

challenging for the EAG to verify the company's claim that utilities are unlikely to differ across arms.

4.2.8 Resources and costs

4.2.8.1 Intervention and comparator costs

The costs of mosunetuzumab for each cycle were made up of drugs acquisition (Table 26) and administration costs (

Table 27). Dosing schedule followed that in GO29781 trial: 1mg on day 1, 2mg on day 8 and 60mg on day 15 of the first cycle. In cycle 2, 60mg was taken on the first day while 30mg was taken on the first day of each subsequent cycle.

A patient access scheme (PAS), comprising a discount of [redacted]% was applied to the mosunetuzumab drug acquisition costs. The base case analysis of the economic evaluation was based on mosunetuzumab drug costs (with PAS applied).

Table 26. Mosunetuzumab dosing and acquisition

Dosing	1/2/60/60/30mg
Dose per cycle	As above
Cost (excluding PAS)	£6,600 (30mg); £220 (1mg)
Cost per dose (excluding PAS)	1mg: £220 2mg: £440 60mg: £13,200 30mg: £6,600
Total (excluding PAS)	1mg: £220 2mg: £440 60mg: £13,200 30mg: £6,600
Abbreviations: PAS: patient access scheme	
Source: Table 38 (Company Submission)	

Table 27. Administration costs for mosunetuzumab

Component	National cost collection for the NHS	Cost	Inflated costs
Administration	Daycase and Reg Day/Night: Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance (SB14Z)	£403.84	NHS Reference Costs 2019 to 2020
Source: Table 39 (Company Submission)			

The CS stated that dosing schedule for comparators followed BNF recommendations. For R[2], rituximab was assumed to be given at 375mg/m^2 on days 1, 8, 15 and 22 in the first cycle and on the first day of cycles 2–5. Lenalidomide was assumed to be given at a dose of 20mg daily for days 1–21 of each 28-day cycle.

For rituximab with bendamustine (RB), rituximab was assumed to be given at 375mg/m^2 every cycle (21days). Bendamustine was given at 90 mg/m^2 on days 1 and 2. For OB, obinutuzumab was assumed to be given at a fixed dose of 1,000mg on days 1,8 and 15 in the first cycle. A fixed dose of 1,000mg was given on day 1 of cycles 2-6 (28-day cycles). A fixed dose of 1,000mg was given on the first day in cycles 7–18. Cycle length for cycles 7-18 was assumed to be 56 days. Bendamustine was given at 90 mg/m^2 on days 2 and 3 in cycle 1 and at 90 mg/m^2 days 1 and 2 of cycles 2–6. Table 28 below shows the estimated costs per cycle for each of the comparator treatments.

Administration costs for comparators were assumed to be the same as for mosunetuzumab for the first cycle. For subsequent cycles, administration costs were costed as subsequent elements of a chemotherapy cycle (

Table 29).

Table 28. Comparator cost per cycle

Comparator	Cost per cycle
Rituximab	£1,349.89
Bendamustine	£257.05
Obinutuzumab	£3,312.00

Lenalidomide	£1,389.50
Source: Table 42 (Company submission)	

Table 29. Comparator administration costs

Component	National cost collection for the NHS	Cost	Inflated costs
Administration (first appointment)	Daycase and Reg Day/Night: Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance (SB14Z)	£403.84	NHS Reference Costs 2019 to 2020
Administration (Subsequent appointments)	Daycase and Reg Day/Night: Subsequent Elements of Chemotherapy Cycle (SB15Z)	£339.46	NHS Reference Costs 2019 to 2020
Administration (oral treatment)	Daycase and Reg Day/Night: Deliver Exclusively Oral Chemotherapy (SB11Z)	£210.79	NHS Reference Costs 2019 to 2020
Abbreviations: NHS: National Health Service Source: Table 43 (Company submission)			

4.2.8.2 Costs of subsequent treatments

Subsequent treatments were included in the model as an average one-off cost to patients entering the PPS heath state, taking into account the mean duration of treatment, the proportion assumed to use each treatment option (i.e. treatments available at third and subsequent lines of FL treatment) and the costs.

The CS stated that data on subsequent therapy use were derived from a survey of clinicians who attended the Clinical Advisory.

Table **30** shows the mean duration and percentage share of each therapy class included in the post-discontinuation treatment for mosunetuzumab, RB, rituximab lenalidomide (RL) and OB and the total post-discontinuation costs which were similar for mosunetuzumab and all comparators (Table 31).

Table 30. Proportion assumed to take each subsequent therapy by arm

Therapy class	Therapy	Mean duration in weeks	% on Mosun	% on RB	% on OB	% on RL
R[2]	R[2]	52.00	35%	35%	35%	35%
R-Chemo	R-Chemo	21.73	25%	25%	25%	25%
Other (non-rituximab chemo)	Other (non-rituximab chemo)	21.73	10%	10%	10%	10%
Palliative care	Palliative care	21.73	10%	10%	10%	10%
Trials	Trials	21,73	20%	20%	20%	20%

Table 31. Total post-discontinuation costs

Treatment	Total cost
Mosunetuzumab	£40,570.93
Rituximab + Bendamustine	£40,570.93
Obinutuzumab + Bendamustine	£40,570.93
Rituximab + Lenalidomide	£40,570.93
Source: Table 46 Company submission	

4.2.8.3 Other health state costs

Supportive and terminal care costs

Resource use was determined by disease state (PFS vs. PPS) rather than treatment arm. The CS stated that resource use values for the PFS health state used in the model were elicited from clinicians following consultation as the resource use values extracted from TA243 were perceived to be an overestimate. Similarly, resource use values for the progressed health state were based on expert opinion, rather than TA604 as clinicians perceived the values from TA604 as no longer representative of current practice/usage. A similar approach was taken for terminal care costs.

Table 32 and Table 33 show the supportive care costs associated with the PFS and PPS health states. A one-off terminal care cost of £6,707.78 was applied in the model.

Table 32. Supportive care cost

Component	Mean cost used in the model (£)
Progression-free state	59.32
Progression state	57.11
One off progression cost	220.78
Source: Table 50 (Company submission)	

Table 33. Supportive care one-off costs associated with progression state

Component	% Patients	Unit cost	Total Cost
Radiological assessments	100	£77.31	£77.31
Biopsy	25	£563.62	£140.91
Blood test	100	£2.56	£2.56
Source: Table 51 (Company Submission)			

4.2.8.4 Adverse event costs and resource use

Table 34 presents unit costs for adverse events applied in the model. The costs of adverse events (AEs) were calculated based on the average number of treatment-related AEs per patient per week in the relevant trial. The probability of event was combined with the costs of each AE in each treatment arm with the resulting cost (Table 35) applied to the proportion remaining on treatment in each cycle in the model.

NHS reference costs 2019/2020 were used to cost adverse events.

Table 34. Costs of AEs included in the model

Event	Mean cost used in the model (£)	Source*
Alanine aminotransferase increased	£125.44	NHS Reference Costs 2019/20: WF01A - Non-Admitted Face-to-Face Attendance, Follow-up
Anaemia	£3,674.74	NHS Reference Costs 2019/20: weighted average of SA03G to SA03H, NEL
Aspartate aminotransferase increased	£125.44	NHS Reference Costs 2019/20: WF01A - Non-Admitted Face-to-Face Attendance, Follow-up
Cytokine release syndrome	£1,619.80	Hospitalisation costs: NHS Reference Costs 2019/20: weighted average of medical adult patients (unspecified specialty) XC01Z to XC07Z, Critical Care, multiplied by mean days hospitalised for G3-4 CRS in GO29781 (median duration 6 days as per trial)

	£102.40	Tocilizumab: List price (BNF) multiplied by average number of tocilizumab doses administered in patients with G3-4 CRS in GO29781 (8mg/kg per dose, weighted average of 1.5)
	£11,050.03	Total cost for CRS event (assuming [redacted] days hospitalised and weighted average doses of tocilizumab received ([redacted]) based on mean weight of cohort (81kg), requiring [redacted] vials)
Hypokalaemia	£356.81	NHS Reference Costs 2019/20: Weighted average of KC05H to KC05N, day case
Hypophosphatemia	£348.17	NHS Reference Costs 2019/20: KC05N - Fluid or Electrolyte Disorders, without Interventions, with CC Score 0-1, NES
Lymphopenia	£457.41	NHS Reference Costs 2019/20: Weighted average of SA08G to SA08J, day case
Neutropenia	£2,042.93	NHS Reference Costs 2019/20: SA08J - Other Haematological or Splenic Disorders, with CC Score 0-2, NEL
Neutrophil count decreased	£2,042.93	NHS Reference Costs 2019/20: SA08J - Other Haematological or Splenic Disorders, with CC Score 0-2, NEL
Rash erythematous	£441.46	NHS Reference Costs 2019/20: Weighted average of JD07E-K, NES
Thrombocytopenia	£3,573.86	NHS Reference Costs 2019/20: Weighted average of SA12G to SA12K, NEL
Tumour flare	£0.00	No costs assumed
Upper respiratory tract infection	£1,794.53	NHS Reference Costs 2019/20: Weighted average of WH07A to WH07G across NEL, NES, DC
Diarrhoea	£1,685.34	NHS Reference Costs 2019/20: FD01J - Gastrointestinal Infections without Interventions, with CC Score 0-1, NEL
Febrile neutropenia	£6,933.22	NICE guidelines NG52; Appendix A (A.3.2.3); inflated to 2019/20 based on PSSRU
Infections	£1,794.53	NHS Reference Costs 2019/20: Weighted average of WH07A to WH07G across NEL, NES, DC
Vomiting	£1,922.00	NHS Reference costs 19/20: FD10M -Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2, NEL
Infections	£1,922.44	NHS Reference costs 19/20: FD10M -Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2, NEL
Leukopenia	£4,205.55	NHS Reference Costs 2019/20: SA31E - Malignant Lymphoma, including Hodgkin's and Non-Hodgkin's, with CC Score 2-3, NEL
Cutaneous reaction	£612.78	NHS Reference Costs: 2019/20: SA31E - Malignant Lymphoma, including Hodgkin's and Non-Hodgkin's, with CC Score 2-3, NES
Source: Table 53 (Company submission)		

Table 35. Adverse event costs per cycle

Drug regimen	Cost per model cycle (weekly) (£)
Mosunetuzumab	17.34
Rituximab and Bendamustine	18.53
Obinutuzumab and Bendamustine	15.71
Rituximab and Lenalidomide	13.66
Source: Table 54 (Company submission)	

EAG Comments:

The main concerns of the EAG relate to:

- subsequent treatments that were included as a one-off cost and were therefore potentially underestimated
- exclusion of retreatment costs in submission which may have slightly biased the cost-effectiveness in favour of mosunetuzumab. Retreatment with mosunetuzumab was allowed in the study protocol. However, data for this were not presented in initial company submission (but were provided at clarification [Question A13]) as the company stated that retreatment following progression is not covered in the marketing authorisation.
    - The EAG agrees but is concerned that retreatments may have slightly biased the cost-effectiveness in favour of mosunetuzumab if it improved survival as costs for these treatments would not have been captured.

5 COST EFFECTIVENESS RESULTS

5.1 *Company's cost effectiveness results*

The company's base case assumed the following distribution for the PFS and OS for the mosunetuzumab population and comparators in Table 36.

Table 36 Company base case: Parametric extrapolation used

	Model base case
Mosun PFS distribution	Log-normal
Mosun OS distribution	Log-normal
RB PFS distribution	Log-logistic
RB OS distribution	Exponential
R^2 PFS distribution	Log-normal
R^2 OS distribution	Weibull
OB PFS distribution	Log-normal
OB OS distribution	Log-normal
PH assumption – OB	Yes

The discounted and undiscounted Life Years (Lys) and quality adjust life years (QALYs) between mosunetuzumab and its comparators are presented in Table 37 below.

Table 37 Company base case: Discounted and undiscounted LYs and QALYs

	Undiscounted LY		Undiscounted QALYs		Discounted LY		Discounted QALYs	
	PFS	**PPS**	**PFS**	**PPS**	**PFS**	**PPS**	**PFS**	**PPS**
Mosun, unweighted	4.03	12.22	[redacted]	[redacted]	3.19	7.93	[redacted]	[redacted]
Mosun, R^2 weighted	10.55	7.41	[redacted]	[redacted]	7.17	5.01	[redacted]	[redacted]
R^2	5.83	8.43	[redacted]	[redacted]	4.61	5.75	[redacted]	[redacted]
Mosun, RB weighted	6.10	11.03	[redacted]	[redacted]	4.56	7.07	[redacted]	[redacted]
RB	3.19	8.00	[redacted]	[redacted]	2.71	5.59	[redacted]	[redacted]
Mosun, OB weighted	2.57	11.67	[redacted]	[redacted]	2.19	7.72	[redacted]	[redacted]
OB	7.35	3.30	[redacted]	[redacted]	5.44	2.40	[redacted]	[redacted]

Disaggregated discounted costs are presented in Table 38 below.

Table 38 Company base case: Discounted disaggregate costs

	Mosun, unweighted	Mosun-R^2	Mosun-RB	Mosun-OB	RB	OB	R^2
PFS							
Treatment cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diagnostic test	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug admin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Supportive care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PPS							
Supportive care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The results for the company's base case cost-effectiveness analysis is presented below. The results in Table 39 below are the mosunetuzumab-comparator weighted population compared to comparators.

Table 39 Company base case: ICER per comparator

	Mosun vs RB	**Mosun vs OB**	**Mosun vs R^2**
Incremental Lys	3.33	2.08	1.82
Incremental QALYs	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]	[redacted]	[redacted]
ICER	£23,504	£7,727	£8,822

The base case probability of cost-effectiveness at a threshold of £20,000 is presented in Table 40 below.

Table 40 Company base case: Probability of cost-effectiveness at £20,000 threshold

	Mosun vs RB	Mosun vs OB	Mosun vs R^2
Mosun Dominant	[redacted]	[redacted]	[redacted]
Mosun-cost effective	[redacted]	[redacted]	[redacted]
Mosun not cost-effective	[redacted]	[redacted]	[redacted]
Mosun dominated	[redacted]	[redacted]	[redacted]

5.2 *Company's sensitivity analyses*

The company presented a range of univariate sensitivity analysis and probabilistic sensitivity analysis using 1000 iterations randomly drawn from the specific distributions.

Univariate sensitivity analysis was performed on 10 parameters considered to be influential by the company. The analysis was performed for each comparator and results presented using a tornado diagram showing its impact on NMB at a threshold of £20,000 as shown in Figure 26, Figure 27, and Figure 28 below.

Figure 26 Tornado diagram showing one-way sensitivity analysis (OWSA) results on NMB – mosunetuzumab vs R^2

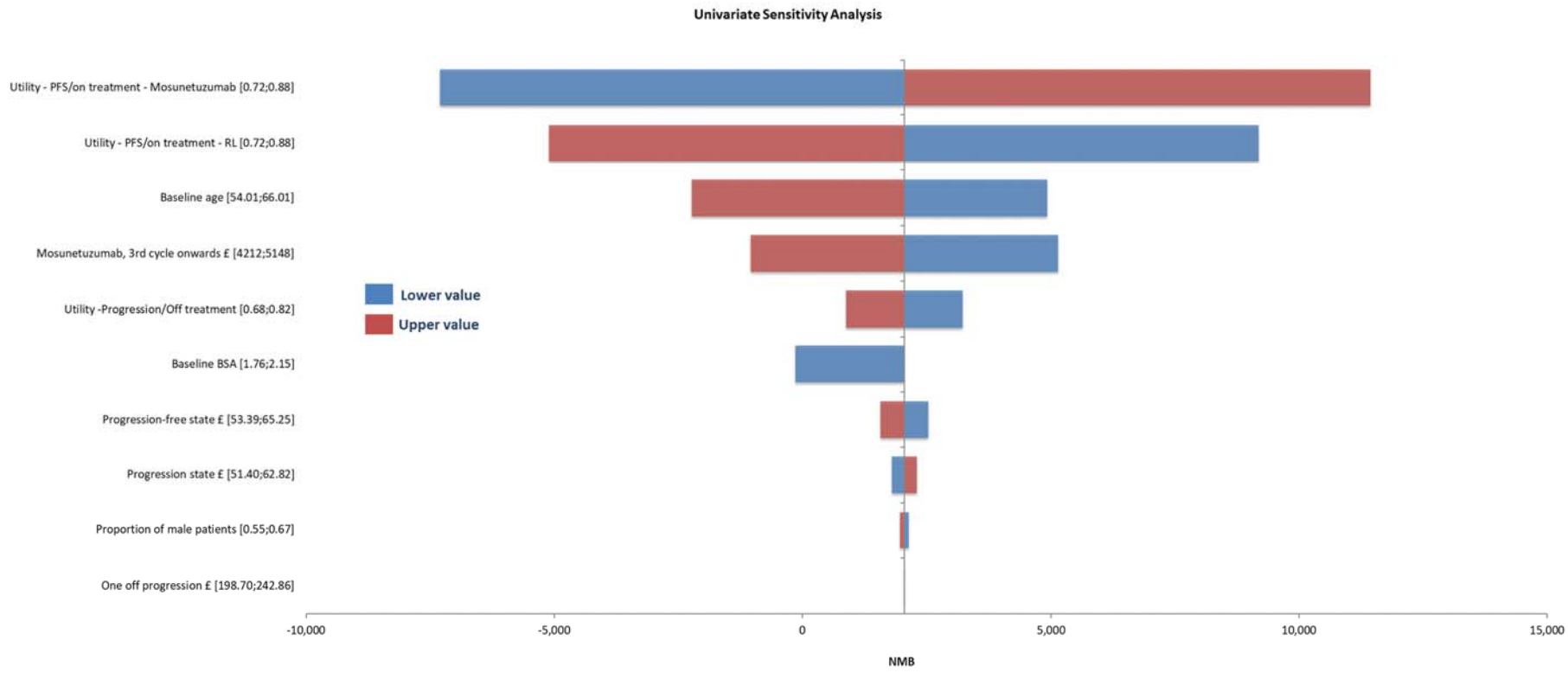


Figure 27 Tornado diagram showing OWSA results on NMB – mosunetuzumab vs RB

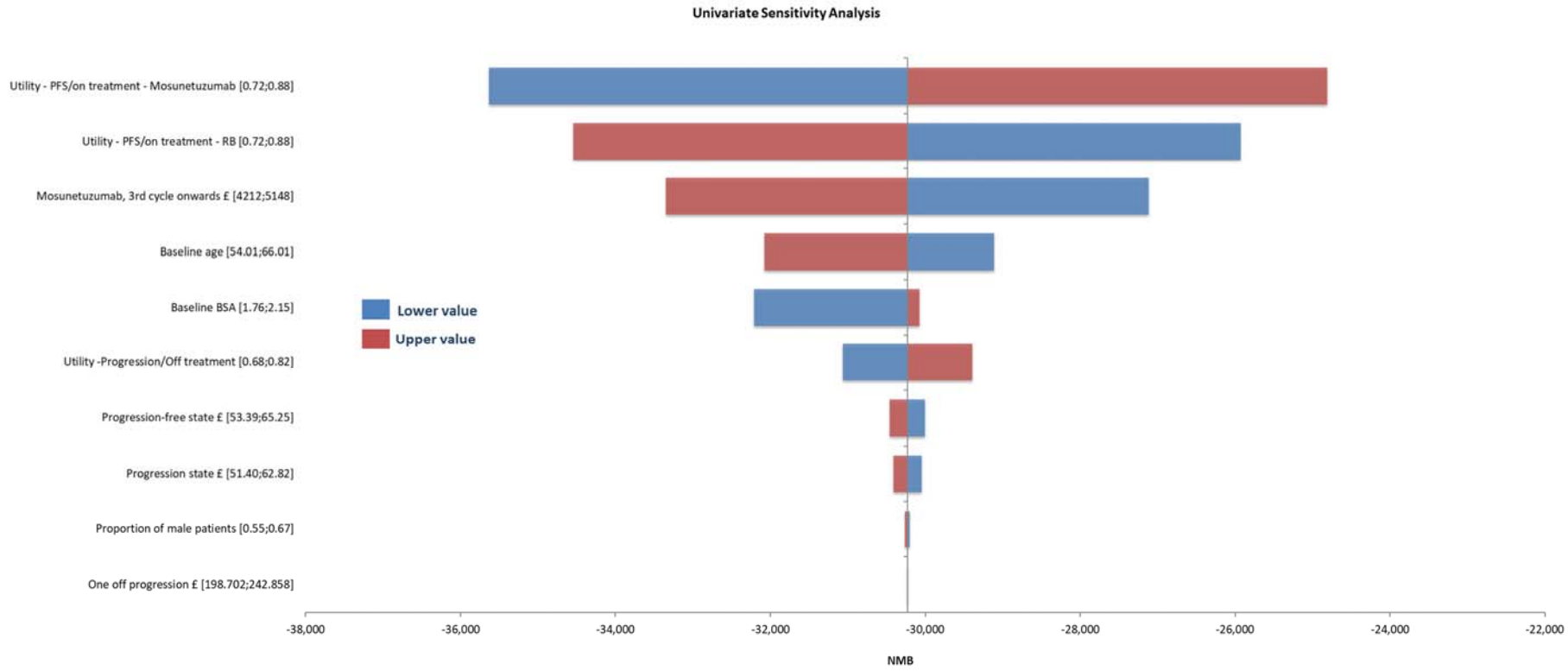
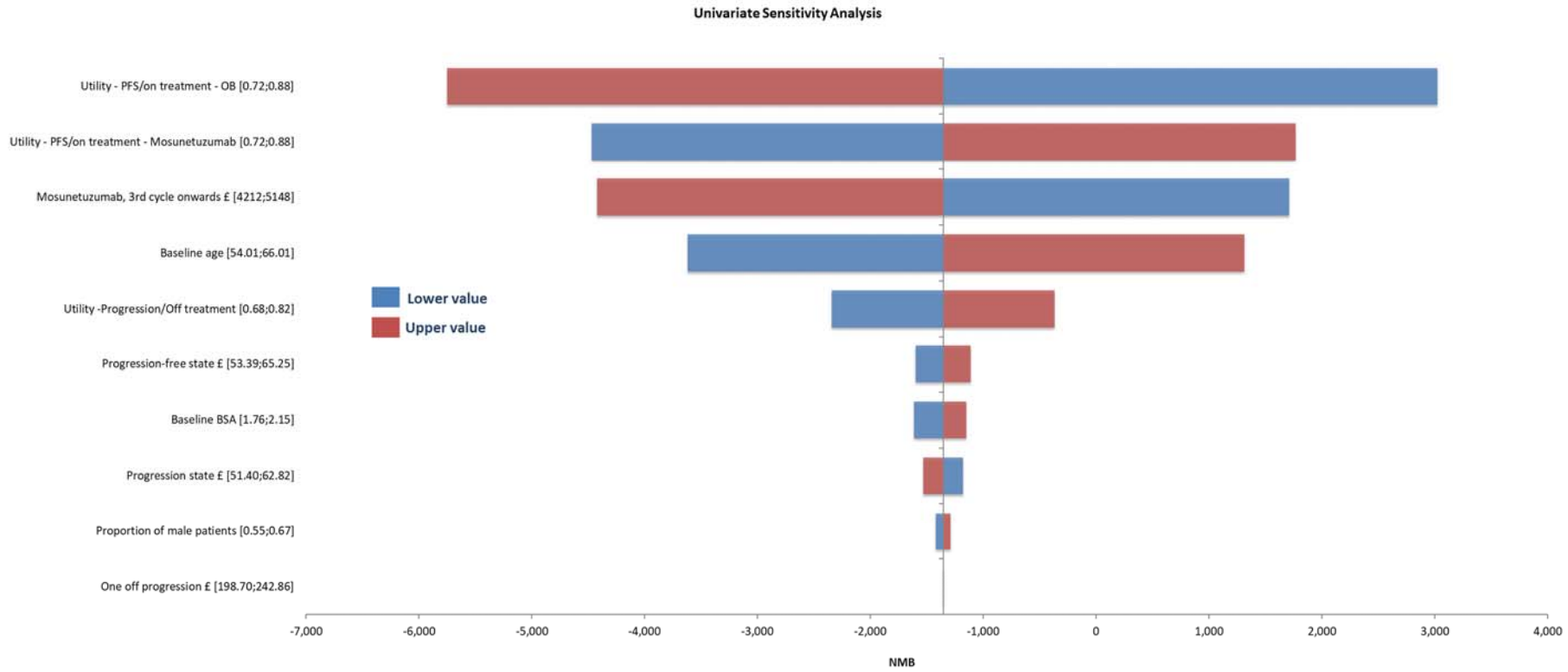


Figure 28 Tornado diagram showing OWSA results on NMB – mosunetuzumab vs OB



The parameter with the largest impact on cost-effectiveness was the utility assigned to the PFS health state for mosunetuzumab vs comparators.

[redacted] The influential parameters were directly related to the time spent in each health state.

The probabilistic sensitivity analysis was similar to the deterministic values as shown in Table 41 below

Table 41 Company base case: Probabilistic sensitivity analysis

	Mosun vs RB	**Mosun vs OB**	**Mosun vs R^2**
Incremental LYs	2.55	1.85	0.79
Incremental QALYs	[redacted]	[redacted]	[redacted]
Incremental costs	[redacted]	[redacted]	[redacted]
ICER	£29,155	£7,664	£14,102

The uncertainty around the cost-effectiveness estimates above were presented using an incremental cost-effectiveness (ICER) plane as shown Figure 29.

The probability that mosunetuzumab is a cost-effective alternative to comparators is presented at various willingness to pay thresholds as presented in Figure 30.

Figure 29. Company base case: ICER plane

Figure 30. Company base case: CEAC

The company presented a range of scenario analysis including the sensitivity of cost-effectiveness estimates to alternative parametric distributions, alternative health state utility sources, assumptions about proportional hazard and a regression adjustment method matching mosunetuzumab with RB population.

- The EAG note that cost-effectiveness estimates are sensitive to alternative parametric distribution for R^2 OS and alternative health state utility sources.

5.3 *Model validation and face validity check*

The EAG conducted an extensive review of the model submitted by the company.

- The model appears to reflect the assumptions made by the company.

6 EXTERNAL ASSESSMENT GROUP'S ADDITIONAL ANALYSES

Table 42 summarises the main issues highlighted by the EAG throughout this report that could impact mosunetuzumab's cost-effectiveness.

It shows the expected direction of bias introduced by these issues and whether these are examined in any exploratory analyses or incorporated in the EAG base-case.

Table 42 Main EAG critique of company's submitted economic evaluation

Issue	Likely direction of bias introduced in ICER	EAG analyses	Addressed in company analyses
Model structure (Section 4.2.2)			
Inappropriate application and use of half-cycle correction	+	Yes, Base case	No
Population, intervention and comparators, perspective, and time horizon (Section 4.2.3 to 4.2.5)			
The base case analysis includes RB which the EAG has some reservations	NA	Base-case	No
Treatment effectiveness and extrapolation (Section 4.2.6)			
Extrapolation of R^2 PFS beyond [redacted] associated with significant uncertainty	+	Base case	No
Magnitude of benefit following parametric extrapolation of R^2 OS not supported by evidence	+	Base-case Scenarios	Scenarios
Potential bias in estimating the proportion of patients remaining on R^2 treatment	+	Scenarios	No
Inconsistent data used in RB PFS extrapolation	+/-	No	No
Choice of parametric extrapolation for RB PFS	+	Base-case Scenarios	Scenarios
Extrapolation of RB OS	+	Base case Scenarios	Scenarios
Potential bias in estimating the proportion of patients remaining on RB treatment	+	Scenarios	No
Extrapolation of OB PFS	+ and -	Base-case Scenarios	Scenarios
Extrapolation of OB OS	+ and -	Base-case Scenarios	Scenarios
Potential bias in estimating the proportion of patients remaining on OB treatment	+	Scenarios	No
Adverse events (Section 4.2.6)			
Incidence of grade 3 and 4 adverse events differ across arms and are likely to affect HRQoL estimates	+/-	No	No
Health-related quality of life (Section 4.2.7)			
Lack of clarity on the source of utility used to inform PPS health state	+/-	No	No

Resource use and cost (Section 4.2.8)			
Subsequent treatments included as a one-off cost and could be potentially underestimated or overestimated.	+/-	No	No
Mosunetuzumab retreatment costs not accounted for	+	No	No
Cost-effectiveness analyses (Section 5.1 to 6.3)			
Different base case assumptions and scenarios were explored	+ and -	Base case Scenarios	Scenarios
Footnotes: Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; '+/-' indicates that the bias introduced by the issue is unclear to the EAG; while '+' indicates that the EAG believes this issue likely induces bias in favour of the intervention versus at least one comparator and '+and -' indicates the EAG believes the potential bias can be positive or negative depending on the assumptions used.			

Based on all considerations in Section 4.2 of this report (summarised in Table 42), the EAG defined a new base case.

6.1 *EAG revised base case*

The adjustments made to the company model are described below and summarised by each pairwise comparison.

6.1.1 Mosunetuzumab vs R^2 comparison

EAG01: Given the high number of censored observations and subsequently very small number of actual observations that inform the mosunetuzumab weighted extrapolated model beyond [redacted], the EAG sets PFS for mosunetuzumab equal to R^2 from [redacted] whilst maintaining the log-normal extrapolation used by the company.

EAG02: For the OS, the EAG prefers a pooled OS with a log-normal extrapolation given the uncertainty around the long-term extrapolation of OS.

EAG03: Remove half cycle correction for TTOT from the mosunetuzumab arm.

6.1.2 Mosunetuzumab vs RB comparison

EAG04: The EAG prefers to revert back to the log-normal distribution to extrapolate PFS in the RB arm as was the case in the original company submission (section 4.2.6.2.1) whilst maintaining log-normal extrapolation for mosunetuzumab.

EAG05: The EAG prefers an exponential model, fitted to pooled OS data from both treatments (section **Error! Reference source not found.**).

EAG06: Remove half cycle correction for TTOT from the mosunetuzumab arm.

6.1.3 Mosunetuzumab vs OB comparison

EAG07: For mosunetuzumab PFS, the EAG maintains a log-normal extrapolation used by the company. Non-proportional hazard is assumed, and an exponential extrapolation is used for OB PFS.

EAG08: For OS, non-proportional hazard is assumed, and an exponential extrapolation is used for mosunetuzumab

EAG09: Half-cycle correction is removed for TTOT from the mosunetuzumab arm.

The impacts of the individual EAG revisions on ICER are shown in Table 43.

Table 43: Impact of individual EAG's preferred model assumptions on ICER

Preferred assumption	**EAG report sections**	**ICER**
mosunetuzumab vs. R^2		
Company base-case	5.1	**£8,822**
EAG01: PFS for R^2 set to equal mosunetuzumab beyond [redacted]	6.1.1	£9.604
EAG02: Log-normal extrapolation for R^2 OS using OS data from both arms	6.1.1	£51,091
EAG03: Removing half-cycle correction for TTOT	6.1.1	£9,781
mosunetuzumab vs. RB		
Company base case	5.1	**£23,504**

EAG04: log-normal distribution to extrapolate PFS in the RB arm	6.1.2	£23,545
EAG05: Exponential model for OS fitted to pooled OS from both treatments	6.1.2	£238,451
EAG06: Remove half-cycle correction for Mosunetuzumab TTOT	6.1.2	£24,025
mosunetuzumab vs. OB		
Company base case	5.1	**£7,727**
EAG07: Exponential model to extrapolate OB PFS	6.1.3	£5,221
EAG08: Exponential distribution to extrapolate mosunetuzumab OS	6.1.3	Mosunetuzumab Dominated
EAG09: Remove half-cycle correction for TTOT in mosunetuzumab arm	6.1.3	£8,699
OS: Overall survival; PFS: Progression-free survival; Mosun: Mosunetuzumab; OB: Obinutuzumab bendamustine; RB: Rituximab bendamustine; TTOT: Time to off-treatment		

6.1.4 EAG deterministic base case results

The cumulative effect of all EAG changes on deterministic cost-effectiveness results for each pairwise comparison is shown in Table 44.

- The EAG emphasises that all EAG analyses are conditional upon the ITC, for which uncertainty could not be incorporated into the economic model.

For R^2, deterministic incremental costs were [redacted] and incremental QALYs were [redacted] which resulted in mosunetuzumab being dominated by R^2. The main driver for the increased ICER was choice of OS curves. Setting PFS for R^2 equal to mosunetuzumab beyond [redacted] and removing half-cycle correction for TTOT worsens the ICER but to a much lesser extent (Table 43).

For mosunetuzumab- RB comparison, deterministic incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £248,335. Using OS data pooled from both treatments was the main driver for the increased ICER whilst using a log-normal distribution to extrapolate PFS in the RB arm and removing half-cycle correction for TTOT in the RB arm only marginally increased the ICER (Table 43).

For the OB comparison, deterministic incremental costs were [redacted] and incremental QALYs were [redacted], resulting in mosunetuzumab being dominated by OB. Similar to R^2, choice of parametric curve to extrapolate mosunetuzumab OS had the greatest impact on the ICER. In contrast, using an exponential model to extrapolate OB PFS improved the ICER to £5,221 whilst half-cycle correction worsens the ICER to £8,699 (Table 43)

Table 44. Deterministic EAG Base Case Cost-effectiveness Results with PAS discount

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	12.52	[redacted]					
R^2	[redacted]	12.52	[redacted]	[redacted]	0	[redacted]	Dominated	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	£ [redacted]	9.23	[redacted]					
RB	[redacted]	9.23	[redacted]	[redacted]	0	[redacted]	£248,335	[redacted]
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	8.49	[redacted]	[redacted]	-0.13	[redacted]	Dominated	[redacted]

6.1.5 EAG's Probabilistic base case cost-effectiveness results

The EAG base case was subject a probabilistic sensitivity analysis using 1000 iterations drawn from the EAG parametric assumptions. For R^2 comparison, probabilistic incremental costs were [redacted] and incremental QALYs were [redacted] resulting in mosunetuzumab being dominated. The probability of mosunetuzumab being cost-effective at £30,000 threshold is [redacted]. The EAG CEAC is presented in Figure 31

Figure 31 Cost-effectiveness acceptability curve- R^2 comparison

For the RB comparison, probabilistic incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £265,231 per QALY gained. Probability of cost-effectiveness at the £20,000/QALY threshold was [redacted] whilst the probability that mosunetuzumab is dominated was [redacted]. The EAG CEAC is presented in Figure 32 below.

Figure 32 Cost-effectiveness acceptability curve - RB comparison

For the OB comparison, probabilistic incremental costs were [redacted] and incremental QALYs were [redacted], resulting in mosunetuzumab being dominated by OB. The probability of mosunetuzumab being cost-effective at £30,000 threshold is [redacted] The EAG CEAC is presented in Figure 33

Figure 33 Cost-effectiveness acceptability curve - OB comparison

6.2 *EAG Scenario Analyses*

6.2.1 Mosunetuzumab vs. R^2 comparison

Given the sensitivity of health state distributions to assumptions about parametric extrapolation, the EAG explores the following scenarios. In all scenarios, non-proportional hazard is assumed, half cycle correction for TTOT is removed and PFS remains the same as the EAG base case.

Scenario 1: Independent OS and a log-normal extrapolation is used for R^2 OS.

Scenario 2: Pooled OS with log-logistic extrapolation.

Scenario 3: Independent OS and a log-normal extrapolation is used for R^2. The number of people who remain on treatment beyond 6 months is increased by 10%, 30% and 50%.

The impact of each scenario on the deterministic ICER is presented in Table 45. Probabilistic analysis did not change the conclusion of the deterministic analysis. mosunetuzumab is dominated by R^2 in all scenarios.

Table 45 R^2: EAG scenario analysis: impact on ICER

	Incremental costs	Incremental QALYs	ICER £/QALY
Scenario 1	[redacted]	[redacted]	-£32,991
Scenario 2	[redacted]	[redacted]	-£1,045,375
Scenario 3	[redacted]	[redacted]	
10% increase			-£36, 139
30% increase			-£42,435
50% increase			-£48,731

6.2.2 Mosunetuzumab vs. RB comparison

Given the sensitivity of health state distributions to assumptions about parametric extrapolation, the EAG explores the following scenarios. In all scenarios, non-proportional hazard is assumed and half cycle correction for TTOT is removed.

Scenario 1: Use exponential distribution for PFS for both arms as it had the lowest BIC and fits well to the KM survival function.

Scenario 2: OS and PFS set to company's base case and the number of people who remain on treatment beyond 6 months is increased by 10%, 30% and 50%. TTOT is set to the EAG base case.

Scenario 3: Use Log-normal distribution for OS for both arms based on pooled OS data

Scenario 4: Assume exponential independent OS

Scenario 5: Assume log-normal independent OS

The impact of each scenario on the deterministic ICER is presented in Table 46 .

Table 46 Impact of individual EAG scenario analysis on ICER for RB comparison

	Incremental costs	Incremental QALYs	ICER £/QALY
Scenario 1	[redacted]	[redacted]	£24,621
Scenario 2	[redacted]	[redacted]	
10% increase			£24,305
30% increase			£24,863
50% increase			£25,421
Scenario 3	[redacted]	[redacted]	£242,780
Scenario 4	[redacted]	[redacted]	£43,519
Scenario 5	[redacted]	[redacted]	£49,439

6.2.3 Mosunetuzumab vs. OB comparison

Given the sensitivity of health state distributions to assumptions about parametric extrapolation, the EAG explores the following scenarios. In all scenarios, non-proportional hazard is assumed.

Scenario 1: Gompertz extrapolation for OB PFS. OS remains the same as company base case. Half cycle correction for TTOT is removed.

Scenario 2: OS and PFS set to company's base case and the number of people who remain on treatment beyond 6 months is increased by 10%, 30% and 50%. TTOT is set to the EAG base case.

Scenario 3: pooled OS with log-normal extrapolation. PFS remains the same as the company base case. TTOT is set to the EAG base case

Scenario 4: Exponential extrapolation for OB OS. PFS remains the same as the company base case. TTOT is set to EAG base case.

The impact of each scenario on the deterministic ICER is presented in Table 47. Mosunetuzumab was cost-effective against OB in all scenarios except scenario 3 where OS is pooled. Probabilistic analysis did not change the conclusion of the deterministic analysis.

Table 47 EAG scenario analyses for OB comparison: impact on ICER

	Incremental costs	**Incremental QALYs**	**ICER £/QALY**
Scenario 1	[redacted]	[redacted]	£4,569
Scenario 2	[redacted]	[redacted]	
10% increase			£8,231
30% increase			£9,240
50% increase			£10,248
Scenario 3	[redacted]	[redacted]	-£31,427
Scenario 4	[redacted]	[redacted]	£5,752

6.3 *Conclusions of the cost effectiveness section*

In summary, the model constructed by the company appears to be logical.

The EAG has the following concerns regarding the cost-effectiveness analysis (as detailed in Section 1.1):

- Uncertainty in the estimates of relative effectiveness of mosunetuzumab compared with comparator interventions obtained from the ITCs; the EAG could not quantify the uncertainty in the model.
- Inconsistency in selection of data to extrapolate and in survival curves chosen, resulting in disconnect between clinical effectiveness and cost-effectiveness evidence.
    - This is particularly pronounced for R^2, where the company's choice of extrapolation for OS in the model results in an ICER that appears too optimistic, and the chosen extrapolation is inconsistent with information provided by the company in the addendum.

Other important factors that had an impact on the cost-effectiveness results included:

- half-cycle correction for TTOT which the EAG considered unnecessary as the cycle length was already precise enough to capture the treatment costs accurately.

The EAG have presented scenarios with a preferred base-case analysis for each pairwise comparison. The ICER has increased compared with the CS.

7 REFERENCES

1. Linton K, Julian C, Gibb A, White E, Armstrong E-F, Li Y, *et al.* Treatment patterns and outcomes in patients with relapsed/refractory follicular lymphoma who received third line therapy at the Christie NHS Foundation Trust in the UK. *Blood* 2021;**138**:5011. http://dx.doi.org/10.1182/blood-2021-148490
2. Budde LE, Sehn LH, Matasar MJ, Schuster SJ, Assouline S, Giri P, *et al.* Mosunetuzumab monotherapy is an effective and well-tolerated treatment option for patients with relapsed/refractory (R/R) follicular lymphoma (FL) who have received ≥2 prior lines of therapy: pivotal results from a phase I/II study. *Blood* 2021;**1387**(Suppl.1):127. http://dx.doi.org/10.1182/blood-2021-145872
3. Leonard JP, Trneny M, Izutsu K, Fowler NH, Hong X, Zhu J, *et al.* AUGMENT: A phase III study of lenalidomide plus rituximab versus placebo plus rituximab in relapsed or refractory indolent lymphoma. *J Clin Oncol* 2019;**37**(14):1188-99. http://dx.doi.org/10.1200/JCO.19.00010
4. Leonard J, Trneny M, Izutsu K, Fowler N, Hong X, Zhang H, *et al.* AUGMENT phase III study: Lenalidomide/rituximab (R2) improved efficacy over rituximab/placebo in relapsed/refractory follicular patients irrespective of POD24 status. *Hematol Oncol* 2019;**37**(52):114-5. http://dx.doi.org/10.1002/hon.75_2629
5. Rummel MJ, Al-Batran SE, Kim SZ, Welslau M, Hecker R, Kofahl-Krause D, *et al.* Bendamustine plus rituximab is effective and has a favorable toxicity profile in the treatment of mantle cell and low-grade non-Hodgkin's lymphoma. *J Clin Oncol* 2005;**23**(15):3383-9. http://dx.doi.org/10.1200/JCO.2005.08.100
6. Lansigan F, Parameswaran V, Yacoub A, Reynolds C, Shao S, Moore S, *et al.* Induction R2 followed by maintenance in patients with relapsed/refractory transformed follicular lymphoma: interim analysis from the phase 3b magnify study. *Hematol Oncol* 2021;**39**(Suppl. 2):404-5. http://dx.doi.org/10.1002/hon.23_2881
7. Lansigan F, Andorsky DJ, Coleman M, Yacoub A, Melear J, Fanning S, *et al.* Completed induction phase analysis of Magnify: phase 3b study of lenalidomide + rituximab (R2) followed by maintenance in relapsed/refractory indolent non-Hodgkin lymphoma [Poster presented at the 63rd American Society of Hematology Annual Meeting and Exposition, Atlanta, GA, December 2021]. *Blood* 2021;**138** 812. http://dx.doi.org/10.1182/blood-2021-145640
8. Gupta A, Schuster S, Svoboda J, Nasta S, Napier E, Walsh K, *et al.* Lenalidomide plus rituximab in patients with rituximab-resistant indolent b-cell and mantle cell lymphomas: 10-year follow-up. *HemaSphere* 2021;**5**(Suppl. 2):368-9. http://dx.doi.org/10.1097/HS9.0000000000000566
9. Chong EA, Ahmadi T, Aqui NA, Svoboda J, Nasta SD, Mato AR, *et al.* Combination of lenalidomide and rituximab overcomes rituximab resistance in patients with indolent B-cell and mantle cell lymphomas. *Clin Cancer Res* 2015;**21**(8):1835-42. http://dx.doi.org/10.1158/1078-0432.Ccr-14-2221
10. Sehn LH, Chua N, Mayer J, Dueck G, Trneny M, Bouabdallah K, *et al.* Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN): a randomised, controlled, open-label, multicentre, phase 3 trial. *Lancet Oncol* 2016;**17**(8):1081-93. http://dx.doi.org/10.1016/S1470-2045(16)30097-3
11. *A study evaluating the safety and efficacy of venetoclax (GDC-0199) plus bendamustine + rituximab (BR) in comparison with BR or venetoclax plus rituximab in*

participants with relapsed and refractory follicular non-Hodgkin's lymphoma (FNHL). ClinicalTrials.gov, U.S. National Library of Medicine 2019. URL: https://www.clinicaltrials.gov/ct2/show/NCT02187861 (Accessed 2 August 2022).
12. Budde LE, Assouline S, Sehn LH, Schuster SJ, Yoon SS, Yoon DH, *et al.* Single-agent mosunetuzumab shows durable complete responses in patients with relapsed or refractory b-cell lymphomas: Phase I dose-escalation study. *J Clin Oncol* 2022;**40**(5):481-91. http://dx.doi.org/10.1200/jco.21.00931
13. Zinzani PL, Flinn IW, Yuen SLS, Topp MS, Rusconi C, Fleury I, *et al.* Venetoclax-rituximab with or without bendamustine vs bendamustine-rituximab in relapsed/refractory follicular lymphoma. *Blood* 2020;**136**(23):2628-37. http://dx.doi.org/10.1182/blood.2020005588
14. Sehn LH, Kamdar M, Herrera AF, McMillan A, Flowers C, Kim WS, *et al.* Randomized phase 2 trial of polatuzumab vedotin (pola) with bendamustine and rituximab (BR) in relapsed/refractory (R/R) FL and DLBCL. *J Clin Oncol* 2018;**36**(15 suppl.):7507. http://dx.doi.org/10.1200/JCO.2018.36.15_suppl.7507
15. van Oers MHJ, Klasa R, Marcus RE, Wolf M, Kimby E, Gascoyne RD, *et al.* Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood* 2006;**108**(10):3295-301. http://dx.doi.org/10.1182/blood-2006-05-021113
16. van Oers MH, Van Glabbeke M, Giurgea L, Klasa R, Marcus RE, Wolf M, *et al.* Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. *J Clin Oncol* 2010;**28**(17):2853-8. http://dx.doi.org/10.1200/JCO.2009.26.5827
17. Cancer Research UK. *Follicular lymphoma*. 2020. URL: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/follicular-lymphoma (Accessed 4 January 2022).
18. European Medicines Agency (EMA). *Committee for Orphan Medicinal Products (COMP): Minutes for the meeting on 5-7 October 2021. EMA/COMP/590190/2021*. 2021. URL: https://www.ema.europa.eu/en/documents/minutes/minutes-comp-meeting_en.pdf (Accessed 2 August 2022).
19. Cancer Research UK. *Stages of non-Hodgkin lymphoma*. 2020. URL: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/stages (Accessed 2 August 2022).
20. European Medicines Agency (EMA). *Lunsumio (mosunetuzumab)*. 2022. URL: https://www.ema.europa.eu/en/medicines/human/EPAR/lunsumio (Accessed 2 August 2022).
21. European Medicines Agency (EMA). *EU/3/21/2517: Orphan designation for the treatment of follicular lymphoma: Mosunetuzumab*. 2021. URL: https://www.ema.europa.eu/en/medicines/human/orphan-designations/eu-3-21-2517 (Accessed 2 August 2022).
22. McNamara C, Montoto S, Eyre TA, Ardeshna K, Burton C, Illidge T, *et al.* The investigation and management of follicular lymphoma. *Br J Haematol* 2020;**191**(3):363-81. http://dx.doi.org/10.1111/bjh.16872
23. Dreyling M, Ghielmini M, Rule S, Salles G, Ladetto M, Tonino SH, *et al.* Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2021;**32**(3):298-308. http://dx.doi.org/10.1016/j.annonc.2020.11.008
24. National Institute for Health and Care Excellence (NICE). *Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]: Final scope*. 2022. URL:

https://www.nice.org.uk/guidance/gid-ta10816/documents/final-scope (Accessed 2 August 2022).
25. National Institute for Health and Care Excellence (NICE). *Technology appraisal guidance [TA137]: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma*. 2008. URL: https://www.nice.org.uk/Guidance/TA137 (Accessed 2 August 2022).
26. National Institute for Health and Care Excellence (NICE). *Technology appraisal guidance [TA627]: Lenalidomide with rituximab for previously treated follicular lymphoma*. 2020. URL: https://www.nice.org.uk/guidance/ta627 (Accessed 2 August 2022).
27. National Institute for Health and Care Excellence (NICE). *Technology appraisal guidance [TA629]: Obinutuzumab with bendamustine for treating follicular lymphoma after rituximab*. 2020. URL: https://www.nice.org.uk/guidance/ta629 (Accessed 2 August 2022).
28. Trotman J, Presgrave P, Carradice DP, Lenton DS, Gandhi M, Cochrane T, *et al.* Lenalidomide consolidation added to rituximab maintenance therapy in patients remaining PET positive after treatment for relapsed follicular lymphoma: phase 2 Australasian Leukaemia & Lymphoma Group NHL26 study. *Blood* 2021;**138**(Suppl. 1):2428. http://dx.doi.org/10.1182/blood-2021-146343
29. Dai D, Heo JH, Rava A, Jousseaume E, Ramos R, Bollu V. Treatment regimens and clinical outcomes among follicular lymphoma patients treated with third-line therapy in the United States: a real-world EHR study. *Blood* 2021;**138**(Suppl. 1):1360. http://dx.doi.org/10.1182/blood-2021-147638
30. Fuji S, Tada Y, Nozaki K, Saito H, Ozawa T, Kida T, *et al.* A multi-center retrospective analysis of patients with relapsed/refractory follicular lymphoma after third-line chemotherapy. *Ann Hematol* 2020;**99**(9):2133-9. http://dx.doi.org/10.1007/s00277-020-04126-y
31. Karadurmus N, Paydas S, Esin E, Surmeli ZG, Yildiz B, Erturk I, *et al.* Effectiveness of bendamustine in relapsed or refractory lymphoma cases: A Turkish oncology group study. *Arch Med Sci* 2021;**17(4)**:920-7. http://dx.doi.org/10.5114/aoms.2019.83000
32. Andorsky DJ, Coleman M, Yacoub A, Melear JM, Fanning SR, Kolibaba KS, *et al.* MAGNIFY phase IIIb interim analysis of induction R2 followed by maintenance in relapsed/refractory indolent NHL. *J Clin Oncol* 2020;**38**(15 suppl.):8046. http://dx.doi.org/10.1200/JCO.2020.38.15_suppl.8046
33. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, *et al.* Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;**25**(5):579-86. http://dx.doi.org/10.1200/jco.2006.09.2403
34. Dreyling M, Morschhauser F, Bouabdallah K, Bron D, Cunningham D, Assouline SE, *et al.* Phase II study of copanlisib, a PI3K inhibitor, in relapsed or refractory, indolent or aggressive lymphoma. *Ann Oncol* 2017;**28**(9):2169-78. http://dx.doi.org/10.1093/annonc/mdx289
35. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, *et al.* iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017;**18**(3):e143-e52. http://dx.doi.org/10.1016/s1470-2045(17)30074-8
36. *Rituximab plus lenalidomide for patients with relapsed / refractory indolent non-hodgkin's lymphoma (follicular lymphoma and marginal zone lymphoma) (AUGMENT)*. ClinicalTrials.gov, U.S. National Library of Medicine; 2022. URL: https://clinicaltrials.gov/ct2/show/NCT01938001 (Accessed 2 August 2022).
37. Rummel M, Andorsky D, Coleman M, Yacoub A, Melear J, Fanning S, *et al.* MAGNIFY: phase IIIb interim analysis of induction lenalidomide + rituximab (R2) followed by maintenance in relapsed/refractory indolent non-Hodgkin lymphoma. *Hemasphere* 2020;**4**(S1):540. http://dx.doi.org/10.1097/HS9.0000000000000404

38. European Medicines Agency (EMA). *Revlimid Assessment Report (EPAR). Procedure No. EMEA/H/C/000717/II/0107*. 2019. URL: https://www.ema.europa.eu/en/documents/variation-report/revlimid-h-c-717-ii-0107-epar-assessment-report-variation_en.pdf (Accessed 16 August 2022).
39. *A study of polatuzumab vedotin (DCDS4501A) in combination with rituximab or obinutuzumab plus bendamustine in participants with relapsed or refractory follicular or diffuse large B-cell lymphoma*. ClinicalTrials.gov, U.S. National Library of Medicine; 2021. URL: https://clinicaltrials.gov/ct2/show/NCT02257567 (Accessed 2 August 2022).
40. Hughes DA. The use of scoring systems in patients with haematological malignancy. *Acta Paediatr Suppl* 2006;**95**(451):47-51. http://dx.doi.org/10.1080/08035320600618866
41. Wake B, Hyde C, Bryan S, Barton P, Song F, Fry-Smith A, *et al.* Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma: a systematic review and economic evaluation. *Health Technol Assess* 2002;**6**(3):1-85. http://dx.doi.org/10.3310/hta6030
42. Cognet M, Taieb V, Bec M, Gauther A, Collins G. Utility values for patients with indolent non-hodgkin lymphoma (INHL) in top five european countries (EU-5). *Haematologica* 2015;**1**:578. https://library.ehaweb.org/eha/2015/20th/100086/magali.cognet.utility.values.for.patients.with.indolent.non-hodgkin.lymphoma.html?f=m1
43. Wild D, Walker M, Pettengell R, Lewis G. PCN62 UTILITY ELICITATION IN PATIENTS WITH FOLLICULAR LYMPHOMA. *Value Health* 2006;**9**(6):A294. http://dx.doi.org/10.1016/S1098-3015(10)63491-2
44. Gallacher D, Kimani P, Stallard N. Extrapolating parametric survival models in health technology assessment: a simulation study. *Med Decis Making* 2021;**41**(1):37-50. http://dx.doi.org/10.1177/0272989x20973201
45. Gallacher D, Kimani P, Stallard N. Biased survival predictions when appraising health technologies in heterogeneous populations. *Pharmacoeconomics* 2022;**40**(1):109-20. http://dx.doi.org/10.1007/s40273-021-01082-x
46. Austin PC, Thomas N, Rubin DB. Covariate-adjusted survival analyses in propensity-score matched samples: Imputing potential time-to-event outcomes. *Stat Methods Med Res* 2020;**29**(3):728-51. http://dx.doi.org/10.1177/0962280218817926
47. Haukaas FS, Ohna A, Krivasi T. Cost-Effectiveness of Obinutuzumab in Combination with Bendamustine Followed by Obinutuzumab Maintenance versus Bendamustine Alone in Treatment of Patients with Rituximab-Refractory Follicular Lymphoma in Norway. *Appl Health Econ Health Policy* 2018;**16**(4):569-77. http://dx.doi.org/10.1007/s40258-018-0401-y
48. Wang H, Smith A, Yu G, Aas E, Bagguley T, Howell D, *et al.* UK Utility Elicitation in Patients with Follicular Lymphoma. *Value Health* 2017;**20**(9):A449. http://dx.doi.org/10.1016/j.jval.2017.08.286
49. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol* 2017;**17**(1):162. http://dx.doi.org/10.1186/s12874-017-0442-1

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

External Assessment Report – factual accuracy check and confidential information check

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and External Assessment Report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the External Assessment 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 Friday 26 August 2022** 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 [redacted], and separately highlight information that is submitted as '[redacted]' in turquoise, all information submitted as '[redacted]' in yellow, and all information submitted as '[redacted] in pink.

Issue 1 Concerns over the suitability of the indirect comparisons performed and presented

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Clarification on baseline data used from AUGMENT (whole patient group or FL specific data used) (p.51)	Text to be amended since the ITC for mosunetuzumab versus R^2 were based on the FL cohort of the AUGMENT study; page 79 of the company appendices states that data from AUGMENT provide both baseline characteristics and results for RR FL patients.	Clarification	The company did not provide information regarding the source of baseline characteristics data specifically for RR FL patients, and the EAG has not been able to verify these data using other references cited in the CS in relation to AUGMENT. However, the EAG has amended the report (p.47 [not 51]) to clarify this.
Incorrect data included in Table 10 (p.53)	The data in the EAG report shows the baseline characteristics after pre-filtering to align eligibility criteria prior to the ITC. The correct columns to use to reflect the characteristics post-ITC (IPTW adjusted) are found on the far right side of Table 15 in the CS.	To include correct values	We have intended to highlight the difference in patient characteristics between GO29781 and GADOLIN after the filtering process to align inclusion/exclusion criteria before IPTW adjustment in this report section, and therefore, the data included in Table 10 are correct. Nevertheless, we have revised the preceding text

			on page 53 to make it clear that the data presented are post-filtering but before IPTW adjustment.
Uncertainty regarding how comparable trial participants are prior to matching (p.55, p57) and inaccuracy in reporting baseline characteristics for the ITT populations in Table 12 (p59)	Clarification required that all trials are in the RR FL setting and only 3L+ patients with ECOG 0-1 were used for comparison with mosun. Table 12 in the EAG report provides baseline characteristics for the ITT populations as per published references, but these are not the patients being used in the comparison therefore the company questions the value of including this. The ITC report contains tables listing the baseline characteristics on all covariates pre-ITC adjustment.	Clarification on approach taken by the company and unsuitability of comparing baseline characteristics for ITT populations from published references	We have clarified the company's approach as suggested (p.55 of EAG report). Given that all data used in the ITC were reported as academic in confidence in the CS and the separate ITC report, the EAG consider it is important to highlight the differences in the original trial populations from which the data for the ITC were drawn.
Clarification on whether all patients who underwent second randomisation were included (p61)	The EAG notes that "it was unclear whether all patients who underwent second randomisation were included, or whether only patients who were subsequently allocated to rituximab maintenance therapy at second randomisation were included". The company can confirm all patients were included since maintenance therapy is optional	Clarification	Not a factual accuracy. This was unclear in the CS. No change made.

Typo in ESS for exploratory analyses in the MAIC versus R^2 (p65)	The value of [redacted] refers to a previous iteration of the ITC report that was not updated in the latest version. The company apologizes for the typo and confirms that the correct value is [redacted], which corresponds to the analysis that was eventually selected as the company base-case. The discrepancy noted in the EAG report is thus no cause of concern and therefore this can be removed.	Clarification on error in ITC report	We have removed the following text. ~~"The company performed some exploratory analyses investigating the inclusion of the low-priority factors. Surprisingly, one of these improves the ESS compared to the company base case ([redacted]), and it is unclear why this analysis was not presented by the company as either its base case or a scenario analysis. The EAG was unable to identify any further details on this analysis that increased the ESS."~~
Further clarity on the absence of measuring DOR in ITC	Amend text to read as follows:	Clarification to reflect the company's response to	Not a factual error.

required (p68)	"The company has not presented a comparison of the duration of response (DOR) of the therapies for any data-cut as DOR and duration of complete response were not in scope for the ITC. Moreover, the company stated in response to clarification questions that baseline characteristics for responders in the FL population of AUGMENT were unavailable, therefore analysing DOR outcomes via MAICs in an unbiased way would require the strong and implausible assumption that the baseline characteristics of responders equal those of both the responder and non-responder groups combined for the comparator patient population. The company also noted that even if baseline characteristics for responders were available, these may be inherently different between treatments so adjusting responder patient populations to make them more similar across trials is not necessarily clinically meaningful and may lead to results that are difficult to interpret and/or generalise"	clarification questions on this issue	We state that data is unavailable as per the company's suggestion. No change made.
Table 15 – incorrect values are reported (p69) and wrong interpretation in text based on this error	The EAG has reported values from the August 2021 data cut for excluding low HgB rather than the [redacted] data cut. The correct values (as noted in response to question A11 in the clarification questions) are as follows:	Correction of wrong data cut used in table and resulting inaccurate interpretation of data	Table 15 has been updated to [redacted] values as requested. Linked text has been

	PFS HR (95% CI, p value): [redacted] OS HR (95% CI, p value): [redacted]As such, the text in this section should also be amended since the HRs for both PFS and OS improve, albeit not at the statistically significant threshold of p=0.05. Nevertheless, the EAGs discussion should be amended to reflect this.		deleted on page 70.
Handling of missing values for covariates (MAIC and propensity score comparisons) (p64 and 71)	The EAG has reported that the submitted ITC report does not state how missing observations were handled. However, Sections 3.1.2.2 and 3.2.2.2. of the submitted ITC report do explain the methodology employed to handle missing data in MAICs and PSAs, respectively (simple imputations using means and modes, or renormalisations). The company can provide the R code which give full details on how many missing values for baseline characteristics used in the ITCs were observed in each treatment arm.	Clarification	Not a factual error. At time of assessment the CS did not state how missing values were handled. No change made.
Misinterpretation in matching method (p72)	Please note that the following sentence is inaccurate: “Optimal pair matching was reportedly the best performing matching method and was	Inaccuracy in reporting of methods. The ITC report states the following “Notably, there was one patient in	We do not consider this a factual error no change made. The EAG consider that the

	compared with IPTW. This is despite an individual in the matching analysis initially being given a weight of 26 which was later reduced to 10 to reduce the individuals influence on the analysis, however it is possible all matching analyses encountered similar problems." This reduction in weighting was a result of IPTW, not optimal pair matching, which employs no weights.	the rituximab plus bendamustine group with a relatively high propensity score (~0.96). This resulted in a relatively high inverse probability of treatment weight (~26), which remained high even after stabilization, as depicted in Appendix G. Therefore, weight trimming to a maximum of 10 was employed to reduce the influence of this outlier weight on covariate balance and results."	results presented by the company are from optimal pair matching and not IPTW. If the company do not suggest optimal pair matching was best it is not clear to us why they would present the results in the clinical section of the CS, and omit the results from IPTW analysis. No change made. This will need to be reviewed with appropriate time during technical engagement.
Typo in ESS for RB in IPTW analysis (p72)	Please amend the sentence so that it reads as: "The ITC report recommends the IPTW analysis which has an ESS of [redacted] for mosunetuzumab and [redacted] for RB."	Clarification on error in previous ITC report, which was corrected in the latest version submitted	The company's error has been amended on page 72. (ITC received during FAC)
Error in reporting inflation in ESS (p72, p78)	The following statements in the EAG report are incorrect: "It is also unclear how the company has inflated the ESS for mosunetuzumab from [redacted] as shown in Figure 15 of the ITC report, however the EAG predicts this may	There was no inflation of ESSs for mosun in the ITCs vs RB or OB, these were [redacted].and [redacted], respectively. The EAG is misinterpreting the number at risk values in the IPTW	The EAG point still stands, it was not clear why the numbers in the KM plot did not match the ESS in the original company CS.

	be from stabilisation of the IPTW weights. If different weightings have been used for different outcomes, this is major limitation to the company's analyses." (p72) "The IPTW weights gave an ESS of [redacted] patients for mosunetuzumab and [redacted] for OB. It is again unclear how the sample sizes for mosunetuzumab and OB are inflated to [redacted] and [redacted] respectively (Figure 15 and 16 of original company submission), but the EAG predicts this is from the stabilisation of the IPTW weights." (p78)	KM curves, where the numbers displayed are the sum of stabilized weights – those KM curves do not display the ESS. Note also, the same weights were used for all outcomes in all IPTW analyses, and that weight stabilization/rescaling has no impact on KM curves or HR/OR point estimates (and therefore is not a limitation to the analysis).	Regarding the ESS on KM not matching – we need to ensure that the company's response is standard practice, prior to amending the EAG report.
Error in reporting adjusted covariates (p73)	There is an error in the following statement – POD24 should not be included. "The EAG infers from the ITC report that the covariates adjusted for in these post-hoc analyses are: time since completion of last therapy, refractory to last line of therapy, ECOG PS, prior ASCT, size of the largest node lesion, Ann Arbor stage III/IV, POD24 and refractory to prior aCD20"	POD24 did not require further adjustment as it was a balanced covariate following both optimal pair matching and IPTW.	We have removed POD24 from page 73.
Inaccuracy in reporting use of optimal pair matching (p73). Also mentioned in the bullet list on p85.	The following statement in the EAG report is not correct: "The EAG present the resulting covariate matching of the two approaches in Table 16, where it is clear that IPTW results in better matching, with the absolute standardised mean difference between fewer variables being above the 0.1	Clarification on company methods and more accurate representation of analysis conducted	See response above re optimal pair matching. We do not consider this a factual error no change made.

	threshold. Hence, the company's text in their original submission which justifies the use of optimal pair matching: "*Optimal pair matching resulted in the greatest number of balanced covariates compared with other methods and was selected as the preferred adjustment method for this comparison*", is factually inaccurate." The wording used in the CS was not clear since this should be referring to optimal pair matching being better than other matching methods that were investigated (e.g. nearest neighbour, genetic matching, etc. – see page 56 of ITC report). For clarity, the base-case analysis was conducted based on IPTW.		
Misinterpretation of approach taken for final propensity score model selection (p71 and p78)	The EAG report notes that it is unclear why specific interaction terms were included in the final propensity score models for the ITCs versus BR but not OB. The ITC report (section 3.2.5.2.1) clarifies that testing of selected interaction terms (suggested by clinician opinion) for inclusion in the PS model was a general procedure conducted for all propensity score analyses. Details as to why these were eventually included for RB but not OB can be found on page 55 and 43 of the ITC report respectively, along with the rationale for these decisions (i.e. maximization of covariate balance).	Clarification on company methods and more accurate representation of analysis conducted	Not a factual error. This is EAG opinion. It is not clear why interactions were felt necessary in one analysis and not the other. No change made.

Issue 2 Inconsistent application of adjusted and unadjusted survival data in economic analyses

Description of problem	Description of proposed amendment	Justification for amendment	
Issue 2 summary table (p.15)	This issue is based on a misunderstanding that the base-case clinical data is based on propensity score matching and the cost-effectiveness results for the comparisons with RB and OB are based on propensity score weighting. The company can confirm that the economic analyses for OB and RB are already based on IPTW data as per the ITC report base-case. This misunderstanding may have resulted from an error when generating the plots which has been corrected – updated plots have been provided supplementary to this form.	Inaccurate description of analysis conducted	[redacted]At the time of assessment, the CS was unclear. The EAG suggest this is formally investigated during technical engagement.
Inaccuracy in reporting on approach used in cost-effectiveness analysis (p72)	As described above, the following statement in the EAG report is incorrect: "The company, however, have instead chosen to implement optimal pair matching across their subsequent clinical and cost-effectiveness analyses. The reason for this inconsistency with the ITC report is not reported which is a cause of concern"	Cost-effectiveness analysis for mosun vs RB (and OB) is based on IPTW, not optimal pair matching.	See response above regarding optimal pair matching. We do not consider this a factual error no change made.

Incorrect interpretation of data used for extrapolation (p101, 102, 103, 109)	The EAG report inaccurately states that the company used unadjusted data for the extrapolations of RB and OB. As described above, the data from the plots provided is shown unadjusted for the comparators due to a bug in the code when generating the plots, but the data in the model corresponds to the adjusted analysis as can be seen in the patient count in the KM sheets in the model.	Inaccurate description of analysis conducted.	See response above. At the time of assessment, the CS was unclear. The EAG suggest this is formally investigated during technical engagement.

Issue 3 Unsupported degree of modelled benefit of mosunetuzumab over its comparators.

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Company reservations to pooling OS not accurately communicated (p97)	The company provided rationale and evidence for not pooling OS data in the response to clarification questions (B1). In addition to reservations in pooling data from heterogeneous populations and inconsistencies to observed differences in PFS, the company also provided a Cox regression analysis and KM plots on a large pool of 3L+ FL patients, demonstrating a strong and significant relationship between being a responder and improved OS in this setting. This has not been reflected in the EAG's report.	Lack of transparency on the company's position to pooling data for OS	Not a factual error. EAG opinion and critique. No change made.
Incorrect ICER recorded for	The table states that mosunetuzumab is dominated when the log-normal	Correction of error	We have made the change

EAG02 in Table 43, (p135)	extrapolation for R^2 OS using OS data from both arms is applied to the company base case. The actual ICER for this is £51,091. The company believes the EAG did not select the pooled OS scenario in the model when applying this to the base case		on page 136 Table 43.

Issue 4 TTOT and half cycle correction

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Inaccurate reflection that the model does not allow for patients receiving a lower dose, pausing or stopping R^2 treatment earlier (p99)	The following statement in the EAG report is inaccurate and should be amended: "This is consistent with the stopping rule for the MA of R^2, however, it does not consider possibility of patients receiving a lower dose, pausing or stopping therapy at an earlier point."	This is incorrect as the company has adjusted the dosing received to account for this. Dose intensity for lenalidomide was set as 0.88. For rituximab it was 0.937. See dosing sheet in the model.	We have amended the text as follows: "*This is consistent with the stopping rule for the MA of R^2, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic analysis*." The company did consider it, in the sense that the model allows for a reduced dose intensity but only when vial sharing is selected. Their base-case analysis does not allow for

			vial sharing (hence no stopping rule) so there is a nuanced argument to be made here.
Inaccurate reflection that the model does not allow for patients receiving a lower dose, pausing or stopping RB treatment earlier (p105)	The following statement in the EAG report is inaccurate and should be amended: "This is consistent with the stopping rule for the MA of RB, however, it does not consider possibility of patients receiving a lower dose, pausing, or stopping therapy at an earlier point."	As above, a dose intensity of 0.926 was assumed for bendamustine (in addition to 0.937 for rituximab)	As above. Page 106 We have amended the text as follows: "*This is consistent with the stopping rule for the MA of RB, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic analysis*."
Inaccurate reflection that the model does not allow for patients receiving a lower dose, pausing or stopping OB treatment earlier (p113)	The following statement in the EAG report is inaccurate and should be amended: "This is consistent with the MA of OB for a population refractory to rituximab, where obinutuzumab maintenance therapy is permitted if patients have either stable disease or a complete or partial response at the first 6 months. However, it does not consider possibility of patients receiving a	As above, a dose intensity of 0.937 was assumed for obinutuzumab	As above. Page 114 *This is consistent with the stopping rule for the MA of OB, however, the model does allow stopping of treatment at an arbitrary time point but this does not inform the economic*

	lower dose, pausing or stopping therapy at an earlier point..”		*analysis*.”
Inaccurate statement on degree of bias when modelling TTOT for mosun vs R^2 (p99)	The following statement in the EAG report is inaccurate and should be amended: “The difference in the modelling of TTOT may be introducing bias in favour of mosunetuzumab.”	This statement is incorrect – it is not due to a difference in modelling, it is due to how treatments are administered as per protocol, so it is a feature of the treatment of mosunetuzumab. Furthermore, the direction of the potential bias would be uncertain, as higher CR rates would imply shorter treatment duration (benefits mosunetuzumab), and lower CR rates would imply higher treatment duration (benefits comparators).	Not a factual error. EAG opinion. It is not necessarily a question of bias (as TTOT was applied to all comparators) but one of application and appropriateness which we feel is not necessary for mosunetuzumab. Removing TTOT does not change the decision on CE in all comparators.
Misinterpretation of administration time points when removing half cycle correction (p100, p106, p114)	The following statement is inaccurate as all treatments, not just mosun, were not administered at specific time points “The EAG will also explore removing the half-cycle correction for TTOT as mosunetuzumab was administered at specific time points. Furthermore, the model cycle length is already precise enough to capture the treatment costs accurately”	The statement is not accurate as there is a time window for treatment in all clinical trials. For instance, for Cycle 2, study drug infusion should occur on Day 1 of the cycle but may be given up to +- 1 day from the scheduled date (with a minimum of 6 days after Cycle 1 Day 15 dosing). For Cycle 3 and beyond, study drug infusions should occur on Day 1 of each 21-day cycle	As per responses above regarding half cycle correction. The information presented in the CS was critiqued. We will consider this additional information at technical engagement (issues 4).

	Similar statements are also noted on pages 106 and 114.	but may be given up to +- 2 days from scheduled date (with a minimum of 19 days between doses) for logistic/scheduling reasons. Other study visits starting in Cycle 2 should occur within +- 2 days from the scheduled date, unless otherwise noted. Given that the model cycles are 7 days in length, half-cycle correction would be appropriate as +-2 days is a sizeable proportion of the model cycle	

Issue 5 Immature data to model post-progression utilities

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Inaccurate representation of when PRO data were collected in GO29781 study and what was included in the clinical effectiveness section vs the economic evaluation (p.40, 115)	Text on page 40 to read as: Patient reported outcomes (PROs) were obtained in GO29781 using EORTC QLQ-C30, FACT-Lym and EQ-5D-5L. The company response to EAG clarification question A14 stated that the data collection as reported in the clinical effectiveness section "*focused on those patients who were still receiving initial treatment of mosunetuzumab*". The clinical effectiveness section of the CS,	Clarification needed as this implies that the data were only collected for patients up to cycle 8 whereas the model reflects data for patients who continued up to the maximum number of cycles as well as extra follow-up.	This is a factual error arising from inaccurate/inadequate information provided by the company. What the EAG has written was based on information supplied by the company in the CS and in the response to EAG clarification questions,

	therefore only analysed and presented PROs up to Cycle 8 of the treatment (approximately 24 weeks post baseline), and the data presented were based on evaluable patients (patients who had a baseline and at least one post-baseline assessment of PRO scales), not based on intention-to-treat. This misunderstanding has also led to inaccurate statements on page 115 ("The HRQoL estimates used in the model present a key area of uncertainty. The company stated in its submission that patient reported outcomes were analysed up to Cycle 8 of treatment") – the values used in the model include several observations for patients in post-progression that go beyond 750 days.		which gave no indication that HRQoL data collected beyond cycle 8 were analysed and included in the company's model. We have edited page 40 to say "*The CS, therefore, only presented PROs up to Cycle 8 of the treatment (approximately 24 weeks post baseline), and the data presented appeared to be based on evaluable patients (patients who had a baseline and at least one post-baseline assessment of PRO scales), not based on intention-to-treat."* We have added the following to page 40. *The EAG had no access to, and was unable to evaluate the dataset used to generate HRQoL estimates for the model.* On page 115 we have

			edited the following sentence: *"The HRQoL estimates used in the model present a key area of uncertainty. The company presented patient reported outcomes up to Cycle 8 of treatment (approximately 24 weeks post baseline) and the data presented appear to be based on patients who had a baseline and at least one post-baseline assessment of PRO scales), not based on intention-to-treat."* We note the following point in the company FAC *"the values used in the model include several observations for patients in post-progression that go beyond 750 days".* However, this was not presented in the CS.
Inaccurate statement from	The statement quoted in the ERG report	Clarification on approach taken	No change made. Not a

company submission leading to inconsistencies in reporting of sources for PPS (p116)	that "*trial data was used where possible, but data were immature for PPS and so estimates were sourced from published evidence*" is based on an error in the company submission and relates to an older version of the model where there was limited follow up from the GO29781 study to provide robust PPS utility values. This was not the case with the final submitted model and PPS utility values were sourced entirely from the GO29781 trial, therefore there is no cause for concern regarding inconsistencies in how these values were sourced	by the company	factual error. In the submission, utility values did not change in both model versions received from the company so it is unclear what the critique here and what point the company is trying to make. The comment seems to imply that the inconsistency arose as a result of them referencing an older version of the model, however this is difficult to verify. It's unclear which older version the company is referring to.

Issue 6 Proposed indication

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Unclear rationale for pursuing third-line plus indication (p.29)	The EAG report states ""The EAG could not determine a rationale for third line positioning rather than second line"	Clarification to avoid misunderstanding on the proposed indication.	Not a factual accuracy. EAG opinion. This was a concern raised by EAG

	The proposed indication is defined in the decision problem table and aligns with the registration cohort – second line use is outside of the expected marketing authorisation indication and eligibility criteria for the pivotal cohort.		clinical advisors and is clearly explained in the text. No change made.

Issue 7 SLR critique

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Description on lack of transparency in supplementary search strategy (p.36)	Full details of approach taken is available in the SLR report that has been supplied supplementary to this form. Text in the EAG report to be amended to reflect this.	To ensure company approach to the SLR is represented appropriately	Not a factual accuracy. At the time of appraisal this information was not provided. No change made.
Description on study selection in MAIC feasibility assessment (p.37)	EAG report states “insufficient details were provided regarding the subsequent process of narrowing down the number of relevant studies, initially to 42 and subsequently to 21, which were considered in the feasibility assessment for MAIC.” Full details of approach taken is available in the SLR report that has been supplied supplementary to this form. Text in the EAG report to be amended to reflect this.	To ensure company approach to the SLR is represented appropriately	Not a factual accuracy. At the time of the appraisal this information was not provided. No change made.

Inaccurate description of study selection criteria for MAIC feasibility assessment (p.38)	Point 2 amended to read: "median lines of prior therapies ≥2, i.e. ≥50% of patients included in these studies would have had at least 2 prior lines of therapy. One exception was applied to allow inclusion of the AUGMENT study as 47% FL patients in this study received 3L+ treatment, which was close to the 50% threshold.	To avoid confusion between number of prior lines received and the therapy setting which a patient is in. Clarification also around the exception made as these only applied to AUGMENT and EORTC 20981, the latter of which is a Roche sponsored study with IPD available, where the exception has no consequences, as propensity score methods would anyway be more appropriate for this comparison.	The first sentence was revised as suggested on page 38. There was no factual error in the EAG's statement, which simply highlights the post hoc nature of these exceptions to the selection criteria.
Inaccurate representation of cost/resource use SLR results (p87)	There is an inaccuracy in the EAG report resulting from a typo in the company submission Document B. Document B is referring to the results from the original SLR, which was subsequently updated in [redacted]. The results from this update are described in Appendix H and are accurate, therefore the company can confirm that 17 publications were identified, not 13 and as such the discrepancy discussed can be removed from the report.	Discrepancy due to typo in Document B. No steps were taken to narrow down the search results.	This error is due to inconsistency in reporting by the company. It does seem to have been an error in Document B (with accurate reporting in H.5-H.6). However, it is more than just a typo as they do go into quite a bit of detail about the 13 studies in B.3.5.1, so information about the additional 4 studies is missing from this

			description. We have edited the final bullet on page 88 to read "*The EAG note an error in reporting in Document B which is inconsistent to Sections H.5 H.6 of the CS appendix. In the CS, the company provide detail about the 13 studies in B.3.5.1. Information about additional identified in the updated search (Jan 2022) is missing from this description."*

Issue 8 Clinical data critique

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
In Table 5 p.43, data cut August 2021: lower 95% confidence limit for median PFS is 10.1, not 10.9 as changed in the EAR	Change the lower confidence limit back to 10.1 in Table 5	All sources provided (CS, CSR) consistently show 10.1 as the lower confidence limit as do the publications (Budde LE et al, ASH 2021 and more recently, Budde LE et al. Lancet Oncol 2022;23(8):1055-65 https://doi.org/10.1016/	Change already identified in the EAG report table 5. Note we have changed * to ^ to avoid repetition in symbols.

		S1470-2045(22)00335-7	
In Table 5 p.43, data cut August 2021: OS at 18 months is now reported from this data cut	In table 5, Data-cut August 2021, OS at 18 months, change 'not reported' to 90% (95%CI: 83, 97)	Subsequent to the CS, this has been published in Budde LE et al. Lancet Oncol 2022;23(8):1055-65 https://doi.org/10.1016/ S1470-2045(22)00335-7	Not a factual error. At the time of assessment this information was not provided in the CS.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

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 [redacted], all information submitted under [redacted], and all information submitted under [redacted] 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 **Monday 10 October 2022**. 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.

NICE National Institute for Health and Care Excellence

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)	Roche Products Ltd
Disclosure 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
Key issue 1: Concerns over the suitability of the indirect comparisons performed and presented (section 3.4.1, 3.4.2)	Yes	The EAG report has highlighted several issues related to the suitability of the ITCs conducted and their interpretation. These are individually addressed below. Please see Appendix 1 for supporting additional analyses. **Residual imbalances for the MAIC vs R^2** On page 65 of the EAG report, it is stated that "*the uncertainty surrounding the other unmatched variables means it is not possible to conclude which way the analysis may be biased*". To mitigate this concern, the company has provided an updated summary table (Table 5) for each of the three MAIC scenarios including all priority baseline characteristics reported, before and after weighting. From this table it is apparent that an important residual bias against mosunetuzumab remains in each scenario for all the factors that were not included in the adjustment (highlighted in red). The Company would also like to provide further clarifications with respect to the other two factors listed that were not considered for adjustment: ***Size of the largest lymph node lesion***: There is no need to control for both this covariate and bulky disease at the same time, as described in sections 3.1.2.1.1 and 3.2.2.1.1 of the submitted ITC report, as the two are inherently correlated given that bulky disease is derived from the size of largest lymph node lesion (longest dimension). ***Prior ASCT***: This factor was not reported in any of the source documents that were searched for baseline characteristics. However, as the majority of patients enrolled in the AUGMENT FL subgroup had only

		received 1 prior line of therapy and SCT is a treatment option considered for second or subsequent lines, the Company believes that having residual bias in this factor favouring mosunetuzumab is highly unlikely. **Misinterpretation of approach taken for final propensity score model selection** The company would like to clarify the approach taken for the selection of covariates included in the final propensity score models for the ITCs vs RB and OB as this seems to have been misinterpreted. This approach was based on the following two complementary strategies (as discussed in section 3.2.5 of the ITC report): • Testing of selected interaction terms deemed to be clinically relevant by internal medical advisors. • Testing of all possible interaction terms in an iterative backward-testing procedure (as there may have been further covariate interactions potentially relevant for our datasets in addition to those flagged by internal medical advisors that could have resulted in improved balance). Both of these strategies were tried and the latter always (and somehow unexpectedly) resulted in model overfitting (i.e. several propensity scores being equal to 0 or 1), which in turn did not result in improved covariate balance compared with not making use of interaction terms (as described in the respective result sections of the ITC report). This does not signal that no covariate interaction should be included in the propensity score model. In fact, ensuring that the propensity score function is sufficiently flexible is an explicit requirement in NICE TSD 17[1] on the analysis of observational data to inform treatment effect estimates (see also QuEENS checklist), but not in TSD 18. Conversely, when the first approach was tested, the inclusion of selected interaction terms resulted in improved covariate balance for the comparison vs RB, but not vs OB, hence why these were included for the former but not the latter (see ITC report). The respective covariate balance plots were not included in the report, but are provided in this response document, for completeness. From the love plots for the ITC vs OB (Figure 8), it is apparent that the inclusion of the selected interaction terms resulted in a roughly similar number of imbalanced covariates for full matching (the best matching method among those tested) compared with not including them (6 vs 7), but at the same time full balance could not be achieved anymore for the better balancing method (IPTW) used for the base case.

		The difference for the ITC vs RB is more nuanced, though still present. In fact, for the better balancing method used for the base case (IPTW), the love plot (Figure 9) indicates that the inclusion of the selected interaction terms resulted in imbalances in one high- and one low-priority prognostic factors compared to two high-priority prognostic factors when interactions were excluded. For optimal pair matching (the best matching method among those tested), it is apparent that the inclusion of the selected interaction terms resulted in a roughly similar number of imbalanced covariates compared to not including them (7 vs 6), though the inclusion resulted in an overall shorter distance from the 0.1 threshold used, especially when considering high priority factors such as age, refractory to last line, refractory to prior aCD20 containing regimens and POD24. To conclude, the final decision for inclusion/exclusion of interaction terms in the ITCs versus RB and OB was solely based on considerations of improvement in overall covariate balance, which is the main goal of any propensity score analysis, and not on subjectivity/medical recommendation. Their exclusion from the propensity score model for the BR ITC, as suggested by the EAG, would thus result in increased bias and yield suboptimal estimates of the relative treatment effect. **Relevance of double refractoriness (DBLRF) to both aCD20 and alkylating agent containing regimens as a confounder** On page 73 of the EAG report, when discussing the ITC vs RB, it is stated that "*This variable also has the potential to be correlated with the other variables based on refractory status already included in the propensity score calculation, effectively prioritising them in the matching analysis. No justification for its inclusion is provided, and it may be detrimental to the analysis*". The Company is unclear what is meant by "effectively prioritising", given that the model coefficients represent the independent contribution of each of the factors included. Nevertheless, the company would like to clarify its meaning behind identifying double refractoriness as an "unclear priority". This was because the feedback received from medical experts (internal and external) suggested that the evidence to assess the value of double refractoriness as an independent prognostic factor was considered to be limited and mainly based on small published subgroup analyses that did not always result in concordant results. However, it was decided to include this factor in the ITCs vs RB and OB as this could represent a confounder when comparing versus these regimens. In fact, it is clinically plausible

		for a patient who was refractory to a prior alkylating agent (such as bendamustine) to be less likely to achieve a response (and if so, of the same durability) to another alkylating agent based regimen. To demonstrate that, the Company has generated survival plots and multivariate Cox models of OS and PFS comparing all 90 FL patients on mosunetuzumab from GO29781 and all 3L+ FL ECOG 0-1 patients on RB and OB that were used in the ITCs (to represent a pooled bendamustine containing regimen arm). Three multivariate models were fitted, one controlling for all covariates that were feasible without DBLRF, one with DBLRF and the other with DBLRF both as an independent factor and interacting it with the treatment arm. This was done to assess its independent impact as a confounder or effect modifier. Plots and regressions are all included in Appendix 1. From Figure 10 and Figure 11, it is clearly evident that double refractoriness can have an important impact on survival outcomes for both mosunetuzumab and the pooled bendamustine containing regimen arm. While the multivariate Cox regressions (Table 6 to Table 13) suggest that the impact of this covariate as an independent prognostic factor may be minor, at the same time it is evident that this is likely to be a strong effect modifier for PFS (a model for OS could not be fit as no deaths occurred in the subset of mosunetuzumab patients that were not double refractory). Therefore, these findings indicate that double refractoriness is an important confounder in an ITC vs RB and OB, confirming the hypothesis put forward by the EAG on page 46 of its report when commenting on the presented GO29781 subgroup analyses ("'*refractoriness to various prior therapies might be an important effect modifier*"). Its exclusion from the propensity score model is thus likely to result in residual confounding and biased estimates of relative treatment effects. **<u>"Inflation" of ESS in the number at risk tables for the KM plots</u>** On pages 73 and 79 of the EAG report, the EAG stated that the Company somehow "inflated the ESS" in the numbers at risk tables below the KM plots, compared to the ESS reported in the main body of the CS/ITC report. The Company would like to confirm no ESS inflation has taken place. This misunderstanding may have arisen around how weights used in KM estimators are subsequently translated into number at risk tables for KM plots. The default output of this process is to generate numbers at risk at time zero which correspond to the sum of the estimated weights, and not to the ESS. To ensure the numbers at risk at time zero

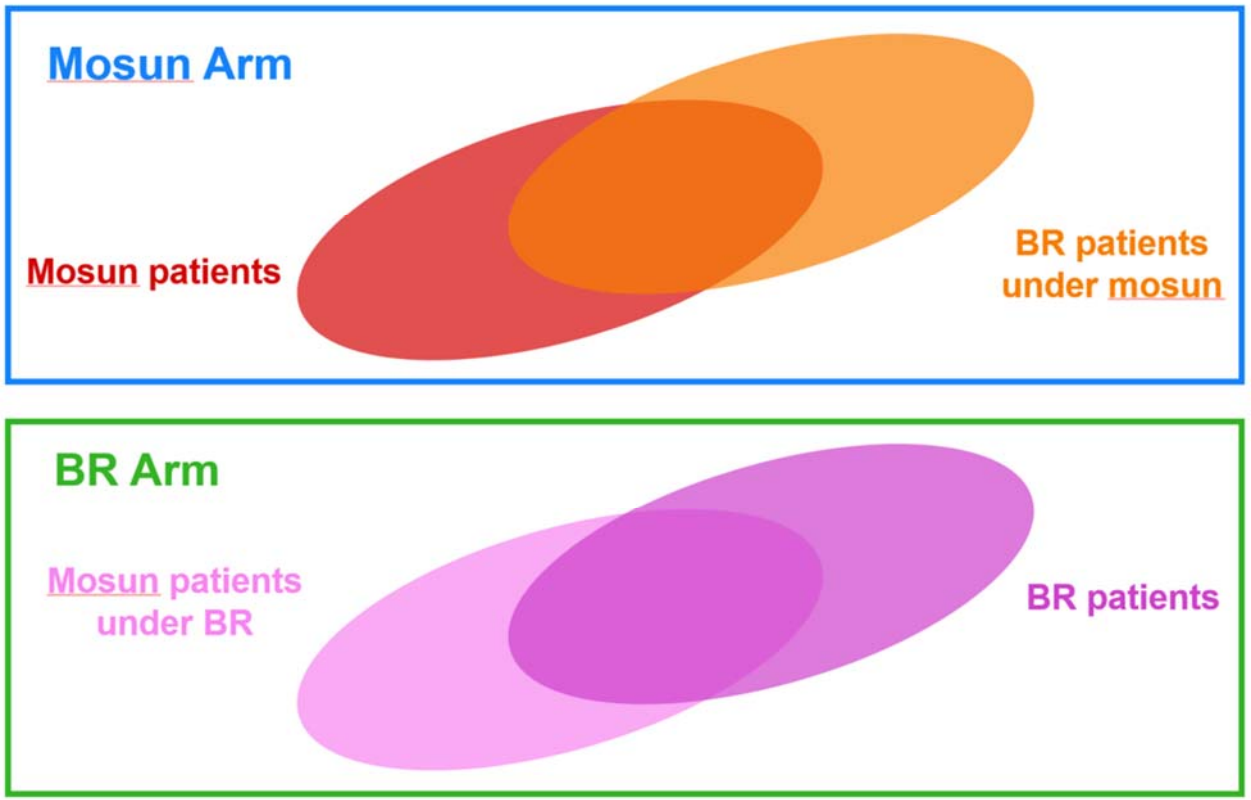
		correspond to the ESS, the estimated weights are rescaled by a common factor equal to the ESS divided by the sum of the weights. Unlike the case of MAICs, such a rescaling was not performed for IPTW and full matching because this procedure may be implemented in different ways when estimating the ATE (target estimand for the Company's propensity score analyses, as per NICE TSD 17)[1], i.e. by arm or pooled, and the Company was initially unsure of what would be the most appropriate approach. In fact, depending on the specific implementation, the sum of these rescaled weights may or may not correspond to the ESS, as some squared terms are involved in estimation of the ESS. It is important to reiterate that weight rescaling has virtually no impact on KM curves or HR/OR point estimates, and therefore not conducting it is not a limitation to the analysis. To help clarify this point, the company has re-run all ITCs by performing such a rescaling by arm. The resulting survival plots and tables with HRs/ORs have been included in Appendix 1 (Figure 12 to Figure 17 and Table 14 to Table 23). These analyses demonstrate that the numbers at risk at time zero correspond to the reported ESS and that the impact on the results is negligible and in most cases not even noticeable. **<u>Presence of individuals with potentially outlier weights</u>** On page 73 of the EAG report, it is stated that *"[..] an individual in the matching analysis initially being given a weight of 26 which was later reduced to 10 to reduce the individuals influence on the analysis, however it is possible all matching analyses encountered similar problems*." The Company would like to clarify that no individual in the matching analysis vs RB was given any weight, as weights are not involved in optimal pair matching. The Company would also like to clarify that weight truncation was only carried out for the IPTW analysis vs RB, as reported. No other cases of outlier weights or propensity scores (as described in the ITC report) were identified. **<u>Outcome model appropriateness in matching and IPTW analyses</u>** The EAG noted that the full output of the outcome models incorporating covariates to control for residual imbalances post-matching/weighting in the ITCs vs RB and OB was not provided. As a result of that, the

		EAG could not be certain that the final model used to estimate the treatment benefit is appropriate and sensible for decision making. Further information regarding the approach taken to deal with cases where outcome model estimation was problematic (e.g. due to some coefficient SEs being very large), i.e. by excluding problematic covariates from the model, is described in section 3.2.5.2 of the ITC report, while information of the specific instances where such problematic cases occurred (and for which covariate(s)), was also provided in the summary of propensity score analysis results table footnotes (e.g. see sections 4.2.2.1 and 4.2.2.5) of the respective comparisons. To help clarify, the Company has generated tables containing the full output of all such regression models, which are available in Appendix 1 (Table 24 to Table 38). These analyses demonstrate that, apart from the problematic cases already described in the ITC report (refractoriness to any prior anti-CD20 mAb regimens being excluded from the OS model for the optimal pair matching plus covariate adjustment analysis, and Ann Arbor Stage III/IV and time since completion of last therapy being excluded as covariates from the discontinuation due to AE model for the optimal pair matching plus covariate adjustment analysis), the size of coefficients and standard errors looks sensible and shows no signs of strange or worrisome behaviour of the regression models. For completeness, please note that the coefficients for variables other than treatment arm are only provided to ensure that the treatment effect estimate is unaffected by residual bias, and as such the interpretation of their size and sign is not informative. Furthermore, on page 75 of the EAG report, the EAG "*recommends estimation of effect sizes excluding all covariates from the final model and relying on the populations as balanced by the propensity score matching/weighting*". The Company wishes to clarify that the approach taken to the adjustment for residual imbalances in outcome models fully aligns with published best practices in observational research methods (see e.g. Ali et al 2019[2] for a recent example, also referenced in the NICE RWE framework[3]), and is also recommended in NICE TSD 17[1]. The Company therefore cautions against removing imbalanced covariates from the outcome models and estimating effect sizes purely based on the populations as (partially) balanced by matching/weighting, particularly as this may yield biased relative treatment effect estimates and it is the reason why such residual imbalances are controlled for in the very first place.

		Disagreement between "expected" results from unadjusted and adjusted KM curves and HRs In its critique of the ITC vs RB, the EAG raised concerns around an alleged disagreement between adjusted KM curves and HRs when compared to unadjusted ones. The Company would like to clarify that in presence of residual imbalances after matching/weighting, including these in a Cox model only controls for them at the HR level, whereas the survival curves remain affected by residual bias, as this cannot be resolved by this second adjustment. This is also mentioned in the discussion section of the ITC report. Therefore, it is not surprising in such situations to see that the difference between adjusted and unadjusted HRs is not of the same size as that between (partially) adjusted and unadjusted KM curves. This is also the reason why the Company/ITC report used IPTW for both comparison base cases rather than matching (lower overall residual balance in both effect estimates and survival curves). Moreover, this also prompted the Company to submit an additional analysis using a method inspired by that described in Austin et al 2020[4] (further elaboration on this below). **Handling of missing covariate values in ITCs** The EAG report states in several sections that "*For variables with missing data, the ITC report does not state how these observations were handled but refers to R code that has not been made available to the EAG*." The Company would like to clarify that Sections 3.1.2.2 and 3.2.2.2. of the ITC report do explain the methodology employed to handle missing data in MAICs and PSAs, respectively (simple imputations using means and modes, or renormalisations). As agreed during the Technical Engagement call, the Company has appended to this response form the first lines of the R codes for each of the submitted ITCs (Appendix 1, Codes 1-3), which give full details on how many missing values for the baseline characteristics used in the ITCs were observed in each treatment arm. From these it can be confirmed that the proportion of missing data for the individual factors was generally very low, which is unsurprising given that the data used comes from prospectively conducted clinical studies. **Concerns about the implementation of the method described in Austin et al 2020**

		Section 4.2.6.2.4 of the EAG report highlights some concerns about the implementation of a pair-wise matching plus regression adjustment approach for the ITC vs RB inspired by that described in Austin et al 2020[4]. The Company would like to clarify that the primary reasons for submitting an analysis based on such an approach were to: • Provide an additional treatment effect estimate that was as closely aligned as possible with the NICE TSD 17 recommendations (see e.g. pg 15, "The treatment effect which is typically of interest in NICE TAs is the ATE")[1], as pairwise matching does not allow to directly estimate the ATE • Comply with the explicit requirement in NICE TSD 17[1] to use at least one method to estimate the treatment effect of interest (see also QuEENS checklist) • Avoid removing the effect of the remaining imbalanced covariates (as this could already be achieved for the relative treatment effect estimate via the presented optimal pair matching plus covariate adjustment approach). The Company would also like to clarify that in the Company's implementation of this approach to estimate the local ATE, there are no residual imbalances in the baseline characteristics between the final patient groups being compared across treatment arms. In fact, to estimate the ATE "*one needs to construct/impute the counterfactuals for both the treated and the control individuals*" (see also pg 25 of NICE TSD 17[1]) and then use all patients to estimate the final treatment effect. By combining patients from the matched mosunetuzumab ("arm A") and RB groups ("arm B") to create two new treatment ("arm A" [observed] + "arm B" [predicted], for mosunetuzumab) and comparator ("arm B" [observed] + "arm A" [predicted], for RB) groups to estimate the local ATE, the method ensures by design that balance in patient baseline characteristics is achieved (as the two groups are now composed by exactly the same patients). This in turn shifts the focus of the analysis from maximising balance in baseline characteristics to ensuring the validity of the regressions models used to impute the missing potential outcomes under treatment and control for patients who received RB and mosunetuzumab, respectively, to ensure the predicted event times and ensuing treatment effect estimates are unbiased.

		This is the main methodological difference between the Company's implementation of this approach and the implementation presented by Austin et al. (apart from obvious adaptations required by the fact that we were dealing with a considerably smaller dataset of patients with an indolent disease, whereas Austin et al used either a simulated dataset [where they could decide how many patients to simulate] or a large testing RW dataset or patients with a very high prevalence condition [heart failure][4]).



In its critique of the Company's approach, the EAG has highlighted some concerns with respect to how outcome regressions and associated event times were estimated. These are individually addressed below.

Impossibility to assess the starting set of candidate covariates in the models

		The Company wishes to clarify that information on which covariates that had to be removed from the starting set of covariates tested for each outcome model (and the reasons for exclusion) was provided in both the addendum document and the updated ITC report. ***Using different regression models by arm which make use of different covariates*** The Company would like to clarify that using regression adjustment to estimate the ATE is standard procedure and is indeed considered an advantage of the method itself (see e.g. pg 37 of NICE TSD 17, "*RA fits different regressions to the treated and control groups and hence offers additional flexibility*")[1]. Forcing the two regression models to have the exact same specification would instead require the additional assumption that the relationship between covariates and outcome is similar between treated and control units (which may not necessarily hold true in practice for cases where sample sizes and/or number of events are limited). Therefore, there is no requirement for such an estimator of having to employ two models with the same specification. Furthermore, the Company notes that as the method presented combines a design-based method of bias reduction (matching) with an outcome model-based method of bias reduction (regression adjustment), the method can be considered to be doubly robust[2,5,6]. This means that "only one of the two models needs to be specified correctly to be able to identify properly the treatment effect" (see also pg 22 of NICE TSD 17[1]). Nevertheless, to further empirically show this in the specific context of the analysis presented, the Company has rerun the analyses using different model specifications than those initially presented (second, third and fourth lowest AIC models among the tested models without convergence issues, vs lowest AIC model in the initial submission). Results from these models have been included in Appendix 1 (Figure 18 to Figure 23, Table 39 to Table 53) and these confirm that the conclusions of the analyses are essentially unaffected by the different outcome models specification used. ***Censoring of event times greater than the maximum observed follow-up*** The EAG notes that the Company's approach of censoring predicted event times greater than the maximum observed follow-up "*introduces bias into the analysis*" and that "*if event times were excessively large, it could be an indication that the model adjustment is inappropriate and not fit for purpose*".

		The Company wishes to clarify that using regression models to extrapolate too far off the range of observed data may result in uncertainty and lead to unreliable predicted model outcomes. This is why the Company used the censoring approach for its implementation of the method. ***Assumption that imputed event times follow a Weibull (and not other) distribution*** The EAG was concerned that the Company inconsistently assumed the survival times follow a Weibull distribution for the missing potential outcome imputation, whereas this distribution was deemed unsuitable to fit well the data for the extrapolations used in the CE model. The Company agrees that the choice of the parametric model may have an impact on the predicted event times, although the same distributional assumption is used here to estimate missing potential outcomes under each treatment. Thus, the overall bias is expected to partially cancel out and its impact on the final treatment effect estimate is likely to be negligible. In order to mitigate this concern, the Company has rerun the analyses using alternative parametric distributions (exponential, lognormal and loglogistic) (gamma, gengamma and Gompertz are not implemented in the survreg function of the SurvParamSim package used) and has provided the results in the response document (Figure 24 to Figure 29 and Table 54 to Table 67). These confirm that the conclusions of the analyses are essentially unaffected by the use of different parametric distributional assumptions for the event times. To conclude, the combination of observed and predicted outcomes to estimate the local ATE has the advantage of increasing the sample size and thus precision of the estimate. This could be particularly beneficial in a setting like 3L+ FL (indolent and relatively rare condition, where it is difficult to find patients and events take a relatively long time to occur) and with a limited follow-up time, particularly for OS. Furthermore, combining matching and regression adjustment adjusts the IPDs twice, by first matching and then estimating individual outcomes for a complementary set of patients under each given treatment in the pooled population, thereby allowing to estimate 'doubly robust' survival curves, which cannot be done via other methods such as IPTW (see also point above). The two advantages above are not only beneficial for estimating the treatment effect (i.e. HR and 95% CIs), but also to extrapolate OS and PFS for cost-effectiveness modeling.

		In fact, this combined sample of patients provides a dataset which is enriched with more events compared to the original one and theoretically less biased survival times, and could potentially lead to more robust and less uncertain long-term extrapolations, thereby also reducing the uncertainty in the final ICER estimates. Because of such advantages, and in light of the rationale and supporting evidence provided as well as of the NICE analysis requirements described in TSD 17[1], the Company is of the opinion that the results from this analysis are both sensible and useful for decision making. Please see Appendix 3 for a scenario analysis implementing this method.
Key issue 2: Inconsistent application of adjusted and unadjusted survival data in economic analyses (section 4.2.6.2, 4.2.6.3)	No	The economic analyses have used MAIC adjusted data (comparing mosunetuzumab MAIC adjusted populations to the corresponding comparator population using pairwise comparisons) whenever appropriate, i.e. where no individual patient data was available for rituximab plus lenalidomide (R^2). When individual patient data was available for the comparator, the economic analyses used IPTW weighted data for both mosunetuzumab and the comparator (RB and OB). The company notes an inaccuracy in the original company submission which may have caused confusion, which originally presented optimal pair-matched results in Section 2.9 while IPTW results were reported in Appendix E. The Company confirms that its base case scenario for the comparison versus RB and OB for the clinical assessment was and remains based on IPTW, and that subsequent economic analyses were/are accordingly also based on IPTW. The confusion regarding the data used in the economic analysis may also have originated from the plots that were generated as part of the company submission. These plots, generated in an independent R code from the data that populates the model, were mistakenly using unweighted data for the comparators that were compared by IPTW (RB and OB). Weighted data were still used for mosunetuzumab in those cases. This means that there was a misalignment between the plots used to evaluate proportional hazards and extrapolations (unweighted data for OB and RB comparators, i.e. the IPTW comparators in the model) and the data that the model used (MAIC and IPTW weighted data, depending on the comparison).

		After noticing the error, the company regenerated the plots that affected the comparisons against RB and OB and were provided in response to the factual inaccuracy check of the EAG draft report. The proportional hazard assessment remained stable and the conclusions did not change with the updated data. Furthermore, the AIC and BIC data provided in the company submission uses the correct adjusted numbers as these numbers were taken directly from the model, so the choice of curve based on the ranking of AIC and BIC still remains valid. To conclude, the company can confirm that there is no reason for concern over the use of the adjusted data, as this was related to errors in the presented plots that did not reflect the data used in the model, nor the AIC and BIC statistics reported in the submission.
Key issue 3: Unsupported degree of modelled benefit of mosunetuzumab over its comparators (section 4.2.6)	Yes	**<u>Summary of changes to company base case</u>** The following changes have been made to the company model to help address this issue: • **Alternative parametric distributions to model survival.** Following technical engagement and consultations with clinical experts, the company has decided to review the choice of parametric distributions, to better reflect numbers of patients alive at different time points, hence presenting the most clinical plausible base case on patient survival. However, the company acknowledges that this represents a conservative view and given the limitations and potential bias against mosun with the current ITC results may not represent the true benefit of mosunetuzumab. Further data collection to provide longer-term follow up data for mosunetuzumab and more robust comparative data representative of 3L+ regimens received in UK clinical practice are required. Table 1 reports the revised base case survival distributions chosen, while Table 2 provides a summary of the proportion of patients alive in the model for each distribution and treatment arm. Values greyed out represent unrealistic estimates of proportions alive at 20 years and therefore unsuitable distributions to be considered in the revised analysis. **Table 1: Selected parametric distributions in revised company base case** | | **PFS** | **OS** |

Mosun vs R^2		
Mosun	Weibull	Weibull
R^2	Log normal	Weibull
Mosun vs RB		
Mosun	Log normal	Exponential
RB	Log normal	Exponential

Table 2: Percentage of patients alive at 20 years in model with each distribution

	Exponential	Weibull	Log normal	Generalised gamma	Log logistic	Gamma
Mosun vs R^2						
Mosun	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mosun vs RB						
Mosun	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RB	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*greyed out values indicated clinically implausible estimates

Please also note the following changes to the economic model have been applied:

- **Removal of obinutuzumab-bendamustine from revised base case analysis.** Analysis for mosunetuzumab vs OB is included in the appendix, but the company has not included it in the base case since it does not consider OB to be a relevant comparator based on current market share and clinical opinion. Clinical experts confirmed to the company that OB is very rarely used in the third-line setting since patients are likely to have received bendamustine and/or obinutuzumab in either the first- or second-line setting, and patients are infrequently retreated with these regimens. This view is supported by updated IPSOS market share data for August 2021 to July 2022, which demonstrated that OB was used in just [redacted] of patients in the third-line setting (n=[redacted]). This estimate has decreased from the previous reported analysis for 2021, which estimated OB had a [redacted] market share during that year.

		• **Application of costs for subsequent treatments**. In the previous submitted economic model, costs for subsequent treatment were applied at the point of previous treatment discontinuation. Follicular lymphoma is characterised by prolonged periods of remission (as observed in the GO29781 study where the median duration of complete response (CR) has still not been reached after a median of [redacted] months follow up in the [redacted] data cut), and applying subsequent therapy costs at the point of treatment discontinuation means these will be applied much earlier than in clinical practice. As such, the costs of subsequent treatments are now applied at the point of disease progression, which is representative of what would happen in clinical practice. In contrast, the following assumptions in the EAG's preferred base case **have not been applied** in the revised economic model (although these have been explored in scenario analyses in Appendix 3): • **Half cycle correction:** Please refer to the response to Key Issue 4 below for rationale for retaining the use of a half-cycle correction in the revised company base case. • **Pooling of OS and assuming equivalence in PFS between mosunetuzumab and R^2 at [redacted].** The company considers this to be an overly conservative assumption for numerous reasons outlined below: o Although analyses were conducted in line with NICE DSU recommended methods, the underlying populations informing the ITC were not perfectly matched since the comparator trials were not true third-line plus studies, which may **bias towards the comparators** and therefore **underestimate the clinical effectiveness of mosunetuzumab**. The assumption that the populations are homogenous enough to assume equivalence in OS therefore disregards this uncertainty. o Assuming equivalent survival between treatments **disregards the available evidence from GO29781**. For instance, the PFS KM curve from the ITC (Figure 3, [redacted] analysis report) demonstrated that [redacted] The EAG's preferred base case opted to assume equivalence in PFS

		after [redacted] however, the company understands this is a conservative assumption not reflective of the total available evidence from GO29781, which suggests the majority of patients treated with mosunetuzumab will have durable periods of CR and those patients will therefore not rapidly progress. o This is further supported from the **updated analysis from the latest data cut** which [redacted]. Moreover, the most recent [redacted], therefore the company considers the assumption that there would be no separation in the OS curves to be overly conservative. o Assuming equivalence in OS also disregards the potential **impact of surrogate endpoints**. As a partitioned survival model is used, CR is not reflected at all in the model despite the 60% CR rate observed in a poorly performing population, which according to clinical expert opinion, exceeds what is typically seen in 3L+ patients with current treatments (CR estimated at 30-40%). Studies have demonstrated that PET CR translates to a survival benefit in first-line treated FL patients[7]. While there are no published data for 3L+ patients, the company conducted an analysis that was submitted at clarification questions which pooled individual patient data ([redacted]) available in-house for RB (CONTRALTO + GO29365 trials), OB (GADOLIN trial), R-CHOP (EORTC trial) and R-monotherapy (GAUSS trial), comparing overall response rate (ORR) and CR vs OS in 3L+ FL patients with ECOG 0–1. [redacted] **Figure 1: KM plot for ORR relationship to OS**

		[redacted] **Figure 2: KM plot for CR relationship to OS** [redacted] o **Additional evidence for the clinical benefit of mosunetuzumab is now available from a comparative analysis of clinical outcomes for mosunetuzumab from GO29781 with real world data (RWD) for third-line+ patients from the US Flatiron Health Database (Appendix 4 and supporting report provided).** [redacted]

[redacted]

- In clinical practice, **mosunetuzumab would offer an additional treatment option**, which could mean that patients treated with mosunetuzumab could be eligible for more lines of treatment. It is reasonable to assume that a survival curve for third-line patients who are eligible for additional active treatments will differ to survival curves for patients who are limited to best supportive care alone. Therefore, the company consider separate parametric distributions to be appropriate in the current analysis.
- A clinical expert consulted by the company stated that they expected there to be a difference in progression-free survival profiles given the **different mechanisms of action** between the bispecific antibody mosunetuzumab, which targets both B and T cells, compared to the immunomodulatory agent lenalidomide that affects T cells only. As such, the company feels separate parametric extrapolations are more appropriate for modelling PFS between mosunetuzumab and R^2.

Details on revised economic model

Parametric survival distribution choice

Mosunetuzumab vs R^2 - Progression-free survival

For the reasons outlined above, the revised economic model does not apply the EAG assumption that PFS of mosunetuzumab and R^2 is equivalent at [redacted] after the curves cross.

AIC and BIC statistics were calculated for the six candidate distributions considered when choosing which model to reflect the base case; however, for the revised base case, the curve selection was guided by the clinical plausibility of the degree of long-term progression free survivors. As such, the Weibull model was

		selected for mosunetuzumab as this reflected the available evidence while providing estimates of patients to be progression-free at 10 ([redacted]) and 20 years ([redacted]) deemed to be appropriate by clinical experts compared to other distributions. The log normal distribution was selected for R^2 as per the EAG base case. **Figure 3: Mosunetuzumab vs R^2 comparison – PFS extrapolations** [redacted] ***Mosunetuzumab vs R^2 - Overall survival*** Similar to the discussion regarding PFS, the company considers not accounting for any survival benefit (pooling OS) to be an overly conservative assumption even in light of the limited follow up, as it does not account for the several factors outlined above. To reflect reasonable estimates of patients being alive at 10 and 20 years, the company has selected distributions that offer clinically plausible estimates at these time points. The Weibull distributions have been selected for both mosunetuzumab and R^2 as these provide realistic estimates of patients alive at 20 years ([redacted] respectively) compared to other distributions. The EAG preferred the log normal distribution to model OS estimates that over [redacted] of patients in both treatment arms will be alive after 20 years. Clinical experts considered this to be unrealistic. Please note however that in the absence of data, the Company has taken a conservative approach, although given the limitations and potential bias against mosun with the current ITC results [redacted], this may underestimate the true benefit of mosunetuzumab. **Figure 4: Mosunetuzumab vs R^2 comparison – OS extrapolations** [redacted] ***Mosunetuzumab vs RB - Progression-free survival***

		The company has amended its previous submitted base case so the distributions used to model PFS for mosunetuzumab and RB reflect those selected by the EAG – the log normal model has been selected for both treatments. **Figure 5: Mosunetuzumab vs RB comparison – PFS extrapolations** [redacted] ***Mosunetuzumab vs RB - Overall survival*** The company has revised its base case so that the exponential model is applied to both the mosunetuzumab and RB curves, as per the EAGs preference; however the company has chosen not to assume equivalence between these survival curves by pooling the OS data as outlined above. In addition to these reasons, the company does not feel it is appropriate to assume equivalence between OS for mosunetuzumab and RB given the modelled difference in PFS between treatment arms (the EAG model estimates [redacted] of patents to be progression-free at 10 years on mosunetuzumab and RB, respectively). The company considers that a PFS benefit with any treatment would translate to a difference in OS curves, even if this was only reflected at earlier time points and the curves eventually converged later on. Together with the fact that mosunetuzumab offers patients an additional line of therapy, plus the fact that [redacted] the company considers it a reasonable assumption that there would be a separation in the curves throughout the duration of the model. The exponential distribution is deemed appropriate since it is not only the best fit to the KM data based on AIC and BIC statistics but also offers clinically plausible estimates of patients alive after 20 years after treatment with mosunetuzumab and RB ([redacted] respectively). While the Weibull distribution also offers reasonable estimates of patients alive at 20 years ([redacted] respectively), the exponential distribution is associated with less uncertainty compared to the Weibull distribution, given that it is only associated with a scale parameter (no shape parameter) and has a constant failure rate, therefore making it more suitable for estimations with very few events, as is the case with OS. Given the limitations and potential underestimation of

		the mosunetuzumab treatment effect from the ITC, fitting simpler parametric models with smaller variation to the available data is likely to be preferable. **Figure 6: Mosunetuzumab vs RB comparison – OS extrapolations** 
		Deterministic base case results The deterministic cost-effectiveness results based on the revised company base case is presented in Table 69 (Appendix 3). For mosunetuzumab vs R^2, incremental costs were [redacted] and incremental QALYs were -[redacted], resulting in a cost saving ICER of £5,484 (NMB -£7,334). For mosunetuzumab vs RB, incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £37,821.
		Probabilistic sensitivity analysis Probabilistic sensitivity analysis (PSA) was performed for 1,000 iterations, randomly drawn from the parametric assumptions (Table 70). The results from the probabilistic analysis are in line with those of the deterministic analysis (ICERs of £3,995 (CS) and £35,235 vs R^2 and RB, respectively). Full results are presented in Appendix 3. **Scenario analysis** Scenario analyses have been run to help provide context to the deterministic cost-effectiveness results and aid decision making. Details of which are provided below and the outcome of each summarised in Table 3. ***Half cycle correction***

		Removing half-cycle correction for TTOT has a minor impact on the cost-effectiveness results (Table 71). Incremental costs are slightly increased resulting in a small increase in the ICERs but not to an extent that changes the overall conclusions. ***Pooling OS*** Applying the EAGs preferred assumption to assume equivalence in OS to the revised company base case has varying effects on the cost-effectiveness estimates (Table 72). For the comparison with R^2, pooling OS results in mosunetuzumab being dominant over R^2. In contrast, pooling OS between mosunetuzumab and RB increases the ICER to £203,383. This scenario analysis demonstrates the uncertainty in the cost-effectiveness results when equivalent survival between mosunetuzumab and the comparators is assumed. For the mosunetuzumab vs R^2 comparison, [redacted] ***Alternative parametric distributions*** The revised base case has applied a conservative approach to ensure modelled OS results in clinically plausible estimates of patients alive at 10 and 20 years in the model. This is despite of the fact that the clinical effectiveness of mosunetuzumab with respect to the comparators is anticipated to be underestimated given the imbalances in the matched ITC populations and underlying differences in the patient populations.

		However, given that the most recent [redacted] data cut demonstrates that [redacted], the actual benefit that relapsed/refractory patients can achieve from mosunetuzumab treatment is still yet to be realised, and only further data collection will enable this to be concluded. As such, a scenario analysis has been conducted to investigate the effect of applying alternative parametric distributions to the mosunetuzumab OS data that reflect the potential benefit offered by mosunetuzumab that may be realised with further follow up (Table 73). Distributions that estimated over [redacted] of patients alive at 20 years were disregarded in this scenario analysis. For mosunetuzumab vs R^2, applying a log logistic distribution (AIC/BIC ranking=3) to the mosunetuzumab OS data, which estimates [redacted] of patients alive at 20 years, results in increased incremental costs so that mosunetuzumab is no longer cost-saving, but also an increased in incremental QALYs ([redacted]), resulting in an ICER of £2,095. Similarly, applying a log logistic distribution (AIC/BIC ranking=3) to the mosunetuzumab OS data in the RB comparison ([redacted] of patients alive at 20 years), reduces the ICER to £29,490. In summary, while conservative base case estimates do not demonstrate cost-effective ICERs for mosunetuzumab vs the comparators, there are scenarios in which mosunetuzumab is demonstrated to be a cost-effective treatment option compared to both R^2 and RB if alternative distributions that estimate an increased OS benefit are selected. The current available data from the GO29781 study indicate that the patients continue to derive benefit from mosunetuzumab, plus comparative data with a RWD data set (Appendix 4) demonstrate [redacted], therefore further data collection will be required to confirm if these alternative scenarios can be supported with more robust data. ***Regression adjustment method for matching mosunetuzumab and RB populations*** In this scenario, the regression adjustment method previously described in our clarification question response and in the response above to Key Issue 1 has been applied to the comparison of mosunetuzumab and RB in order to estimate the missing potential outcomes and address the imbalance in the matched population and likely bias against mosunetuzumab compared to RB seen in the ITC.

The regression adjustment method involves first fitting a regression model in each treatment group based on a selection of baseline characteristics. These regression models can then be used to predict missing potential outcomes for a given treatment for each individual belonging to the other treatment group. Such unobserved outcomes can then be combined with the actual observed outcomes under each treatment, to obtain an estimate of the relative treatment effect in the overall patient population. This method combines a design-based method of bias reduction (matching) with an outcome model based method of bias reduction that is similar in spirit to 'doubly robust' methods, in that only one of the two models being estimated needs to be correctly specified to be able to properly identify the treatment effect.

Applying this method in the model with the revised company base case results in mosunetuzumab being cost-effective compared to RB, with an ICER of £14,042 (Table 74).

Table 3: Summary of scenario analyses

Scenario	ICER	Conclusion
Half-cycle correction	vs R^2: £2,997 (CS) vs RB: 38,854	Minor increase to revised base case ICERs and no impact on overall conclusions
Assuming equivalent OS	vs R^2: Mosun dominant vs RB: £203,383	Assuming equivalent OS and disregarding available evidence results in unstable and unreliable estimates.
Alternative parametric distributions	vs R^2 (log logistic distribution for mosun OS): £2,095 vs RB (log logistic distribution for mosun OS): £29,490	Alternative distributions that reflect [redacted] demonstrate mosunetuzumab to be cost-effective
Regression adjustment method for matching mosunetuzumab and RB populations	vs RB: £14,042	Implementing a combination of a design-based method of bias reduction with an outcome model based method of bias reduction (i.e. 'doubly robust'

				method) demonstrates mosunetuzumab to be cost-effective compared to RB

Conclusions

The base case has been revised to provide what the company believes to be conservative estimates of survival, given the available data. Parametric distributions to model OS were selected also based on the clinical plausibility of the degree of long-term survivors in accordance with clinical expert opinion. While this provides a more conservative view of the potential benefit of mosunetuzumab compared to the original company base case, it should be acknowledged that there is still uncertainty in the clinical effectiveness estimates. Given the imbalances in the matched populations and underlying differences in the populations used in the ITC, there is a potential for the data to be biased against mosunetuzumab, therefore this should be considered when interpreting the results. In light of this, and the updated data from the [redacted] data of GO29781 that demonstrates that patients [redacted] after [redacted] months follow up, scenarios that reflect the uncertainty in the data and the potential benefit offered by mosunetuzumab are provided (Appendix 3).

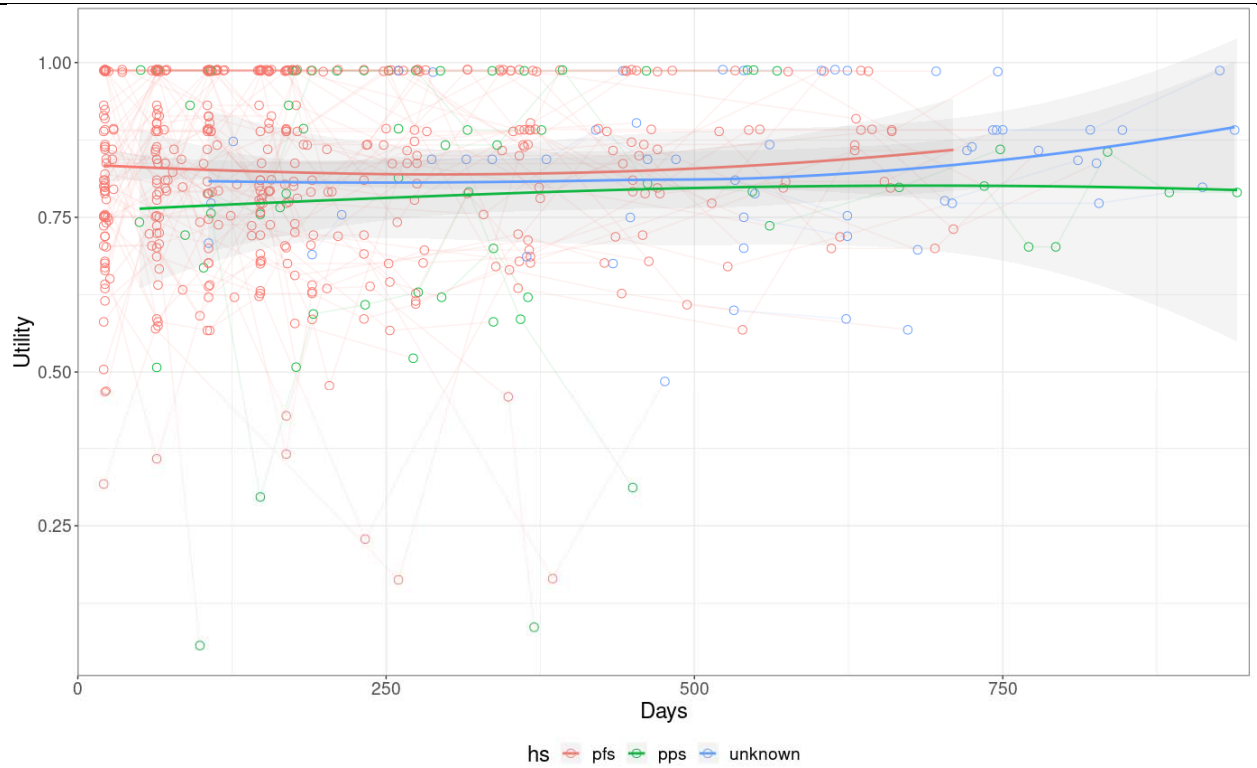
The company base case has also been amended so that costs of subsequent treatment are applied at the point of progression, rather than treatment discontinuation as this is more reflective of when patients receive their next therapy in clinical practice.

For the comparison vs R^2, mosunetuzumab is shown to be cost saving [redacted]. If the EAG's preferred assumption to assume equivalence in OS is applied, mosunetuzumab is dominant over R^2, [redacted], and if the log logistic distribution is applied to model OS, mosunetuzumab will be cost-effective with an ICER of £2,095. Overall, in the most conservative scenario, mosunetuzumab could provide cost savings to the NHS in comparison to R^2, while [redacted] should also be considered in the context of current uncertainty in the data, where the ITC is likely to underestimate the benefit of mosunetuzumab.

Mosunetuzumab is marginally above the £30,000 per QALY gained threshold vs RB in the revised base case; however, there are plausible and potentially likely scenarios in which mosunetuzumab is demonstrated to be

		cost-effective. Again, it is important to consider the uncertainty in the clinical effectiveness estimates from the ITC and the potential benefit patients may actually derive from mosunetuzumab compared to another line of immunochemotherapy, especially since mosunetuzumab is the only treatment available that provides patients with a non-rituximab and non-chemotherapy based regimen. In summary, while the revised base case provides a more conservative view on the degree of survival benefit modelled for mosunetuzumab, the scenario analysis suggests the potential for mosunetuzumab to be cost-effective against R^2 and RB. The latest analysis of the GO29781 data demonstrates that [redacted], thereby supporting the rationale for being possibly an effective and cost effective treatment. The company believes that mosunetuzumab could provide a very important and additional treatment for patients in need of further options, as cited by clinical experts who acknowledge that the complete response rate observed in GO29781 far exceeds what they see with current available treatments. Further data collection in the framework of the Cancer Drugs Fund would provide evidence to facilitate more robust cost-effectiveness estimates of mosunetuzumab.
Key issue 4: Unnecessary half cycle correction applied in the model (section 4.2.2)	No	Despite the short cycle length, the company considers it appropriate to apply a half cycle correction for TTOT since mosunetuzumab is not administered at a specific time point in the cycle. For instance: • For Cycle 2, study drug infusion should occur on Day 1 of the cycle but may be given up to +/- 1 day from the scheduled date (with a minimum of 6 days after Cycle 1 Day 15 dosing) • For Cycle 3 and beyond study drug infusions should occur on Day 1 of each 21-day cycle but may be given up to +/- 2 days from the scheduled date (with a minimum of 19 days between doses) for logistic/scheduling reasons. • Other study visits starting in Cycle 2 should occur within +/- 2 days from the scheduled date, unless otherwise noted.

		Given that the model cycles are 7 days in length, half cycle correction is appropriate as +/-2 days is a sizeable proportion of the model cycle. Furthermore, in TA604, the company did not consider it necessary to incorporate a half cycle correction given the cycle length was one week. However, the ERG deemed this to be necessary to ensure consistent application of total costs and QALY calculations. As such, the revised company base case retains the use of a half-cycle correction to align with previous appraisals, although a scenario in which this is not applied is provided (see Appendix). This scenario demonstrates that removing half-cycle correction for TTOT has a minor impact on the cost-effectiveness results; incremental costs are slightly increased resulting in a small increase in the ICERs but not to an extent that changes the overall conclusions. Therefore, inclusion of the half-cycle correction is inconsequential to the cost-effectiveness estimates.
Key issue 5: Immature data to model post-progression utilities (section 4.2.7.2)	No	All data up to the most recent follow-up was used in order to estimate the utilities in the model. Utilities included observations up to 2.5 years follow-up using the [redacted] data cut. 63 observations identified in the post-progression health state were used in the regression. Of these, 19 observations were recorded later than one year after the starting period. Observations that could not be identified due to censoring were treated as a different group. The figure below displays the utility for each patient in each of the health states identified (progression-free, post-progression, unknown). **Figure 7: Post-baseline utilities per health state through time fitted with LOESS**

		 To conclude, the company believes there is no reason for concern as the number of observations in the progressed health state is enough to be estimated with precision and the values are stable with respect to time.
NICE technical team query: Please provide details of any new studies and further data	No	Further analyses providing more robust, long-term data for mosunetuzumab from the GO29781 pivotal cohort is planned, with annual outcomes analyses conducted until at least [redacted] In the context of the available evidence to date, an additional 2–3 years follow up of the GO29781 pivotal cohort would provide valuable data to inform the long term extrapolation which may help to resolve the current uncertainties in the degree of benefit derived from mosunetuzumab.

collection, including, start date, design, comparators and locations, that could support a managed access proposal.

In addition, [redacted]

Table 4: Summary of [redacted]

Study design	[redacted]
Population	[redacted]
Intervention(s)	[redacted]
Comparator(s)	[redacted]
Outcomes	[redacted]
Indicate if study used in the NICE economic model	[redacted]
Trial start date	[redacted]
Data cut submitted to NICE	[redacted]
Anticipated data cut after a period of managed access	[redacted]

Furthermore, the company plans to generate more robust comparator data for the control arms of the ITC through sponsored projects and supporting investigator-initiated analyses of real world data. In addition, data collection through SACT could help address EAG issues 7 and 8 below, by facilitating a retrospective cohort of

		patients receiving individual chemotherapy regimens. As highlighted by the comparative analysis in the Flatiron data set (Appendix 4), [redacted] therefore a UK-specific analysis would help demonstrate the clinical benefit of mosunetuzumab in a population representative of UK patients. The company believes that the proposed data collection methods are consistent with the evidence package submitted and appraised in recent NICE CDF reviews in haematological indications. For instance, daratumumab monotherapy was recommended as an option for multiple myeloma (TA783) after CDF review on the basis of further follow up of the single arm MMY2002 study, with comparator data sourced from the SACT dataset since this represented patients in UK clinical practice. Furthermore, venetoclax was recently recommended for the treatment of CLL following its CDF review in TA796, with this reappraisal conducted on the basis of SACT collected data. In summary, the company is committed to collecting further data for mosunetuzumab and the relevant comparators, thereby reducing the uncertainty, and it considers the proposed approach to be acceptable for future decision making based on the precedence of recent CDF reviews.

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
Other issue 6: Inclusion of RB as a comparator	Section 3.3.3, Table 4	No	As highlighted in the decision problem meeting, the only rituximab plus chemotherapy option that could be assessed in the ITCs was rituximab plus bendamustine. This is because the SLR could not identify any relevant studies for R-CVP, while a propensity score analysis against R-CHOP based on the EORTC 20981 trial was not feasible due to several limitations associated with small sample sizes, unavailability of information on some important prognostic factors, and important residual imbalances after adjusting for differences in the available factors. This issue has also been discussed in previous appraisals. In TA627, clinical experts confirmed that there were no available data for R-CHOP or R-CVP in previously treated follicular lymphoma, while the committee agreed that it was inappropriate to conduct a comparison with data from EORTC 20981 given the limitations in the evidence base and missing data. Moreover, in TA604, the manufacturer was unable to conduct a comparison with individual chemotherapy regimens having not identified trials or registries that could provide these data. The committee acknowledged that this evidence would be difficult to source.
Other issue 7: Lack of suitable clinical effectiveness data for the comparison with R-CHOP	Section 3.3.4	No	

			In the current appraisal, the Company has been able to conduct a propensity score analysis using individual patient data for R-bendamustine from two company-sponsored trials. This remains the best and only option available to facilitate a comparison between mosunetuzumab and R-chemotherapy. The company acknowledges the limitation of representing a comparison of mosunetuzumab to R-chemotherapy with one regimen that may not necessarily be commonly used in the third-line plus setting. However, clinical experts confirmed to Roche that if a patient was to receive R-bendamustine in the third-line setting, the observed data would reflect what they would expect to see with this regimen in clinical practice. The company is planning to generate further comparator data for the control arms of the ITC through company-sponsored projects and the support of investigator-initiated analyses of real world data. Moreover, retrospective data collected via SACT during a managed access period may also generate a cohort of patients receiving individual chemotherapy regimens. These robust comparator data will help to address the limitations and imbalances between patient populations that is a feature of the current ITC.
Other issue 8: Generalisability of the patient cohort to the NHS	Section 3.2.1, 3.2.2	No	Roche has consulted with several UK clinical experts in the treatment of follicular lymphoma to gain feedback on whether the pivotal cohort population is representative of the UK patients that are eligible for treatment of relapsed or refractory follicular lymphoma following at least two prior therapies. There are some differences in the pivotal cohort population compared with the UK treatable population but the conclusion from the UK clinical experts is that the GO29781 pivotal cohort is representative of the patients they would treat in this indication in the context of the type of patients that are recruited for later line clinical trials in follicular lymphoma.

			The median age in the pivotal cohort population (60 years) is younger than the UK treatable population, where the median age at diagnosis of follicular lymphoma is reported as 65.6 by the Haematological Malignancy Research Network (HMRN; https://hmrn.org/statistics/quickstats). This is representative of the age of patients that are referred for clinical trials. The clinical experts highlighted that the UK median age is within the interquartile range of the age in the pivotal cohort population (53–67) years, with the range 29-90 years and the subgroup analysis demonstrated that the CR rate is consistent for patients aged <65 and >65. The ethnicity of the pivotal cohort population may not be representative of some areas of the UK. Again this reflects the types of patients that are entered into clinical trials. At a national level, the differential ethnicity categories is representative of the population; in the pivotal cohort population 82%, 9% and 4% were white, Asian and Black or African American, respectively. A between-census publication from the Office of National Statistics in 2019 reported that 85% of the England and Wales population were white, 8% were Asian/Asian British and 3.5% were Black / African / Caribbean / Black British (https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/populationestimatesbyethnicgroupandreligionenglandandwales/2019). Risk prognostic factors such as ECOG performance score, Ann Arbor stage, disease bulk, FLIPI score and refractoriness to last previous therapy in the pivotal cohort baseline characteristics are well representative of the UK treatable population according to the clinical experts. There is a higher proportion of POD24 patients in the pivotal cohort population than is expected in UK clinical practice; the proportion of patients classed as POD24 is approximately 20% of the patients completing first line treatment for follicular lymphoma, however, this proportion tends to increase in the population treated with later lines of therapy.

			The previous lymphoma therapies in the pivotal cohort population is representative of UK patients treated in the third line and beyond, with all patients exposed to alkylator and anti-CD20 therapy. However, due to the limited access to PI3K inhibitors in the UK, fewer patients would have been treated with these agents.

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)
Key Issue 3 Company's base case following technical engagement (or revised base case)	**Original base case from updated analysis:** Mosun vs R^2: £8,822 Mosun vs RB: £23,504	• Alternative parametric distributions chosen • Costs of subsequent treatment applied at disease progression (instead of treatment discontinuation) • Removal of OB	**Revised base case analysis:** Mosun vs R^2: £5,484 (cost-saving) Mosun vs RB: £37,821

Sensitivity analyses around revised base case

Please see Key Issue 3 and Appendix 3 for details.

Appendix 1: ITC analysis – additional evidence

Table 5: Summary of baseline characteristics after MAIC weighting (comparison vs R-Len)

Variable	Mosun unweighted (n=90)	Mosun-weighted (ESS=32.9) Base-case	Mosun-weighted (ESS=20.9) Scenario 1	Mosun-weighted (ESS=35.3) EAG preferred scenario	R^2 (n=147)
Age (mean)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ECOG (1 vs 0) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III–IV (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Refractory to last line (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Only one prior line of therapy (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Refractory to prior anti-CD20 containing regimen (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Low Hgb level (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone marrow involvement (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
POD24 (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bulky disease (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy >2 years (yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Presence of B-symptoms (Yes) (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 8: Love plots showing the impact of including selected interaction terms in the PS model on covariate balance (ITC vs OB)

Figure 9: Love plots showing the impact of including selected interaction terms in the PS model on covariate balance (ITC vs RB)

Figure 10: KM plot of Progression-Free Survival by Treatment (Mosunetuzumab vs Bendamustine-containing regimens) and Double Refractory Status (Yes vs No)

Figure 11: KM plot of Overall Survival by Treatment (Mosunetuzumab vs Bendamustine-containing regimens) and Double Refractory Status (Yes vs No)

Table 6:Results from Cox multivariate regression models for PFS

	Hazard ratio (95% CI) (p-value) Mosun vs Comparator (PFS)
Multivariate Cox model w/o DLBRF	[redacted]
Multivariate Cox model w/ DLBRF	[redacted]
Multivariate Cox model w/ TRT*DLBRF	[redacted]

Table 7:Results from Cox multivariate regression models for OS

	Hazard ratio (95% CI) (p-value) Mosun vs Comparator (OS)
Multivariate Cox model w/o DLBRF	[redacted]
Multivariate Cox model w/ DLBRF	[redacted]
Multivariate Cox model w/ TRT*DLBRF	[redacted]

Table 8: Summary of outcome model parameters Multivariate Cox model without DBLRF (PFS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]

Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]
Time since completion of last therapy	[redacted]	[redacted]	[redacted]

Table 9: Summary of outcome model parameters Multivariate Cox model with DBLRF (PFS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
Double Refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]

Time since completion of last therapy	[redacted]	[redacted]	[redacted]

Table 10: Summary of outcome model parameters Multivariate Cox model with DBLRF*TRT interaction (PFS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
Double Refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]
Time since completion of last therapy	[redacted]	[redacted]	[redacted]
Double refractory * Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]

Table 11: Summary of outcome model parameters Multivariate Cox model without DBLRF (OS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]

Refractory to last line	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]
Time since completion of last therapy	[redacted]	[redacted]	[redacted]

Table 12: Summary of outcome model parameters Multivariate Cox model with DBLRF (OS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]
Time since completion of last therapy	[redacted]	[redacted]	[redacted]

Table 13: Summary of outcome model parameters Multivariate Cox model with DBLRF*TRT interaction (OS)

Variable	Estimate	SE	p-value
Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior therapies ≥3	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Size of the largest node lesion	[redacted]	[redacted]	[redacted]
Low Hgb	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Bone marrow involvement	[redacted]	[redacted]	[redacted]
Time since completion of last therapy	[redacted]	[redacted]	[redacted]
Double refractory * Treatment (pooled OB - RB)	[redacted]	[redacted]	[redacted]

Figure 12: ITC results vs BR (using ESS rescaled weights) - OS KM Plots IPTW

[redacted]

Figure 13: ITC results vs BR (using ESS rescaled weights) - PFS KM Plots IPTW

Table 14:ITC results vs RB (using ESS rescaled weights) - PFS HR

	Hazard ratio (95% CI) Mosun vs Comparator (PFS)
Inverse probability of treatment weighting plus covariate adjustment	[redacted]

Table 15: ITC results vs RB (using ESS rescaled weights) - OS HR

	Hazard ratio (95% CI) Mosun vs Comparator (OS)
Inverse probability of treatment weighting plus covariate adjustment	[redacted]

Table 16: ITC results vs RB (using ESS rescaled weights) - ORR OR

	Odds ratio (95% CI) Mosun vs Comparator (ORR)
Inverse probability of treatment weighting plus covariate adjustment	[redacted]

Table 17:ITC results vs RB (using ESS rescaled weights) - ORR CR

	Odds ratio (95% CI) Mosun vs Comparator (CR)
Inverse probability of treatment weighting plus covariate adjustment	[redacted]

Table 18: ITC results vs RB(using ESS rescaled weights) - ORR WITHDRAWAL_AE

	Odds ratio (95% CI) Mosun vs Comparator (Withdrawal due to AEs)
Inverse probability of treatment weighting plus covariate adjustment	[redacted]

Figure 14:ITC results vs OB (using ESS rescaled weights) - PFS KM Plots Full matching

Figure 15: ITC results vs OB (using ESS rescaled weights) - OS KM Plots Full matching

Figure 16: ITC results vs OB (using ESS rescaled weights) - PFS KM Plots IPTW

Figure 17: ITC results vs OB (using ESS rescaled weights) - OS KM Plots IPTW

[redacted]

Table 19: ITC results vs OB (using ESS rescaled weights) - PFS HR

	Hazard ratio (95% CI) Mosun vs Comparator (PFS)
Full matching plus covariate adjustment	[redacted]
Inverse probability of treatment weighting	[redacted]

Table 20: ITC results vs OB (using ESS rescaled weights) - OS HR

	Hazard ratio (95% CI) Mosun vs Comparator (OS)
Full matching plus covariate adjustment	[redacted]
Inverse probability of treatment weighting	[redacted]

Table 21: ITC results vs OB (using ESS rescaled weights) - ORR OR

	Odds ratio (95% CI) Mosun vs Comparator (ORR)
Full matching plus covariate adjustment	[redacted]
Inverse probability of treatment weighting	[redacted]

Table 22: ITC results vs OB (using ESS rescaled weights) - CR OR

	Odds ratio (95% CI) Mosun vs Comparator (CR)
Full matching plus covariate adjustment	[redacted]
Inverse probability of treatment weighting	[redacted]

Table 23: ITC results vs OB (using ESS rescaled weights) - WITHDAWAL_AE OR

	Odds ratio (95% CI) Mosun vs Comparator (Withdrawal due to AEs)
Full matching plus covariate adjustment	[redacted]
Inverse probability of treatment weighting	[redacted]

ITC results vs OB (summary of outcome model parameters)

Table 24: Summary of outcome model parameters – PFS – Full matching plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (OB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]

FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]

Table 25: Summary of outcome model parameters – OS – Full matching plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (OB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]

Table 26: Summary of outcome model parameters – ORR – Full matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (OB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]

Table 27:Summary of outcome model parameters – CR – Full matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (OB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]

Table 28: Summary of outcome model parameters – WITHDRAWAL_AE – Full matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (OB)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
FLIPI ≥3	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]
Double refractory	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]

ITC results vs RB (summary of outcome model parameters)

Table 29: Summary of outcome model parameters – PFS – Optimal pair matching plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]

Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy (months)	[redacted]	[redacted]	[redacted]

Table 30: Summary of outcome model parameters – OS – Optimal pair matching plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen *	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy (months)	[redacted]	[redacted]	[redacted]

* Covariate could not be adjusted for as its inclusion was associated with a SE >1000 (no OS events in patients with BR who were not refractory to prior anti-CD20 regimens)

Table 31: Summary of outcome model parameters – ORR – Optimal pair matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]

Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy (months)	[redacted]	[redacted]	[redacted]

Table 32: Summary of outcome model parameters – CR – Optimal pair matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy (months)	[redacted]	[redacted]	[redacted]

Table 33: Summary of outcome model parameters – WITHDRAWAL_AE – Optimal pair matching plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV *	[redacted]	[redacted]	[redacted]
Prior ASCT	[redacted]	[redacted]	[redacted]
Refractory to last line	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen	[redacted]	[redacted]	[redacted]

Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy (months) *	[redacted]	[redacted]	[redacted]

* Covariates could not be adjusted for as their inclusion was associated with SEs >1000 (very few discontinuation events observed in the BR arm)

Table 34: Summary of outcome model parameters – PFS – IPTW plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]

Table 35: Summary of outcome model parameters – OS – IPTW plus covariate adjustment

Variable	Estimate	SE	p-value
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]

Table 36: Summary of outcome model parameters – ORR – IPTW plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]

Table 37: Summary of outcome model parameters – CR – IPTW plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]

ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]

Table 38:Summary of outcome model parameters – WITHDRAWAL_AE – IPTW plus covariate adjustment

Variable	Estimate	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Arm (RB)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0)	[redacted]	[redacted]	[redacted]
High LDH	[redacted]	[redacted]	[redacted]

Initial lines of R code reporting how many missing values for BL characteristics were observed in the ITC cohorts and how they were handled

Code 1 - ITC vs RB

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Code 2 - ITC vs OB

[redacted]

[redacted]

[redacted]

[redacted]

Code 3 - MAIC vs R^2

[redacted]

[redacted]

[redacted]

[redacted]

Sensitivity analyses for the matching plus regression adjustment approach (Austin 2020) for the comparison vs RB

Using different sets of covariates for the final survival time models

Figure 18: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - PFS KM Plots

Figure 19: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - OS KM Plots

Table 39: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming log-normal distr. of event times)
Regression adjustment and matching using Austin 2020 (PFS)	[redacted]
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 40: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]
POD24	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 41: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]

Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 42: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0) (%)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm]	[redacted]	[redacted]	[redacted]
Low Hgb (Yes) (%)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 43: ITC results vs RB (Austin 2020, using second lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Figure 20: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - PFS KM Plots

[redacted]

Figure 21: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - OS KM Plots

[redacted]

Table 44: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming log-logistic distr. of event times)
Regression adjustment and matching using Austin 2020 (PFS)	[redacted]
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 45: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 46: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 47: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]

Low Hgb (Yes) (%)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 48: ITC results vs RB (Austin 2020, using third lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Figure 22: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - PFS KM Plots

[redacted]

Figure 23: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - OS KM Plots

[redacted]

Table 49: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming exponential distr. of event times)
Regression adjustment and matching using Austin 2020 (PFS)	[redacted]
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 50: ITC results vs BR (Austin 2020, using fourth lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value

(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 51: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 52: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
Prior therapies ≥3 (%)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 53: ITC results vs RB (Austin 2020, using fourth lowest AIC model w/o convergence issues) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value

(Intercept)	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Assuming different parametric distributions for event times

Figure 24: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - PFS KM Plots

[redacted]

Figure 25: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - OS KM Plots

[redacted]

Table 54: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming log-normal distr. of event times)
Regression adjustment and matching using Austin 2020 (PFS)	[redacted]
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 55: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 56: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 57: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0) (%)	[redacted]	[redacted]	[redacted]
Ann Arbor Stage III/IV (Yes) (%)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
POD24 (Yes) (%)	[redacted]	[redacted]	[redacted]
Low Hgb (Yes) (%)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 58: ITC results vs RB (Austin 2020, Assuming log-normal distribution of event times) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Figure 26: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - PFS KM Plots

Figure 27: ITC results vs BR (Austin 2020, Assuming log-logistic distribution of event times) - OS KM Plots

Table 56. ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming log-logistic distr. of event times)
Regression adjustment and matching using Austin 2020 (PFS)	[redacted]
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 59: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 60: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 61: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
ECOG PS (1 vs 0) (%)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Low Hgb (Yes) (%)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Table 62: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
FLIPI ≥3 (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
Log (scale)	[redacted]	[redacted]	[redacted]

Figure 28: ITC results vs RB (Austin 2020, Assuming exponential distribution of event times) - PFS KM Plots

[redacted]

Figure 29: ITC results vs RB (Austin 2020, Assuming exponential distribution of event times) - OS KM Plots

[redacted]

Table 63: ITC results vs RB (Austin 2020, Assuming exponential distribution of event times) - HRs

	Hazard ratio (95% CI) Mosun vs Comparator (assuming exponential distr. of event times)
Regression adjustment and matching using Austin	[redacted]

2020 (PFS)	
Regression adjustment and matching using Austin 2020 (OS)	[redacted]

Table 64: ITC results vs RB (Austin 2020, Assuming exponential distribution of event times) - Summary of model parameters- PFS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to any prior anti-CD20 mAb containing regimen (Yes) (%)	[redacted]	[redacted]	[redacted]

Table 65: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- PFS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]
Time since completion of last therapy [months] (mean)	[redacted]	[redacted]	[redacted]

Table 66: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- OS final regression adjustment model for rituximab plus bendamustine

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Age (mean)	[redacted]	[redacted]	[redacted]
Prior therapies ≥3 (%)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]
High LDH (Yes) (%)	[redacted]	[redacted]	[redacted]

Table 67: ITC results vs RB (Austin 2020, Assuming log-logistic distribution of event times) - Summary of model parameters- OS final regression adjustment model for mosunetuzumab

Variable	Estimate	SE	p-value
(Intercept)	[redacted]	[redacted]	[redacted]
Refractory to last line (Yes) (%)	[redacted]	[redacted]	[redacted]
Size of the largest node lesion [cm] (mean)	[redacted]	[redacted]	[redacted]

Appendix 2: Updated clinical efficacy from GO29781

This Appendix provides a summary of the updated clinical effectiveness results for mosunetuzumab in relapsed/refractory follicular lymphoma in patients who have received 2 or more prior lines of therapy. This analyses is based on the [redacted] data cut of the GO29781 study, which provides further follow up based on the previous data cut [redacted]) presented in the updated analysis that was submitted with the responses to the EAG's clarification questions. This data cut is limited to investigator-assessed endpoints only and as such should be used for descriptive purposes only to demonstrate that the clinical benefit of mosunetuzumab is ongoing with a median of [redacted] months follow up. Due to the absence of independent review-facility assessed endpoints, the cost-effectiveness analysis and indirect treatment comparisons have not been updated.

Overall response rate

As of the [redacted] data cut-off, investigator assessed overall response rate (ORR) was [redacted]) and investigator-assessed CR rate was [redacted] [8].

Further details of response assessment are provided in Table 68.

Table 68. Tumour response in the GO29781 pivotal cohort, July 2022 data cut (N=90)

Best overall response by IRF - with or without PET Scan*	
Responders, n (% with 95% CI)	[redacted]
Non-responders, n (%)	[redacted]
Response classification by IRF	
CR, n (% with 95% CI), primary endpoint	[redacted]
PR, n (% with 95% CI)	[redacted]
SD, n (% with 95% CI)	[redacted]
PD**, n (% with 95% CI)	[redacted]
Not evaluable	[redacted]
Missing or not done	[redacted]

* FDG PET and CT scans were required for response assessments during study treatment and CT scans with or without PET scans could be utilised during post-treatment follow up.
** PD includes missing, not evaluable and not done (ND) assessments where the patient has otherwise had a PD observed by an adverse event of Malignant Neoplasm Progression, Discontinuation due to Progressive Disease or Non-Radiological Progression observed by Investigator.
Abbreviations: CI, confidence interval; CR, complete response; IRF, independent review facility; PD, progressive disease; PET, positron emission tomography; PR, partial response; SD, stable disease.

DOR per investigator

As of the [redacted] data cut-off, of the [redacted] patients who achieved a response as determined by the investigator, [redacted] patients ([redacted]) subsequently experienced disease progression ([redacted]) or died ([redacted]).

[redacted]). A total of [redacted], [redacted] and [redacted] of patients respectively, remained in response at 12, 18 months and 24 months.

This is consistent with the [redacted]data cut, in which , [redacted] patients ([redacted]) had experienced disease progression ([redacted]) or died ([redacted]).
[redacted], and a total of [redacted] and [redacted] of patients of patients respectively, remained in response at 12 and 18 months.

The updated data confirms that patients continue to have durable responses to mosunetuzumab.

Duration of CR per investigator

As of the [redacted] data cut-off, among the [redacted] patients who achieved a CR as determined by the investigator, [redacted] subsequently had disease progression. [redacted].
Of patients achieving CR per investigator assessment, [redacted] remained in CR at 12, 18 and 24 months respectively.

Duration of response in complete responders

As of the [redacted] data cut-off,
[redacted]. Of patients achieving CR per investigator assessment, a response was ongoing in [redacted] of patients at 12, 18 and 24 months respectively.

Progression-free survival

As of the [redacted] data cut-off, [redacted] had a PFS event as assessed by the investigator, including [redacted] patients experiencing disease progression and [redacted] deaths. [redacted] The 12, 18, and 24- months progression-free rates were [redacted] respectively.

This is consistent with the data from the [redacted]data cut, in which [redacted] patients ([redacted] had a PFS event (disease progression n=[redacted], or death n=[redacted]) and [redacted]. The 12- and 18-month PFS rates were [redacted]

Overall survival

At the time of the [redacted] data cut off, [redacted], while 12-, 18- and 24-month survival rates were [redacted], respectively.

This updated data is consistent with the results from the [redacted]data cut-off, in which [redacted] patients had died ([redacted] while 12- and 18-month survival rates were [redacted] and [redacted] respectively.

Summary

This most recent data from the GO29781 pivotal cohort provides further follow up for the clinical efficacy of mosunetuzumab, as assessed by investigators, with a median follow up of

[redacted] months. The current data confirms that [redacted] In summary, these data confirm that the true benefit that heavily pre-treated patients can achieve from mosunetuzumab treatment is still yet to be realised, therefore there is a need for further follow up of this cohort of the GO29781 study.

Appendix 3: Revised cost-effectiveness analysis

Deterministic base case results

The deterministic cost-effectiveness results based on the revised company base case is presented in Table 69.

For mosunetuzumab vs R^2, incremental costs were [redacted] and incremental QALYs were -[redacted], resulting in a cost saving ICER of £5,484 (NMB -£7,334).

For mosunetuzumab vs RB, incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £37,821.

Table 69. Deterministic company base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.780	[redacted]	£5,484 (CS)	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	£37,821	[redacted]

Probabilistic base case results

Table 70. Probabilistic company base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	8.98	[redacted]					
R^2	[redacted]	9.96	[redacted]	[redacted]	-0.98	[redacted]	£3,995 (CS)	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.93	[redacted]					
RB	[redacted]	8.20	[redacted]	[redacted]	1.74	[redacted]	£35,235	[redacted]

Figure 30: Cost-effectiveness acceptability curve- R^2 comparison

[redacted]

Figure 31: Incremental cost-effectiveness plane - R^2 comparison

[redacted]

Figure 32: Cost-effectiveness acceptability curve - RB comparison

[redacted]

Figure 33: Incremental cost-effectiveness plane - RB comparison

[redacted]

Scenario analyses results

Table 71. Scenario analysis – no half cycle correction (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.78	[redacted]	£2,997 (CS)	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	38,854	[redacted]

Table 72. Scenario analysis – pooled OS (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	10.51	[redacted]					
R^2	[redacted]	10.51	[redacted]	[redacted]	0	[redacted]	Mosun dominant	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.23	[redacted]					
RB	[redacted]	9.23	[redacted]	[redacted]	0	[redacted]	£203,383	[redacted]

Table 73. Scenario analysis – alternative distributions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs R^2								
Base case: Mosun OS – Weibull; R^2 OS - Weibull								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.780	[redacted]	£5,484 (CS)	[redacted]
Scenario: Mosun OS – log logistic; R^2 OS - Weibull								
Mosunetuzumab	[redacted]	10.84	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	0.48	[redacted]	£2,095	[redacted]
Mosunetuzumab vs RB								
Base case: Mosun OS – Exponential; RB OS - Exponential								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	£37,821	[redacted]
Scenario: Mosun OS – Log logistic; RB OS - Exponential								
Mosunetuzumab	[redacted]	10.49	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	2.19	[redacted]	£29,490	[redacted]

Table 74. Scenario analysis – regression adjustment for mosun vs RB (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	12.97	[redacted]					
RB	[redacted]	7.78	[redacted]	[redacted]	5.19	[redacted]	£14,042	[redacted]

Mosunetuzumab vs OB

Analysis for mosunetuzumab vs OB is included for completeness as OB was included in the scope of the current appraisal, but the company would like to reiterate that it does not consider OB to be a relevant comparator based on current market share and clinical opinion. Clinical experts confirmed to the company that OB is very rarely used in the third-line setting since patients are likely to have received bendamustine and/or obinutuzumab in either the first- or second-line setting, and patients are infrequently retreated with these regimens. This view is supported by updated IPSOS market share data for August 2021 to July 2022, which demonstrated that OB was used in just [redacted] of patients in the third-line setting ([redacted]). This estimate has decreased from the previous reported analysis for 2021, which estimated OB had a [redacted] market share.

Parametric survival distribution choice

Progression-free survival

The company has amended its previous submitted base case so the distributions used to model PFS for mosunetuzumab and OB reflect those selected by the EAG – the log normal model has been selected for mosunetuzumab and exponential for OB.

Figure 34: Mosunetuzumab vs OB comparison – PFS extrapolations

[redacted]

Overall survival

The company has amended its base case so that the exponential distribution is applied to both treatments, since this model is the best fit to the data based on AIC and BIC statistics, and it provides clinically plausible estimates of patients alive after 20 years ([redacted] for mosunetuzumab and OB, respectively). The log-normal distribution was not deemed appropriate due to a clinically implausible crossing of the curves at around 120 months.

Figure 35: Mosunetuzumab vs OB comparison – OS extrapolations

[redacted]

Deterministic results for mosunetuzumab vs OB

Table 75. Deterministic company base case cost-effectiveness results with revised assumptions (PAS discount) – Mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	7.48	[redacted]	[redacted]	0.89	[redacted]	£1,760	[redacted]

Probabilistic base case results for mosun vs OB

Table 76. Probabilistic company base case cost-effectiveness results with revised assumptions (PAS discount) – Mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	7.49	[redacted]	[redacted]	0.87	[redacted]	£1,358	[redacted]

Figure 36: Cost-effectiveness acceptability curve - OB comparison

Figure 37: Incremental cost-effectiveness plane - OB comparison

Scenario analyses results for mosun vs OB

Table 77. Scenario analysis – no half cycle correction (PAS discount) – mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	7.48	[redacted]	[redacted]	0.89	[redacted]	£3,895	[redacted]

Appendix 4: Comparative analysis of outcomes of relapsed/refractory FL patients after at least two prior lines of systemic therapy treated with mosunetuzumab in GO29781 with patients treated in the routine clinical practice setting in the US

Please refer to the confidential report provided with this response for full details of this analysis.

References

1 Faria, R., Hernandez Alava, A., M,, Manca, A. & Wailoo, A., J. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data. *Available from http://www.nicedsu.org.uk* (2015).

2 Ali, M. S. *et al.* Propensity Score Methods in Health Technology Assessment: Principles, Extended Applications, and Recent Advances. *Frontiers in pharmacology* **10**, 973, doi:10.3389/fphar.2019.00973 (2019).

3 National Institute for Health and Care Excellence. *NICE real-world evidence framework*, <https://www.nice.org.uk/corporate/ecd9/chapter/overview> (2022).

4 Austin, P. C., Thomas, N. & Rubin, D. B. Covariate-adjusted survival analyses in propensity-score matched samples: Imputing potential time-to-event outcomes. *Stat Methods Med Res* **29**, 728-751, doi:10.1177/0962280218817926 (2020).

5 Funk, M. J. *et al.* Doubly Robust Estimation of Causal Effects. *American Journal of Epidemiology* **173**, 761-767, doi:10.1093/aje/kwq439 (2011).

6 Kreif, N., Grieve, R. & et al. Regression-adjusted matching and double-robust methods for estimating average treatment effects in health economic evaluation. *Health Services and Outcomes Research Methodology* **12**, 174-202 (2013).

7 Tan, T. D. *et al.* The Impact of First Complete Remission by PET-CT and Time to Next Treatment on Survival of Follicular Lymphoma Patients. *Clinical hematology international* **1**, 168-172, doi:10.2991/chi.d.190528.001 (2019).

8 Budde, L. E. *et al.* Safety and efficacy of mosunetuzumab, a bispecific antibody, in patients with relapsed or refractory follicular lymphoma: a single-arm, multicentre, phase 2 study. *The Lancet Oncology* **23**, 1055-1065, doi:10.1016/S1470-2045(22)00335-7 (2022).

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma

Updated Analysis

November 2022

File name	Version	Contains confidential information	Date
ID3931_Mosun RRFL_Updated Analysis_Nov2022_CIC	1.0	Yes	30 November 2022

Summary

This document provides updated cost effectiveness results for mosunetuzumab in relapsed/refractory follicular lymphoma in patients who have received 2 or more prior lines of therapy. This analyses is based on the [redacted] data cut of the GO29781 study and takes into account a revised PAS, which has been increased from [redacted] in the original company submission to [redacted]

This amendment to the PAS has been made to facilitate NICE committee deliberations on the suitability of mosunetuzumab for entry into the Cancer Drugs Fund. This update reflects the [redacted]. The revised analysis detailed below demonstrates plausibly cost-effective ICERs for mosunetuzumab compared to both comparators in the company base case, and consequently this represents the company's full and final offer to last for the duration of any time spent in the CDF only. The Company is confident that further data collection in the CDF will enable the cost-effectiveness estimates for mosunetuzumab to improve even further and will therefore allow for this price to be re-evaluated at the end of the CDF funding period.

Results

Deterministic base case results

The deterministic cost-effectiveness results based on the revised company base case with updated PAS is presented in Base-case results

Table 1.

For mosunetuzumab vs rituximab-lenalidomide (R^2), incremental costs were [redacted] and incremental QALYs were [redacted], resulting in a cost saving ICER of £27,422 (NMB [redacted]).

For mosunetuzumab vs rituximab-bendamustine (RB), incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £28,333.

Sensitivity analysis

Probabilistic ICERs are consistent with the deterministic base case (Table 2). These results are in line with those of the deterministic analysis in terms of the estimated QALY and LY gains and the estimated incremental costs demonstrating that the deterministic base case results are likely to represent the average experience per person treated with mosunetuzumab.

Scenario analyses

These analyses demonstrate scenarios in which mosunetuzumab is even more cost-effective compared to R^2 and RB (Table 3–Table 6), i.e. assuming alternative distributions for mosunetuzumab OS (vs both comparators), assuming equivalent survival for

mosunetuzumab and R^2, and applying the regression adjustment method for matching mosunetuzumab and RB populations.

Base-case results

Table 1. Deterministic company base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.78	[redacted]	£27,422 (CS)	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	£28,333	[redacted]

Deterministic sensitivity analysis

Figure 1. Tornado diagram showing OWSA results on NMB – Mosun vs R^2

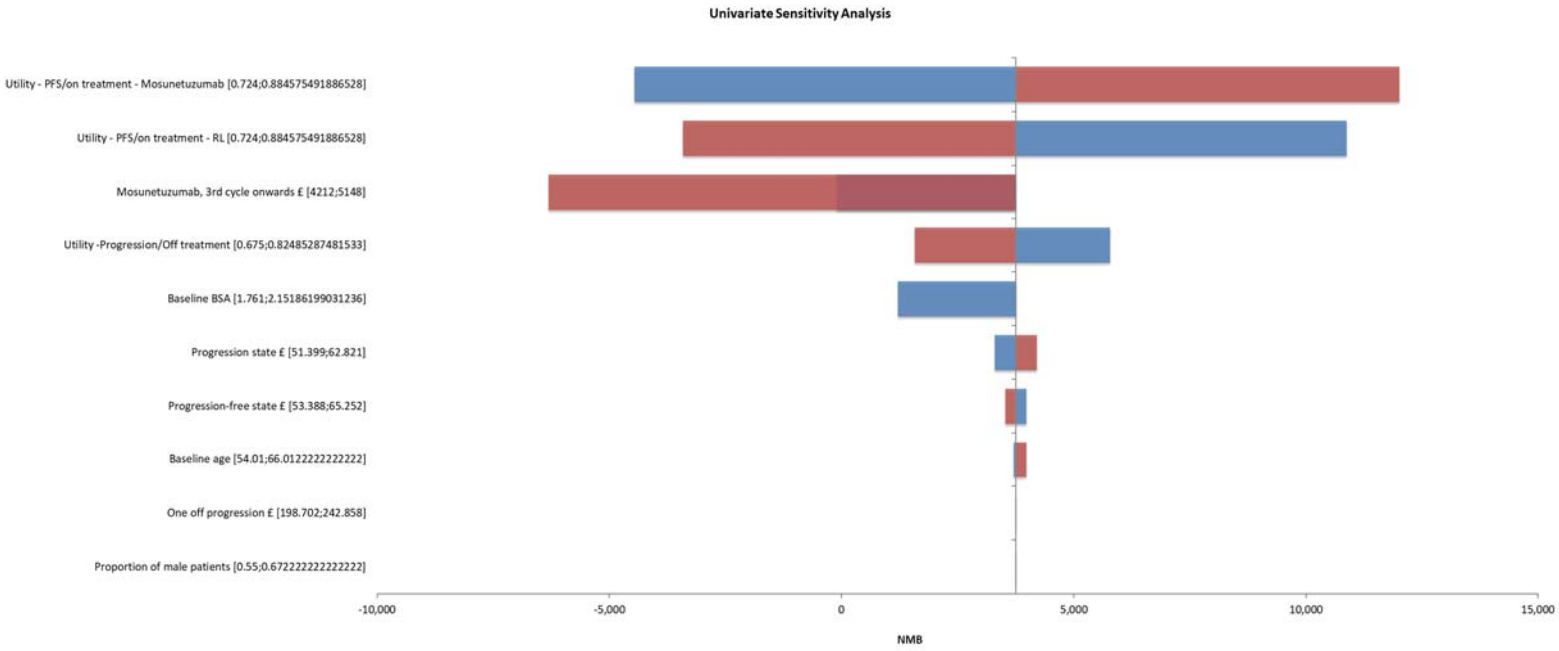
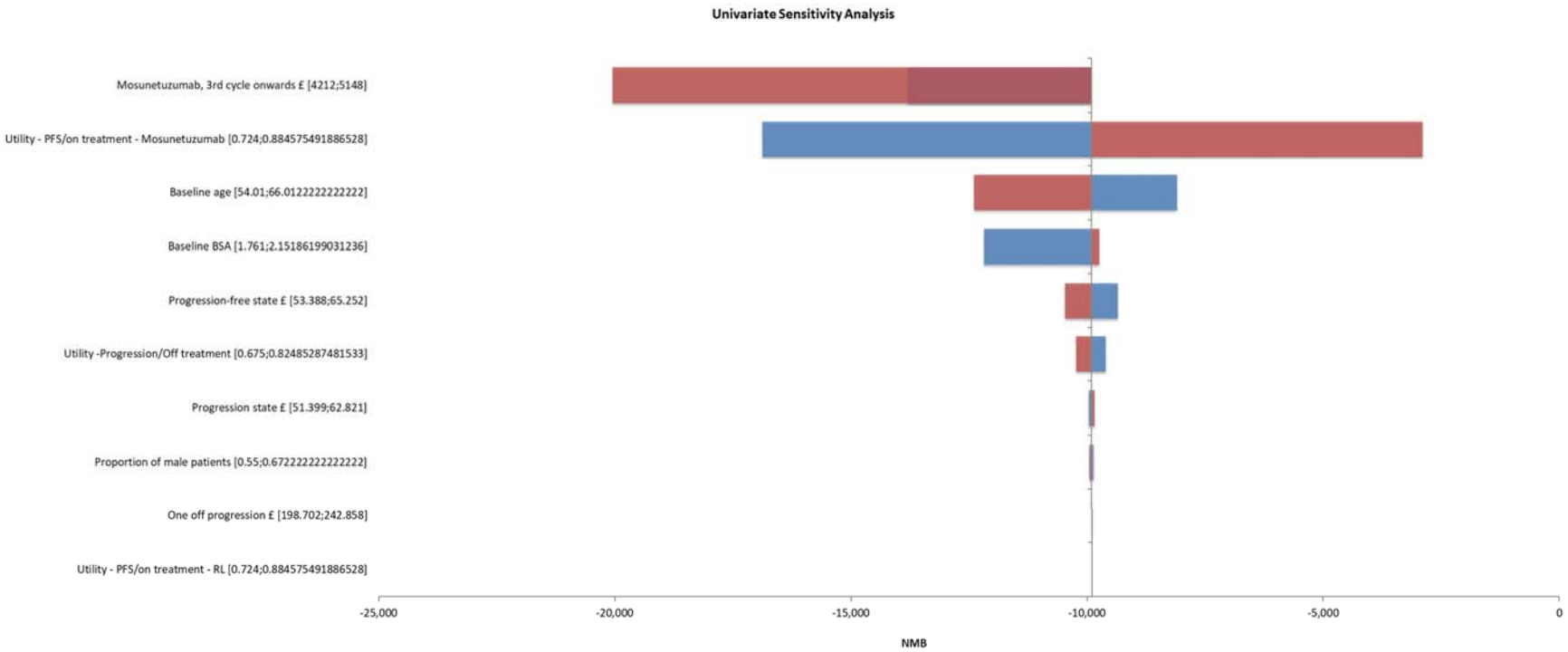


Figure 2. Tornado diagram showing OWSA results on NMB – Mosun vs RB



Probabilistic sensitivity analysis

Table 2. Probabilistic company base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	9.18	[redacted]					
R^2	[redacted]	9.94	[redacted]	[redacted]	-0.77	[redacted]	£23,102	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.95	[redacted]					
RB	[redacted]	8.27	[redacted]	[redacted]	1.69	[redacted]	£27,069	[redacted]

Figure 3: Cost-effectiveness acceptability curve- R^2 comparison

Figure 4: Incremental cost-effectiveness plane - R^2 comparison

Figure 5: Cost-effectiveness acceptability curve - RB comparison

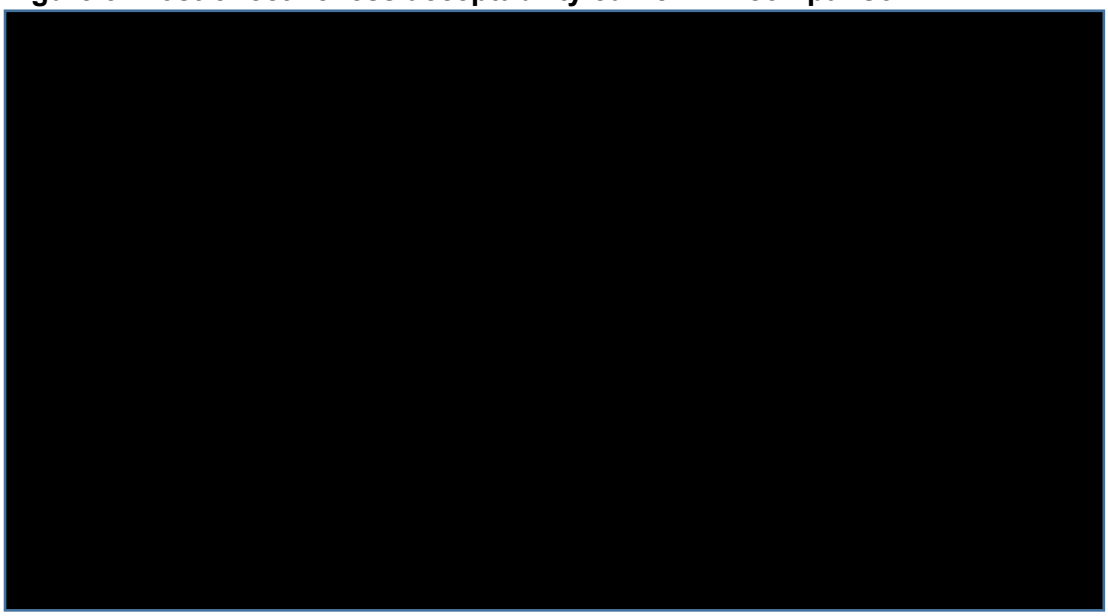


Figure 6: Incremental cost-effectiveness plane - RB comparison

Scenario analyses

Table 3. Scenario analysis – no half cycle correction (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.78	[redacted]	£25,493 (CS)	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	£29,134	[redacted]

Table 4. Scenario analysis – pooled OS (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs R^2								
Mosunetuzumab	[redacted]	10.51	[redacted]					
R^2	[redacted]	10.51	[redacted]	[redacted]	0	[redacted]	Mosun dominant	[redacted]
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	9.23	[redacted]					
RB	[redacted]	9.23	[redacted]	[redacted]	0	[redacted]	£143,234	[redacted]

Table 5. Scenario analysis – alternative distributions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs R^2								
Base case: Mosun OS – Weibull; R^2 OS - Weibull								
Mosunetuzumab	[redacted]	9.58	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	-0.78	[redacted]	£27,422 (CS)	[redacted]
Scenario: Mosun OS – log logistic; R^2 OS - Weibull								
Mosunetuzumab	[redacted]	10.84	[redacted]					
R^2	[redacted]	10.36	[redacted]	[redacted]	0.48	[redacted]	Mosun dominant	[redacted]
Mosunetuzumab vs RB								
Base case: Mosun OS – Exponential; RB OS - Exponential								
Mosunetuzumab	[redacted]	9.90	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	1.60	[redacted]	£28,333	[redacted]
Scenario: Mosun OS – Log logistic; RB OS - Exponential								
Mosunetuzumab	[redacted]	10.49	[redacted]					
RB	[redacted]	8.30	[redacted]	[redacted]	2.19	[redacted]	£22,910	[redacted]

Table 6. Scenario analysis – regression adjustment for mosun vs RB (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs RB								
Mosunetuzumab	[redacted]	12.97	[redacted]					
RB	[redacted]	7.78	[redacted]	[redacted]	5.19	[redacted]	£11,020	[redacted]

Mosunetuzumab vs obinutuzumab-bendamustine (OB)

Analysis for mosunetuzumab vs OB is included for completeness as OB was included in the scope of the current appraisal, but the company would like to reiterate that it does not consider OB to be a relevant comparator based on current market share and clinical opinion. Clinical experts confirmed to the company that OB is very rarely used in the third-line setting since patients are likely to have received bendamustine and/or obinutuzumab in either the first- or second-line setting, and patients are infrequently retreated with these regimens. This view is supported by updated IPSOS market share data for August 2021 to July 2022, which demonstrated that OB was used in just [redacted] of patients in the third-line setting ([redacted]). This estimate has decreased from the previous reported analysis for 2021, which estimated OB had a [redacted] market share.

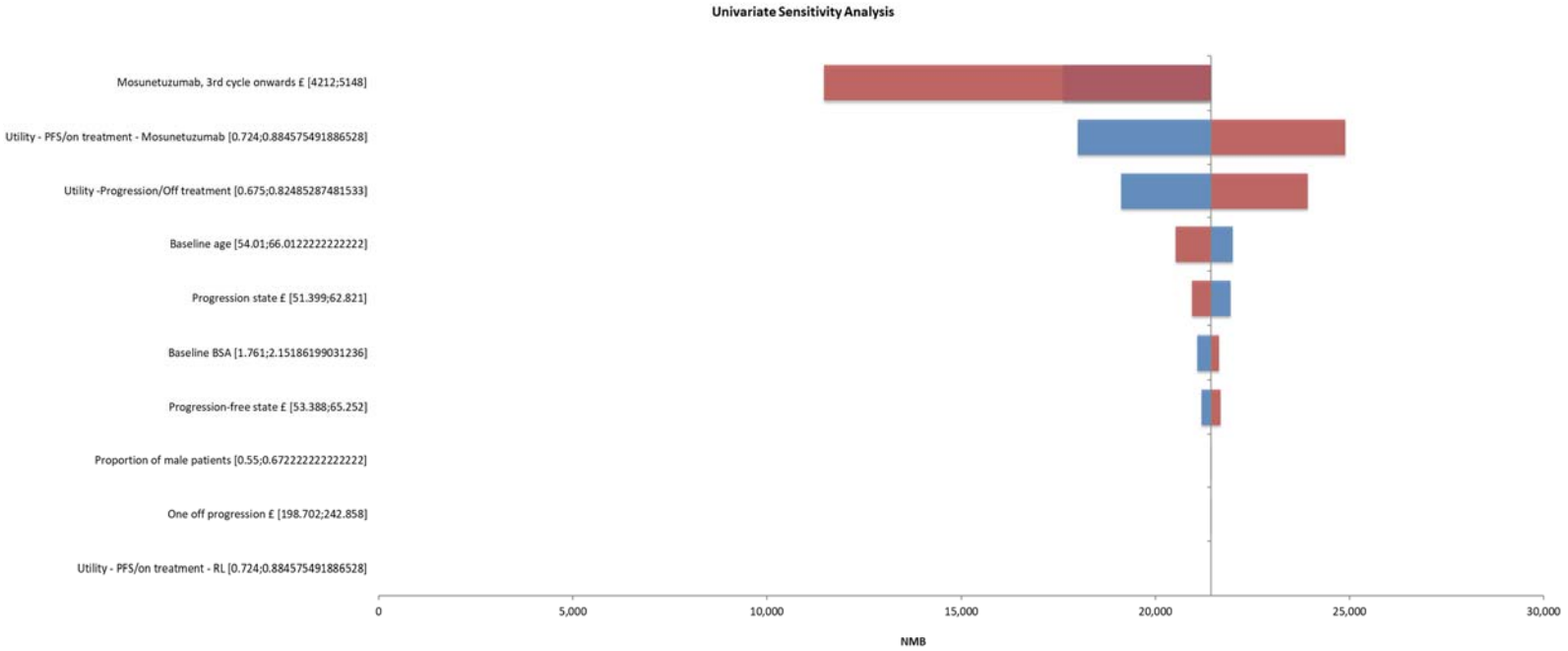
Deterministic results for mosunetuzumab vs OB

Table 7. Deterministic company base case cost-effectiveness results with revised assumptions (PAS discount) – Mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	7.48	[redacted]	[redacted]	0.89	[redacted]	Mosun dominant	[redacted]

Deterministic sensitivity analysis for mosun vs OB

Figure 7. Tornado diagram showing OWSA results on NMB – Mosun vs OB



Probabilistic base case results for mosun vs OB

Table 8. Probabilistic company base case cost-effectiveness results with revised assumptions (PAS discount) – Mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.35	[redacted]					
OB	[redacted]	7.50	[redacted]	[redacted]	0.85	[redacted]	Mosun dominant	[redacted]

Figure 8: Cost-effectiveness acceptability curve - OB comparison

Figure 9: Incremental cost-effectiveness plane - OB comparison

Scenario analyses results for mosun vs OB

Table 9. Scenario analysis – no half cycle correction (PAS discount) – mosun vs OB

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	NMB
Mosunetuzumab vs OB								
Mosunetuzumab	[redacted]	8.36	[redacted]					
OB	[redacted]	7.48	[redacted]	[redacted]	0.89	[redacted]	Mosun dominant	[redacted]

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

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. 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 **Monday 10 October 2022**. 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.

NICE National Institute for Health and Care Excellence

Part 1: Treating relapsed or refractory follicular lymphoma and current treatment options

Table 1 About you, aim of treatment, place and use of mosunetuzumab, sources of evidence and equality

1. Your name	Kim Linton
2. Name of organisation	The Christie NHS Foundation Trust and The University of Manchester
3. Job title or position	Clinical Senior Lecturer and Honorary Consultant in Medical Oncology
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 follicular lymphoma? ☒ A specialist in the clinical evidence base for follicular lymphoma or mosunetuzumab ? ☐ 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)	☐
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Nothing to declare
8. What is the main aim of treatment for relapsed or refractory follicular lymphoma? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	The main aim of treatment is to stop disease progression and associated disease symptoms for as long as possible. This is on the basis that advanced FL is incurable; most people have a long clinical course punctuated by multiple relapses. Using therapy that does not significantly compromise quality of life is

	therefore another important aim. Patients who tolerate treatment well and go into complete remission are most likely to achieve these aims.
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)	A treatment response resulting in low/no disease burden and resolution of symptoms for at least 18 months in the third line or later treatment setting
10. In your view, is there an unmet need for patients and healthcare professionals in relapsed or refractory follicular lymphoma?	Yes, patients have a finite number of licensed treatment options, which they may exhaust during the course of a long illness. Multiply treated patients exposed to multiple rounds of immunochemotherapy – the mainstay of treatment at relapse – may become resistant to chemotherapy or intolerant of its toxicity, leading to inferior outcomes. There is an unmet need to extend the array of treatments to include effective, safe, non-chemotherapy, novel options.
11. How is relapsed or refractory follicular lymphoma 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?	• BSH guidelines on the investigation and management of follicular lymphoma, McNamara et al, BJH 2020 • Relapsed treatment pathways are not well defined for FL. Therapy choices are based on patient age, fitness, previous treatment outcomes and tolerance, available options, clinician/centre experience and patient preference. Licensed options available in the UK for treatment of relapsed and refractory FL (r/r FL) include RCVP, RCHOP, R-bendamustine (BR), rituximab-lenalidomide (R2), obinutuzumab-bendamustine, and platinum-based immunochemotherapy. In current practice, BR and R2 are the most commonly used third line treatments. Rarely used (<10% of patients) are obinutuzumab-bendamustine for patients with rituximab-refractory disease and platinum-based immunochemotherapy in selected transplant eligible patients. • Approval of mosunetuzumab would provide an additional treatment line for patients with r/r FL, offering a non-cross-resistant, safe and effective novel option to prolong PFS and improve QOL.
12. Will mosunetuzumab be used (or is it already used) in the same way as current care in NHS clinical practice?	• Mosunetuzumab is currently delivered within trials at specialised centres with experience of this relatively new class of agents.

• 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)	• Patients receiving mosunetuzumab require monitoring for treatment emergent cytokine release syndrome and neurological events. This differs from current care (other treatments). However, since most TLS/NE events are low grade and occur in the first cycle of treatment, this additional requirement has a small resource impact. • Owing to the favourable safety profile, this drug could be delivered within non-specialist centres. Training will be needed for non-specialist centres.
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?	• Yes, especially where current care is 'best supportive care' for patients who have exhausted standard options or have become refractory to standard immunochemotherapy. Since patients with FL experience multiple relapses, many go on to receive all/most available treatments during the course of their illness. This technology should be viewed as providing an additional line of therapy rather than replacing an existing one. In this context, comparing to current care is an inherently flawed strategy. • As a well-tolerated, non-chemotherapy option delivering high and durable complete remission rates, HR-QOL may be better than current care (mostly immunochemotherapy) both during and after treatment.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	• No, this drug is active across all subgroups, including high risk and patients of all ages.
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	• There are some challenges to delivering mosunetuzumab including monitoring and management of cytokine release syndrome and neurological events. There are however well developed mitigation strategies in place to manage these events. The severity of these toxicities is much lower for mosunetuzumab than CAR-T therapy and as most events are mild and occur early, hospitalisation is not mandated and escalation of care (e.g. admission to ITU) is rare. Nevertheless, from a practical perspective, these services

acceptability or ease of use or additional tests or monitoring needed)	need to be in place and the delivery teams need to receive appropriate training.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	• The same rules as for any other treatment apply to starting/stopping this treatment. No additional testing is necessary to select patients to start treatment. Patients will stop treatment if they progress, develop unacceptable toxicity or reach the end of the planned course. • Patients who achieve complete remission (60% of the enrolled trial population) can also stop early (at 8 cycles)
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	• Mosunetuzumab can be delivered subcutaneously as well as intravenously. SC delivery offers advantages over standard iv options (shorter treatment time, more convenient and acceptable to patients). • The QALY calculation may not capture the cumulative burden of toxicity (physical and psychological) experienced by heavily pre-treated patients exposed to multiple rounds of immunochemotherapy. Re-treatment with standard therapy will have a greater negative impact on QOL than patients who have had fewer lines of 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?	• This technology is the first in class going for approval. As such it is both highly innovative and a step-change in the management of relapsed FL • There is an unmet need to extend the array of treatments to include effective, safe, non-chemotherapy, novel options for patient with FL that have become refractory to standard therapy. Mosunetuzumab is active in refractory disease (70-80% of the enrolled trial population had refractory disease) and provides an additional treatment line for patients with r/r FL, offering a non-cross-resistant, safe and effective novel option to prolong PFS and improve QOL. • There is also an unmet need to improve outcomes for patients who experience early disease progression (including so-called POD24 patients). These patients have inferior survival not overcome by standard therapies. Mosunetuzumab is just as effective in POD24, which made up over 50% of

	the enrolled trial population, than non-POD24. This represents a significant advance in care.
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?	• Most side effects are mild and treatment related deaths are very rare (~2% in the trial). Very few patients stopped treatment due to toxicity (~4% in the trial). High grade toxicities (neutropenia mainly) are easy to manage and associated neutropenic sepsis is exceptionally rare (unlike standard immunochemotherapy where rates of neutropenic sepsis are 15-25%). • The toxicity profile of mosunetuzumab does not significantly impact HR-QOL and in my clinical experience, delivers improved QOL compared to immunochemotherapy.
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 trial broadly reflects UK clinical practice. Compared to real world patients, data from an ongoing German MAIC (Marschner et al, HemaSphere | 2022; 6:S3) confirms that the trial population has a higher proportion of patients with factors known to be associated with poorer outcomes in 3L+ FL, e.g. refractory to prior anti-CD20 therapy, double- refractory to prior anti-CD20 therapy and alkylator therapy. • ORR, CR, PFS and safety are the most important outcomes, and all were measured in the trial. • Surrogates were not used. • No new adverse events have emerged since the trial.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	• No
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance for rituximab in combination with chemotherapy [TA137]; lenalidomide with rituximab [TA627]; obinutuzumab with bendamustine followed by obinutuzumab maintenance [TA629]?	An ongoing MAIC in Germany aims to compare real world data from the Tumor Registry Lymphatic Neoplasms extension with data from the mosunetuzumab single arm trial (Marschner et al, HemaSphere | 2022; 6:S3, presented at the EHA 2022 Hybrid Congress).

23. How do data on real-world experience compare with the trial data?	• There is no real world experience of mosunetuzumab (all treated in trials)
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. 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.	• There are no equalities issues

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Concerns over the suitability of the indirect comparisons performed and presented (section 3.4.1, 3.4.2 of EAR)	• R2 and BR are valid comparators. OB is seldom used in the UK and could be excluded in my view. • The most important clinically relevant variables were included. Missing some high priority variables in some of the comparisons, e.g. refractoriness to anti-CD20, prior ASCT, bulk disease, are unlikely to lead to significant imbalance or bias. • Number of lines of therapy 3 vs > 3 or median is an important prognostic variable that was not included in the MAIC vs R2. Most patients receiving R2 in AUGMENT were second line; this is the main reason for longer PFS for R2 that any 3L treatment
Key issue 2: Inconsistent application of adjusted and unadjusted survival data in economic analyses (section 4.2.6.2, 4.2.6.3)	• Unable to comment

Key issue 3: Unsupported degree of modelled benefit of mosunetuzumab over its comparators (section 4.2.6)	• The degree of modelled benefit is clinically plausible. To date, patients there have been few relapses for patients who have achieved CR, with the median DoCR not reached. This is reflected in a plateau emerging in the survival curves for CR patients. The median follow-up duration (23 months) already exceeds the median 3L PFS (13 months) reported in real world studies (Batlevei et al, 2020). Based on these encouraging data, it is plausible that the modelled PFS for R2 could exceed that for R2 and RB, however longer follow-up would increase certainly. • It is difficult to comment on the OS models. As many patients with FL will receive all available therapies during the course of their illness (if fit enough), OS may be similar unless treatment sequence is prognostic. We have no data on the latter.
Key issue 4: Unnecessary half cycle correction applied in the model (section 4.2.2)	• Unable to comment
Key issue 5: Immature data to model post-progression utilities (section 4.2.7.2)	• In the trial 23% of patients received < 8 cycles and 59% received 8 cycles, i.e. 82% of patients had up to 8 cycles of treatment. Most patients achieved best response by the first response scan. In my experience, by this time most patients achieved their optimal QOL due to a combination of treatment response and recovery from any early side effects. Therefore, in my view, PRO data collected up to cycle 8 is likely to capture the benefits of mosunetuzumab in improving HR-QOL. • The post progression utilities assumed by the company are also reasonable (Table 34, page 134). The company used a utility value of 0.75 for PPS (vs 0.804 for PFS state). An HMRN study of 181 FL patients reported 0.83 for disease remission and 0.74 for people on treatment. The UK population reference is 0.857.
Other issue 6: Inclusion of RB as a comparator (section 3.3.3, table 4)	• RB is one of the most common regimens used in the 3L+ treatment setting in the UK. As such, it is a valid comparator. Other treatments given to patients instead of mosunetuzumab in 3L+ include R2, RCVP and RCHOP • There are very limited comparative data between these options to guide therapy choice at relapse. In the front-line setting RB delivers longer PFS than RCHOP and similar CR rates to pooled RCVP/RCHOP data. RB also delivers higher MRD negative rates than RCHOP and RCVP. RB is however more toxic than RCVP and RCHOP. • There are no published data for standard therapy in r/r FL apart from the Van Oers RCHOP data, which are not relevant to today's practice as the population did not receive prior rituximab.

	• Extrapolating from front line evidence, experts generally agree that RB is the most effective immunochemotherapy treatment at relapse, but also the most toxic. • Due to this evidence gap, 3L+ treatment is very heterogenous and all the population data publications in r/r FL (Casulo et al, Scholar-5, RECORD-FL) report pooled data for 3L+ treatment. • The choice of therapy for 3L+ is therefore not driven by comparative efficacy evidence, but instead based on the following: o Previous treatment - most clinicians to not re-administer RCHOP or RB due to cumulative toxicity risks o Patient age and fitness - full dose RB is less frequently given to patients aged over 70 due to excess toxicity, and RCHOP is avoided in patients with cardiac co-morbidity and usually reserved for treatment of high grade transformation o Treatment intent – high risk or transplant fit patients are less likely to receive RCVP o Availability of trial options – a novel experimental therapy is more attractive than another round of standard therapy for patients with early relapse, multiple previous lines, refractory to last therapy, cumulative toxicity or poor tolerance of standard therapy • RB as a comparator sets a high bar because a) it is probably more effective and b) prognosis is biased by a higher proportion of younger/fitter patients
Other issue 7: Lack of suitable clinical effectiveness data for the comparison with R-CHOP (section 3.3.4)	• There are no published data for standard therapy in r/r FL apart from the Van Oers RCHOP data, which are not relevant to today's practice as the population did not receive prior rituximab. • This is a recognised knowledge gap in the literature. The R2 TA used RCHOP/RCVP combined data from HMRN for their comparator. • The question arises whether the comparison of mosunetuzumab with rituximab + bendamustine can be generalised to a comparison of mosunetuzumab with R-CHOP, i.e. would similar outcomes be expected? We have absolutely no data to challenge this assumption. There may be differences in outcomes at 1L but experts agree that differences are less evident in the relapsed setting. In the HMRN dataset, there were no differences between RCVP and RCHOP in the relapsed setting, hence these data were pooled in the comparator arm for the NICE R2 TA.
Other issue 8: Generalisability of the patient cohort to the	• The trial population is generalisable to the UK population and broadly reflects the patients we see apart from having more high risk patients than would be expected in secondary care, and ~20% PI3k pre-treated patients (Pi3k agents are not licensed in the UK). These differences are inconsequential.

NHS (section 3.2.1, 3.2.2)	
NICE technical team query: Please provide details of any new studies and further data collection, including, start date, design, comparators and locations, that could support a managed access proposal.	• The Chronos 4 clinical trial in r/r FL has a RCHOP comparator arm. Primary completion is expected in Feb 2023. • The UK NCRI REFACT trial (co-CIs Linton and Bishton) is a randomised phase 2 platform trial in r/r FL comparing novel therapies with standard immunochemotherapy (RB, RCVP, RCHOP, R2). The trial is running in the UK, opening in Q1 2023. The primary endpoint for the first round (n=126) will report in Q3 2025.
Are there any important issues that have been missed in EAR?	• No

NICE National Institute for Health and Care Excellence

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Mosunetuzumab is a first in class agent offering an innovative step change in the management of r/r FL.
- It provides an additional line of therapy in a disease prone to multiple relapses but with finite treatment options.
- Moreover, it is effective in high-risk patients including those with early relapse or refractory to previous therapy, thus addressing key areas of unmet need.
- Its manageable and mostly low-grade toxicity profile makes this agent suitable for treatment of all patients, including those who are older, frailer or less likely to tolerate standard therapy based on previous toxicity experience.
- Mosunetuzumab is suitable for delivery in non-specialist centres that have received appropriate training.

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

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

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. 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 **Monday 10 October 2022**. 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.

NICE National Institute for Health and Care Excellence

Part 1: Treating relapsed or refractory follicular lymphoma and current treatment options

Table 1 About you, aim of treatment, place and use of mosunetuzumab, sources of evidence and equality

1. Your name	Mark Bishton
2. Name of organisation	Nottingham University Hospitals NHS 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 follicular lymphoma? ☒ A specialist in the clinical evidence base for follicular lymphoma or mosunetuzumab ? ☐ 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.	NA
8. What is the main aim of treatment for relapsed or refractory follicular lymphoma? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	Most patients have incurable, advanced stage disease that follows a long relapsing/remitting course (rrFL) with progressively shorter remissions despite multiple lines of therapy. Treatment aim is to keep the disease at bay (PFS) whilst maintaining quality of life.

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 metabolic response.
10. In your view, is there an unmet need for patients and healthcare professionals in relapsed or refractory follicular lymphoma?	Yes. Treatment resistance, early progression, and poor survival occurs in 20-25%, whilst increasing numbers experience cumulative complications or treatment resistance after multiple therapies, and eventually exhaust treatment options.
11. How is relapsed or refractory follicular lymphoma 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?	There are no standard treatment pathways, no randomised trials comparing experimental with current therapies, and limited outcome data from current therapy in rrFL. This has created immense difficulties around therapy choice and approval of novel agents. There are 3 effective therapies RCHOP/CVP, R2 or BR, and rituximab maintenance can be used. These can be sequenced at the physicians discretion and so 2nd line+ therapies vary significantly. Dependant on the age and fitness of the patient, autologous stem cell transplantation, and those patients who did not receive maintenance antibody after first line immune-chemotherapy may receive following second (or third) line therapy. As well tolerated Mosun would be used 3rd line + and may replace on of the options above, or be used 4th line when all completed.
12. Will mosunetuzumab 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)	Mosun will be given in Daycase of level 2+ centres in secondary care. No extra investment required.

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?	Improvements in progression free survival are expected – overall survival is less likely and difficult to predict with current data.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No.
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)	High grade Cytokine release syndrome (CRS) is a rare but recognised specific complication, which clinicians need to be aware of. This is confined to cycles 1 and 2 in the vast majority of cases.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	If in CR after cycle 8 can stop, otherwise up to to cycle 12 if disease responding.
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	No chemotherapy so no anti-emetics and less infection prophylaxis needed (eg PJP).

may be more easily administered (such as an oral tablet or home treatment) than current standard of care	
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	It is the first of a new class of drugs for multiply relapsed FL, an area where there is no current standard of care, and so supplies an unmet need of the patient population.
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?	High grade Cytokine release syndrome (CRS) is a rare but recognised specific complication, which clinicians need to be aware of. This is confined to cycles 1 and 2 in the vast majority of cases. Otherwise the treatment is extremely well tolerated.
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?	Mosun has been evaluated in single arm studies and has shown impressive progression free survival with low toxicity. The results can be extrapolated to the UK setting. No adverse effects that were not apparent in clinical trials have come to light.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	There have been a few further real world publications not considered which specifically consider 3L+ therapies for follicular lymphoma. These papers suggest that R-chemotherapy is received in only around 1/3 of patients, so necessarily the numbers receiving BR or RCHOP are less than this, and I believe are reflective of real world practice in an older, more frail population..

	https://pubmed.ncbi.nlm.nih.gov/35358443/ The Scholar 5 data is seen in two papers: https://onlinelibrary.wiley.com/doi/full/10.1002/hon.26_2880 https://pubmed.ncbi.nlm.nih.gov/35679476/
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance for rituximab in combination with chemotherapy [TA137]; lenalidomide with rituximab [TA627]; obinutuzumab with bendamustine followed by obinutuzumab maintenance [TA629]?	No.
23. How do data on real-world experience compare with the trial data?	I am not aware of any real world data with Mosun to date.
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	I am not aware of any equality issues.

• 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. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Technical engagement questions for clinical experts

We welcome your comments on the key issues below, but you may want to concentrate on issues that are in your field of expertise. If you think an issue that is important to clinicians or patients has been missed in the EAR, please also advise on this in the space provided at the end of this section.

The text boxes will expand as you type. Your responses to the following issues will be considered by the committee and may be summarised and presented in slides at the committee meeting.

For information: the professional organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR. These will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Concerns over the suitability of the indirect comparisons performed and presented (section 3.4.1, 3.4.2 of EAR)	This is a technical/statistical query I can't comment on. However, comparing separate studies directly is always fraught and there will be significant caveats in the conclusions derived. There has been very extensive analysis and comparison of sub-groups and I believe this sort of data is open to significant criticism – it would never be accepted in a peer-reviewed journal. It is also clear that many of the patients in the Mosun study will have received ALL or MOST of the regimens in the separate comparator studies already ie RCHOP, BR, R2 etc. You are comparing patients who have received superior 1st and 2nd line therapies in the Mosun study with comparator studies where they have not.
Key issue 2: Inconsistent application of adjusted and unadjusted survival data in economic analyses	This is a technical/statistical query I can't comment on.

(section 4.2.6.2, 4.2.6.3)	
Key issue 3: Unsupported degree of modelled benefit of mosunetuzumab over its comparators (section 4.2.6)	This is a technical/statistical query I can't comment on, as I do not know how the modelling is undertaken.
Key issue 4: Unnecessary half cycle correction applied in the model (section 4.2.2)	This is a technical/statistical query I can't comment on.
Key issue 5: Immature data to model post-progression utilities (section 4.2.7.2)	Follow up is short, however, 80% of patients had QoL responses at 24 weeks. QoL data is very demanding for patients and do not expect >95% uptake. The QoL studies will include patient experience whilst on drug for the majority. If patients have responded and stopped the drug then I expect god QoL scores.
Other issue 6: Inclusion of RB as a comparator (section 3.3.3, table 4)	There is no standard of care for any line of therapy for FL. Please note comments for key issue 1. Moreover, there is now a huge reluctance to use bendamustine due to the well established adverse event of severe lymphopenia and concerns re COVID19. Considering the likely average age of patients treated 3rd line off study is likely well in excess of 65 years, there would be concerns in the community about using bendamustine 3rd line in heavily treated patients.
Other issue 7: Lack of suitable clinical effectiveness data for the comparison with R-CHOP (section 3.3.4)	As for key issue 1, but in addition, on several occasions, the EAG refer to RCHOP being one of the most commonly used treatments for 3rd line follicular lymphoma in the UK. The reference used is in abstract form only, and comprises a total of six patients treated with RCHOP 3rd line. I do not believe this is a credible reference. I would suggest again this shows an absence of good data in the 3rd line setting.

Other issue 8: Generalisability of the patient cohort to the NHS (section 3.2.1, 3.2.2)	Most of the study was conducted in the USA. Treatments for FL are similar. I am not concerned re generalisability. Very few drugs would be approved if this was the case.
NICE technical team query: Please provide details of any new studies and further data collection, including, start date, design, comparators and locations, that could support a managed access proposal.	NCARS culd be asked directly for relevant SACT data, although only OS and TTNT could be provided.
Are there any important issues that have been missed in EAR?	No.

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Mosun is a highly novel therapy which provides a new treatment option

Mosun is very well tolerated

Comparator studies are very difficult to analyse credibly to the extent in the EAR

Many of the patients in the Mosun study will have received ALL or MOST of the regimens in the separate comparator studies already

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

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

Patient expert statement

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with relapsed or refractory follicular lymphoma or caring for a patient with relapsed or refractory follicular lymphoma. The text boxes will expand as you type.

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).

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 **Wednesday 9 November 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with relapsed or refractory follicular lymphoma

Table 1 About you, relapsed or refractory follicular lymphoma, current treatments and equality

1. Your name	Zoe Drymoussi
2. Are you (please tick all that apply)	☐ A patient with follicular lymphoma? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with follicular lymphoma? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	The Follicular Lymphoma Foundation
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:

6. What is your experience of living with relapsed or refractory follicular lymphoma? **If you are a carer (for someone with relapsed or refractory follicular lymphoma) please share your experience of caring for them**	As a patient organisation, we are in touch with a large patient base, with whom we have regular contact. I personally do not have a lived experience of FL or any of the relevant treatments. However, I am representing the feelings and experience of our patient community. There is a range of patient experiences from very low to very high tumour burden; from those who are on watch and wait to those who are waiting for their 6th line of treatment for example.
7a. What do you think of the current treatments and care available for relapsed or refractory follicular lymphoma on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	The current treatments available for R/R FL patients are very limited, unless they transform to DLBCL, where CART is already approved.
8. If there are disadvantages for patients of current NHS treatments for relapsed or refractory follicular lymphoma (for example, how they are given or taken, side effects of treatment, and any others) please describe these	
9a. If there are advantages of mosunetuzumab 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 mosunetuzumab 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**	The biggest advantage of mosunetuzumab is that it offers a completely new option for R/R FL patients. There is still very little understood in the patient community around bispecific antibodies. However, they have seen the potential life-changing impact of CAR-T, and are very hopeful that bispecifics will be a step up from that.

10. If there are disadvantages of mosunetuzumab over current treatments on the NHS please describe these. For example, are there any risks with mosunetuzumab ? If you are concerned about any potential side effects you have heard about, please describe them and explain why	There is always the worry and anxiety around side effects, and how much of those will be long-lasting. Patient understanding around bispecifics is fairly limited, although they have heard the side effects are potentially more manageable than CAR-T, which would be encouraging.
11. Are there any groups of patients who might benefit more from mosunetuzumab 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	This would be highly significant for those who have run out of options and would be in line for systemic chemo or preparing to join a clinical trial.
12. Are there any potential equality issues that should be taken into account when considering relapsed or refractory follicular lymphoma and mosunetuzumab 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.	There is a concern that accessibility may be an issue – would this treatment potentially be only available in the larger hospital settings in a small number of locations across the country? How could all eligible patients access this treatment?
13. Are there any other issues that you would like the committee to consider?	

NICE National Institute for Health and Care Excellence

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Patients are very excited and hopeful for a new option for R/R FL, in a space where there is very little available.
- Mosunetuzumab has shown some excellent response in FL patients, and so it would be hugely significant should it be approved.
- Patients have limited understanding of bispecifics, as the last few years have had a lot of attention on CAR-T.
- There is anxiety over the expected side-effects, and patients will want reassurance over that.
- There is concern over accessibility, that no matter of location and background, eligible patients would be able to access this treatment

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

Mosunetuzumab for treating relapsed or refractory follicular lymphoma [ID3931]

Technical engagement response form – EAG Response

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 1: Key issues

Key issue	Does this response contain new evidence, data or analyses ?	Response	EAG response
Key issue 1: Concerns over the suitability of the indirect comparisons performed and presented (section 3.4.1, 3.4.2)	Yes	The EAG report has highlighted several issues related to the suitability of the ITCs conducted and their interpretation. These are individually addressed below. Please see Appendix 1 for supporting additional analyses. **Residual imbalances for the MAIC vs R^2** On page 65 of the EAG report, it is stated that "*the uncertainty surrounding the other unmatched variables means it is not possible to conclude which way the analysis may be biased*". To mitigate this concern, the company has provided an updated summary table (**Error! Reference source not found.**) for each of the three MAIC scenarios including all priority baseline characteristics reported, before and after weighting. From this table it is apparent that an important residual bias against mosunetuzumab remains in each scenario for all the factors that were not included in the adjustment (highlighted in red).	Please see the text at the end of this document for the EAG response to Key Issue 1.

		The Company would also like to provide further clarifications with respect to the other two factors listed that were not considered for adjustment: ***Size of the largest lymph node lesion***: There is no need to control for both this covariate and bulky disease at the same time, as described in sections 3.1.2.1.1 and 3.2.2.1.1 of the submitted ITC report, as the two are inherently correlated given that bulky disease is derived from the size of largest lymph node lesion (longest dimension). ***Prior ASCT***: This factor was not reported in any of the source documents that were searched for baseline characteristics. However, as the majority of patients enrolled in the AUGMENT FL subgroup had only received 1 prior line of therapy and SCT is a treatment option considered for second or subsequent lines, the Company believes that having residual bias in this factor favouring mosunetuzumab is highly unlikely. **Misinterpretation of approach taken for final propensity score model selection** The company would like to clarify the approach taken for the selection of covariates included in the final propensity score models for the ITCs vs RB and OB as this seems to have been misinterpreted. This approach was based on the following two complementary strategies (as discussed in section 3.2.5 of the ITC report): • Testing of selected interaction terms deemed to be clinically relevant by internal medical advisors. • Testing of all possible interaction terms in an iterative backward-testing procedure (as there may have been further covariate interactions potentially relevant for our datasets in addition to those flagged by internal medical advisors that could have resulted in improved balance).	

		Both of these strategies were tried and the latter always (and somehow unexpectedly) resulted in model overfitting (i.e. several propensity scores being equal to 0 or 1), which in turn did not result in improved covariate balance compared with not making use of interaction terms (as described in the respective result sections of the ITC report). This does not signal that no covariate interaction should be included in the propensity score model. In fact, ensuring that the propensity score function is sufficiently flexible is an explicit requirement in NICE TSD 17 on the analysis of observational data to inform treatment effect estimates (see also QuEENS checklist), but not in TSD 18. Conversely, when the first approach was tested, the inclusion of selected interaction terms resulted in improved covariate balance for the comparison vs RB, but not vs OB, hence why these were included for the former but not the latter (see ITC report). The respective covariate balance plots were not included in the report, but are provided in this response document, for completeness. From the love plots for the ITC vs OB (**Error! Reference source not found.**), it is apparent that the inclusion of the selected interaction terms resulted in a roughly similar number of imbalanced covariates for full matching (the best matching method among those tested) compared with not including them (6 vs 7), but at the same time full balance could not be achieved anymore for the better balancing method (IPTW) used for the base case. The difference for the ITC vs RB is more nuanced, though still present. In fact, for the better balancing method used for the base case (IPTW), the love plot (**Error! Reference source not found.**) indicates that the inclusion of the selected interaction terms resulted in imbalances in one high- and one low-priority prognostic factors compared to two high-priority prognostic factors when interactions were excluded. For optimal pair matching (the best matching method among those tested), it is apparent that the inclusion of the selected interaction terms resulted in a roughly similar number of imbalanced covariates compared to not including them (7 vs 6), though the inclusion resulted in an overall shorter distance from the 0.1 threshold used, especially when considering high priority factors such as age, refractory to last line, refractory to prior aCD20 containing regimens and POD24.	

		To conclude, the final decision for inclusion/exclusion of interaction terms in the ITCs versus RB and OB was solely based on considerations of improvement in overall covariate balance, which is the main goal of any propensity score analysis, and not on subjectivity/medical recommendation. Their exclusion from the propensity score model for the BR ITC, as suggested by the EAG, would thus result in increased bias and yield suboptimal estimates of the relative treatment effect. **Relevance of double refractoriness (DBLRF) to both aCD20 and alkylating agent containing regimens as a confounder** On page 73 of the EAG report, when discussing the ITC vs RB, it is stated that "*This variable also has the potential to be correlated with the other variables based on refractory status already included in the propensity score calculation, effectively prioritising them in the matching analysis. No justification for its inclusion is provided, and it may be detrimental to the analysis*". The Company is unclear what is meant by "effectively prioritising", given that the model coefficients represent the independent contribution of each of the factors included. Nevertheless, the company would like to clarify its meaning behind identifying double refractoriness as an "unclear priority". This was because the feedback received from medical experts (internal and external) suggested that the evidence to assess the value of double refractoriness as an independent prognostic factor was considered to be limited and mainly based on small published subgroup analyses that did not always result in concordant results. However, it was decided to include this factor in the ITCs vs RB and OB as this could represent a confounder when comparing versus these regimens. In fact, it is clinically plausible for a patient who was refractory to a prior alkylating agent (such as bendamustine) to be less likely to achieve a response (and if so, of the same durability) to another alkylating agent based regimen. To demonstrate that, the Company has generated survival plots and multivariate Cox models of OS and PFS comparing all 90 FL patients on mosunetuzumab from GO29781 and all 3L+ FL ECOG 0-1 patients on RB and OB that were used in the ITCs (to represent a pooled bendamustine containing regimen arm). Three multivariate	

		models were fitted, one controlling for all covariates that were feasible without DBLRF, one with DBLRF and the other with DBLRF both as an independent factor and interacting it with the treatment arm. This was done to assess its independent impact as a confounder or effect modifier. Plots and regressions are all included in Appendix 1. From **Error! Reference source not found.** and **Error! Reference source not found.**, it is clearly evident that double refractoriness can have an important impact on survival outcomes for both mosunetuzumab and the pooled bendamustine containing regimen arm. While the multivariate Cox regressions (**Error! Reference source not found.** to **Error! Reference source not found.**) suggest that the impact of this covariate as an independent prognostic factor may be minor, at the same time it is evident that this is likely to be a strong effect modifier for PFS (a model for OS could not be fit as no deaths occurred in the subset of mosunetuzumab patients that were not double refractory). Therefore, these findings indicate that double refractoriness is an important confounder in an ITC vs RB and OB, confirming the hypothesis put forward by the EAG on page 46 of its report when commenting on the presented GO29781 subgroup analyses ("'*refractoriness to various prior therapies might be an important effect modifier*"). Its exclusion from the propensity score model is thus likely to result in residual confounding and biased estimates of relative treatment effects. **<u>"Inflation" of ESS in the number at risk tables for the KM plots</u>** On pages 73 and 79 of the EAG report, the EAG stated that the Company somehow "inflated the ESS" in the numbers at risk tables below the KM plots, compared to the ESS reported in the main body of the CS/ITC report. The Company would like to confirm no ESS inflation has taken place. This misunderstanding may have arisen around how weights used in KM estimators are subsequently translated into number at risk tables for KM plots. The default output of this process is to generate numbers at risk at time zero which correspond to the sum of the estimated weights, and not to the ESS. To ensure the numbers at risk at time zero	

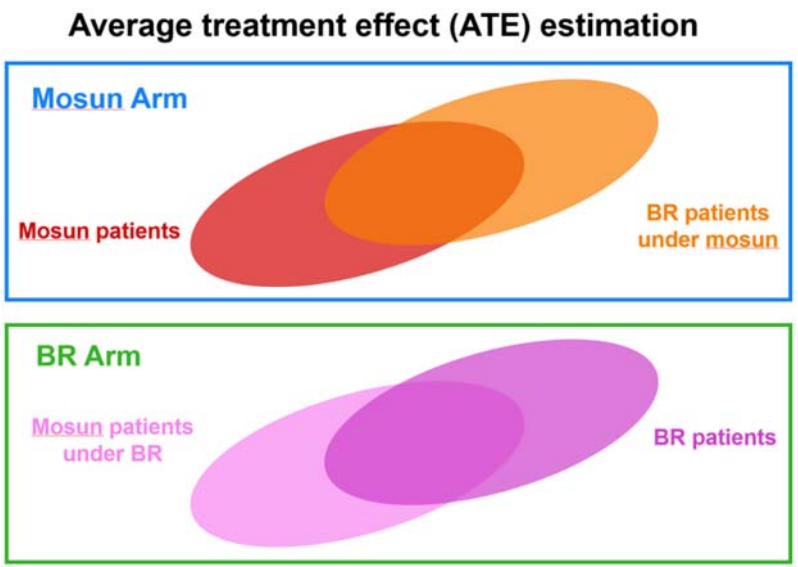
		correspond to the ESS, the estimated weights are rescaled by a common factor equal to the ESS divided by the sum of the weights. Unlike the case of MAICs, such a rescaling was not performed for IPTW and full matching because this procedure may be implemented in different ways when estimating the ATE (target estimand for the Company's propensity score analyses, as per NICE TSD 17), i.e. by arm or pooled, and the Company was initially unsure of what would be the most appropriate approach. In fact, depending on the specific implementation, the sum of these rescaled weights may or may not correspond to the ESS, as some squared terms are involved in estimation of the ESS. It is important to reiterate that weight rescaling has virtually no impact on KM curves or HR/OR point estimates, and therefore not conducting it is not a limitation to the analysis. To help clarify this point, the company has re-run all ITCs by performing such a rescaling by arm. The resulting survival plots and tables with HRs/ORs have been included in Appendix 1 (**Error! Reference source not found.** to **Error! Reference source not found.** and **Error! Reference source not found.** to **Error! Reference source not found.**). These analyses demonstrate that the numbers at risk at time zero correspond to the reported ESS and that the impact on the results is negligible and in most cases not even noticeable. **<u>Presence of individuals with potentially outlier weights</u>** On page 73 of the EAG report, it is stated that *"[..] an individual in the matching analysis initially being given a weight of 26 which was later reduced to 10 to reduce the individuals influence on the analysis, however it is possible all matching analyses encountered similar problems*." The Company would like to clarify that no individual in the matching analysis vs RB was given any weight, as weights are not involved in optimal pair matching. The Company would also like to clarify that weight truncation was only carried out for the IPTW analysis vs RB, as reported. No other cases of outlier weights or propensity scores (as described in the ITC report) were identified.	

		Outcome model appropriateness in matching and IPTW analyses The EAG noted that the full output of the outcome models incorporating covariates to control for residual imbalances post-matching/weighting in the ITCs vs RB and OB was not provided. As a result of that, the EAG could not be certain that the final model used to estimate the treatment benefit is appropriate and sensible for decision making. Further information regarding the approach taken to deal with cases where outcome model estimation was problematic (e.g. due to some coefficient SEs being very large), i.e. by excluding problematic covariates from the model, is described in section 3.2.5.2 of the ITC report, while information of the specific instances where such problematic cases occurred (and for which covariate(s)), was also provided in the summary of propensity score analysis results table footnotes (e.g. see sections 4.2.2.1 and 4.2.2.5) of the respective comparisons. To help clarify, the Company has generated tables containing the full output of all such regression models, which are available in Appendix 1 (**Error! Reference source not found.** to **Error! Reference source not found.**). These analyses demonstrate that, apart from the problematic cases already described in the ITC report (refractoriness to any prior anti-CD20 mAb regimens being excluded from the OS model for the optimal pair matching plus covariate adjustment analysis, and Ann Arbor Stage III/IV and time since completion of last therapy being excluded as covariates from the discontinuation due to AE model for the optimal pair matching plus covariate adjustment analysis), the size of coefficients and standard errors looks sensible and shows no signs of strange or worrisome behaviour of the regression models. For completeness, please note that the coefficients for variables other than treatment arm are only provided to ensure that the treatment effect estimate is unaffected by residual bias, and as such the interpretation of their size and sign is not informative. Furthermore, on page 75 of the EAG report, the EAG "*recommends estimation of effect sizes excluding all covariates from the final model and relying on the populations as balanced by the propensity score matching/weighting*". The Company wishes to clarify that the approach taken to the adjustment for residual imbalances in outcome models fully aligns with published best practices in observational research methods (see e.g. Ali	

		et al 2019 for a recent example, also referenced in the NICE RWE framework), and is also recommended in NICE TSD 17. The Company therefore cautions against removing imbalanced covariates from the outcome models and estimating effect sizes purely based on the populations as (partially) balanced by matching/weighting, particularly as this may yield biased relative treatment effect estimates and it is the reason why such residual imbalances are controlled for in the very first place. **Disagreement between "expected" results from unadjusted and adjusted KM curves and HRs** In its critique of the ITC vs RB, the EAG raised concerns around an alleged disagreement between adjusted KM curves and HRs when compared to unadjusted ones. The Company would like to clarify that in presence of residual imbalances after matching/weighting, including these in a Cox model only controls for them at the HR level, whereas the survival curves remain affected by residual bias, as this cannot be resolved by this second adjustment. This is also mentioned in the discussion section of the ITC report. Therefore, it is not surprising in such situations to see that the difference between adjusted and unadjusted HRs is not of the same size as that between (partially) adjusted and unadjusted KM curves. This is also the reason why the Company/ITC report used IPTW for both comparison base cases rather than matching (lower overall residual balance in both effect estimates and survival curves). Moreover, this also prompted the Company to submit an additional analysis using a method inspired by that described in Austin et al 2020 (further elaboration on this below). **Handling of missing covariate values in ITCs** The EAG report states in several sections that "*For variables with missing data, the ITC report does not state how these observations were handled but refers to R code that has not been made available to the EAG*." The Company would like to clarify that Sections 3.1.2.2 and 3.2.2.2. of the ITC report do explain the methodology employed to	

		handle missing data in MAICs and PSAs, respectively (simple imputations using means and modes, or renormalisations). As agreed during the Technical Engagement call, the Company has appended to this response form the first lines of the R codes for each of the submitted ITCs (Appendix 1, Codes 1-3), which give full details on how many missing values for the baseline characteristics used in the ITCs were observed in each treatment arm. From these it can be confirmed that the proportion of missing data for the individual factors was generally very low, which is unsurprising given that the data used comes from prospectively conducted clinical studies. **<u>Concerns about the implementation of the method described in Austin et al 2020</u>** Section 4.2.6.2.4 of the EAG report highlights some concerns about the implementation of a pair-wise matching plus regression adjustment approach for the ITC vs RB inspired by that described in Austin et al 2020. The Company would like to clarify that the primary reasons for submitting an analysis based on such an approach were to: • Provide an additional treatment effect estimate that was as closely aligned as possible with the NICE TSD 17 recommendations (see e.g. pg 15, "The treatment effect which is typically of interest in NICE TAs is the ATE"), as pairwise matching does not allow to directly estimate the ATE • Comply with the explicit requirement in NICE TSD 17 to use at least one method to estimate the treatment effect of interest (see also QuEENS checklist) • Avoid removing the effect of the remaining imbalanced covariates (as this could already be achieved for the relative treatment effect estimate via the presented optimal pair matching plus covariate adjustment approach). The Company would also like to clarify that in the Company's implementation of this approach to estimate the local ATE, there are no residual imbalances in the baseline characteristics between the final patient groups being compared across treatment arms. In fact, to estimate the ATE "*one needs to construct/impute the counterfactuals for both*	

		the treated and the control individuals" (see also pg 25 of NICE TSD 17) and then use all patients to estimate the final treatment effect. By combining patients from the matched mosunetuzumab ("arm A") and RB groups ("arm B") to create two new treatment ("arm A" [observed] + "arm B" [predicted], for mosunetuzumab) and comparator ("arm B" [observed] + "arm A" [predicted], for RB) groups to estimate the local ATE, the method ensures by design that balance in patient baseline characteristics is achieved (as the two groups are now composed by exactly the same patients). This in turn shifts the focus of the analysis from maximising balance in baseline characteristics to ensuring the validity of the regressions models used to impute the missing potential outcomes under treatment and control for patients who received RB and mosunetuzumab, respectively, to ensure the predicted event times and ensuing treatment effect estimates are unbiased. This is the main methodological difference between the Company's implementation of this approach and the implementation presented by Austin et al. (apart from obvious adaptations required by the fact that we were dealing with a considerably smaller dataset of patients with an indolent disease, whereas Austin et al used either a simulated dataset [where they could decide how many patients to simulate] or a large testing RW dataset or patients with a very high prevalence condition [heart failure]).	

		 In its critique of the Company's approach, the EAG has highlighted some concerns with respect to how outcome regressions and associated event times were estimated. These are individually addressed below. ***Impossibility to assess the starting set of candidate covariates in the models*** The Company wishes to clarify that information on which covariates that had to be removed from the starting set of covariates tested for each outcome model (and the reasons for exclusion) was provided in both the addendum document and the updated ITC report. ***Using different regression models by arm which make use of different covariates*** The Company would like to clarify that using regression adjustment to estimate the ATE is standard procedure and is indeed considered an advantage of the method itself (see e.g. pg 37 of NICE TSD 17, "*RA fits different regressions to the treated and control groups and hence offers additional flexibility*"). Forcing the two regression models to have the exact	

		same specification would instead require the additional assumption that the relationship between covariates and outcome is similar between treated and control units (which may not necessarily hold true in practice for cases where sample sizes and/or number of events are limited). Therefore, there is no requirement for such an estimator of having to employ two models with the same specification. Furthermore, the Company notes that as the method presented combines a design-based method of bias reduction (matching) with an outcome model-based method of bias reduction (regression adjustment), the method can be considered to be doubly robust. This means that "only one of the two models needs to be specified correctly to be able to identify properly the treatment effect" (see also pg 22 of NICE TSD 17). Nevertheless, to further empirically show this in the specific context of the analysis presented, the Company has rerun the analyses using different model specifications than those initially presented (second, third and fourth lowest AIC models among the tested models without convergence issues, vs lowest AIC model in the initial submission). Results from these models have been included in Appendix 1 (**Error! Reference source not found.** to **Error! Reference source not found.**, **Error! Reference source not found.** to **Error! Reference source not found.**) and these confirm that the conclusions of the analyses are essentially unaffected by the different outcome models specification used. ***Censoring of event times greater than the maximum observed follow-up*** The EAG notes that the Company's approach of censoring predicted event times greater than the maximum observed follow-up "*introduces bias into the analysis*" and that "*if event times were excessively large, it could be an indication that the model adjustment is inappropriate and not fit for purpose*". The Company wishes to clarify that using regression models to extrapolate too far off the range of observed data may result in uncertainty and lead to unreliable predicted model	

		outcomes. This is why the Company used the censoring approach for its implementation of the method. ***Assumption that imputed event times follow a Weibull (and not other) distribution*** The EAG was concerned that the Company inconsistently assumed the survival times follow a Weibull distribution for the missing potential outcome imputation, whereas this distribution was deemed unsuitable to fit well the data for the extrapolations used in the CE model. The Company agrees that the choice of the parametric model may have an impact on the predicted event times, although the same distributional assumption is used here to estimate missing potential outcomes under each treatment. Thus, the overall bias is expected to partially cancel out and its impact on the final treatment effect estimate is likely to be negligible. In order to mitigate this concern, the Company has rerun the analyses using alternative parametric distributions (exponential, lognormal and loglogistic) (gamma, gengamma and Gompertz are not implemented in the survreg function of the SurvParamSim package used) and has provided the results in the response document (**Error! Reference source not found.** to **Error! Reference source not found.** and **Error! Reference source not found.** to **Error! Reference source not found.**). These confirm that the conclusions of the analyses are essentially unaffected by the use of different parametric distributional assumptions for the event times. To conclude, the combination of observed and predicted outcomes to estimate the local ATE has the advantage of increasing the sample size and thus precision of the estimate. This could be particularly beneficial in a setting like 3L+ FL (indolent and relatively rare condition, where it is difficult to find patients and events take a relatively long time to occur) and with a limited follow-up time, particularly for OS. Furthermore, combining matching and regression adjustment adjusts the IPDs twice, by first matching and then estimating individual outcomes for a complementary set of patients under each given treatment in the pooled population, thereby allowing to estimate 'doubly robust' survival curves, which cannot be done via other methods such as IPTW (see also point above).	

		The two advantages above are not only beneficial for estimating the treatment effect (i.e. HR and 95% CIs), but also to extrapolate OS and PFS for cost-effectiveness modeling. In fact, this combined sample of patients provides a dataset which is enriched with more events compared to the original one and theoretically less biased survival times, and could potentially lead to more robust and less uncertain long-term extrapolations, thereby also reducing the uncertainty in the final ICER estimates. Because of such advantages, and in light of the rationale and supporting evidence provided as well as of the NICE analysis requirements described in TSD 17, the Company is of the opinion that the results from this analysis are both sensible and useful for decision making. Please see Appendix 3 for a scenario analysis implementing this method.	
Key issue 2: Inconsistent application of adjusted and unadjusted survival data in economic analyses (section 4.2.6.2, 4.2.6.3)	No	The economic analyses have used MAIC adjusted data (comparing mosunetuzumab MAIC adjusted populations to the corresponding comparator population using pairwise comparisons) whenever appropriate, i.e. where no individual patient data was available for rituximab plus lenalidomide (R^2). When individual patient data was available for the comparator, the economic analyses used IPTW weighted data for both mosunetuzumab and the comparator (RB and OB). The company notes an inaccuracy in the original company submission which may have caused confusion, which originally presented optimal pair-matched results in Section 2.9 while IPTW results were reported in Appendix E. The Company confirms that its base case scenario for the comparison versus RB and OB for the clinical assessment was and remains based on IPTW, and that subsequent economic analyses were/are accordingly also based on IPTW. The confusion regarding the data used in the economic analysis may also have originated from the plots that were generated as part of the company submission. These plots, generated in an independent R code from the data that populates the model, were mistakenly using unweighted data for the comparators that were compared by IPTW (RB and OB). Weighted data were still used for mosunetuzumab in those cases. This means that there was a misalignment between the plots used to evaluate proportional hazards and extrapolations (unweighted data for OB and RB comparators, i.e. the IPTW	The company presented a significant amount of information in their first two submissions which contained data that was not relevant to the decision problem. This error was queried by the EAG in the first set of clarification questions. The company either failed to notice or report the mistake as no change was made to the CS. During FAC the company acknowledged the errors in their submission. At this late stage, the errors limit the

		comparators in the model) and the data that the model used (MAIC and IPTW weighted data, depending on the comparison). After noticing the error, the company regenerated the plots that affected the comparisons against RB and OB and were provided in response to the factual inaccuracy check of the EAG draft report. The proportional hazard assessment remained stable and the conclusions did not change with the updated data. Furthermore, the AIC and BIC data provided in the company submission uses the correct adjusted numbers as these numbers were taken directly from the model, so the choice of curve based on the ranking of AIC and BIC still remains valid. To conclude, the company can confirm that there is no reason for concern over the use of the adjusted data, as this was related to errors in the presented plots that did not reflect the data used in the model, nor the AIC and BIC statistics reported in the submission.	value of the EAG critique for committee. At the point of TE, the company's approach to modelling in the clinical and cost-effectiveness sections are consistent and use the same IPTW adjusted data. Therefore, the company's original errors (issue 2) have been resolved.
Key issue 3: Unsupported degree of modelled benefit of mosunetuzumab over its comparators (section 4.2.6)	Yes	**<u>Summary of changes to company base case</u>** The following changes have been made to the company model to help address this issue: • **Alternative parametric distributions to model survival.** Following technical engagement and consultations with clinical experts, the company has decided to review the choice of parametric distributions, to better reflect numbers of patients alive at different time points, hence presenting the most clinical plausible base case on patient survival. However, the company acknowledges that this represents a conservative view and given the limitations and potential bias against mosun with the current ITC results may not represent the true benefit of mosunetuzumab. Further data collection to provide longer-term follow up data for mosunetuzumab and more robust comparative data representative of 3L+ regimens received in UK clinical practice are required. Table 1 reports the revised base case survival distributions chosen, while Table 2 provides a summary of the proportion of patients alive in the model for each distribution and treatment arm. Values greyed out represent unrealistic estimates	Please see the text at the end of this document for the EAG response to Key Issue 3.

of proportions alive at 20 years and therefore unsuitable distributions to be considered in the revised analysis.

Table 1: Selected parametric distributions in revised company base case

	PFS	OS
Mosun vs R^2		
Mosun	Weibull	Weibull
R^2	Log normal	Weibull
Mosun vs RB		
Mosun	Log normal	Exponential
RB	Log normal	Exponential

Table 2: Percentage of patients alive at 20 years in model with each distribution

	Exponential	Weibull	Log normal	Generalised gamma	Log logistic	Ga
Mosun vs R^2						
Mosun	[redacted]	[redacted]	****	****	[redacted]	**
R^2	****	[redacted]	****	****	[redacted]	[redacted]
Mosun vs RB						
Mosun	[redacted]	[redacted]	****	****	[redacted]	**
RB	[redacted]	[redacted]	****	**	[redacted]	[redacted]

*greyed out values indicated clinically implausible estimates

Please also note the following changes to the economic model have been applied:

- **Removal of obinutuzumab-bendamustine from revised base case analysis.** Analysis for mosunetuzumab vs OB is included in the appendix, but the company has not included it in the base case since it does not consider OB to be a relevant comparator based on current market share and clinical opinion. Clinical experts confirmed to the company that OB is very rarely used in the third-line setting since patients are likely to have received bendamustine and/or obinutuzumab in either the first- or second-line setting, and patients are infrequently retreated with these regimens. This view is supported by updated IPSOS market share data for

		August 2021 to July 2022, which demonstrated that OB was used in just [redacted] of patients in the third-line setting (n=[redacted]). This estimate has decreased from the previous reported analysis for 2021, which estimated OB had a [redacted] market share during that year. • **Application of costs for subsequent treatments**. In the previous submitted economic model, costs for subsequent treatment were applied at the point of previous treatment discontinuation. Follicular lymphoma is characterised by prolonged periods of remission (as observed in the GO29781 study where the median duration of complete response (CR) has still not been reached after a median of [redacted] months follow up in the [redacted] data cut), and applying subsequent therapy costs at the point of treatment discontinuation means these will be applied much earlier than in clinical practice. As such, the costs of subsequent treatments are now applied at the point of disease progression, which is representative of what would happen in clinical practice. In contrast, the following assumptions in the EAG's preferred base case **have not been applied** in the revised economic model (although these have been explored in scenario analyses in Appendix 3): • **Half cycle correction:** Please refer to the response to Key Issue 4 below for rationale for retaining the use of a half-cycle correction in the revised company base case. • **Pooling of OS and assuming equivalence in PFS between mosunetuzumab and R^2 at [redacted].** The company considers this to be an overly conservative assumption for numerous reasons outlined below: o Although analyses were conducted in line with NICE DSU recommended methods, the underlying populations informing the ITC were not perfectly matched since the comparator trials were not true third-line plus studies, which may **bias towards the comparators** and therefore **underestimate the clinical effectiveness of mosunetuzumab**. The assumption that the	

		populations are homogenous enough to assume equivalence in OS therefore disregards this uncertainty. o Assuming equivalent survival between treatments **disregards the available evidence from GO29781**. For instance, the PFS KM curve from the ITC (Figure 3, [redacted] analysis report) demonstrated that [redacted]. The EAG's preferred base case opted to assume equivalence in PFS after [redacted]; however, the company understands this is a conservative assumption not reflective of the total available evidence from GO29781, which suggests the majority of patients treated with mosunetuzumab will have durable periods of CR and those patients will therefore not rapidly progress. o This is further supported from the **updated analysis from the latest data cut** which [redacted]. Moreover, the most recent [redacted], therefore the company considers the assumption that there would be no separation in the OS curves to be overly conservative. **Assuming equivalence in OS also disregards the potential impact of surrogate endpoints. As a partitioned survival model is used, CR is not reflected at all in the model despite the 60% CR rate observed in a poorly performing population, which according to clinical expert opinion, exceeds what is typically seen in 3L+ patients with current treatments (CR estimated at 30-40%). Studies have demonstrated that PET CR translates to a survival benefit in first-line treated FL patientss. While there are no published data for 3L+ patients, the company conducted an analysis that was submitted at clarification questions which pooled individual patient data ([redacted])**	

		available in-house for RB (CONTRALTO + GO29365 trials), OB (GADOLIN trial), R-CHOP (EORTC trial) and R-monotherapy (GAUSS trial), comparing overall response rate (ORR) and CR vs OS in 3L+ FL patients with ECOG 0–1. [redacted] **Figure 1: KM plot for ORR relationship to OS**	

		Figure 2: KM plot for CR relationship to OS	

		Additional evidence for the clinical benefit of mosunetuzumab is now available from a comparative analysis of clinical outcomes for mosunetuzumab from GO29781 with real world data (RWD) for third-line+ patients from the US Flatiron Health Database (Appendix 4 and supporting report provided). [redacted] o In clinical practice, **mosunetuzumab would offer an additional treatment option**, which could mean that patients treated with mosunetuzumab could be eligible for more lines of treatment. It is	

		reasonable to assume that a survival curve for third-line patients who are eligible for additional active treatments will differ to survival curves for patients who are limited to best supportive care alone. Therefore, the company consider separate parametric distributions to be appropriate in the current analysis. o A clinical expert consulted by the company stated that they expected there to be a difference in progression-free survival profiles given the **different mechanisms of action** between the bispecific antibody mosunetuzumab, which targets both B and T cells, compared to the immunomodulatory agent lenalidomide that affects T cells only. As such, the company feels separate parametric extrapolations are more appropriate for modelling PFS between mosunetuzumab and R^2. **Details on revised economic model** **Parametric survival distribution choice** ***Mosunetuzumab vs R^2 - Progression-free survival*** For the reasons outlined above, the revised economic model does not apply the EAG assumption that PFS of mosunetuzumab and R^2 is equivalent at [redacted] after the curves cross. AIC and BIC statistics were calculated for the six candidate distributions considered when choosing which model to reflect the base case; however, for the revised base case, the curve selection was guided by the clinical plausibility of the degree of long-term progression free survivors. As such, the Weibull model was selected for mosunetuzumab as this reflected the available evidence while providing estimates of patients to be progression-free at 10 ([redacted]) and 20 years ([redacted]) deemed to be appropriate by clinical experts compared to other distributions. The log normal distribution was selected for R^2 as per the EAG base case.	

		Figure 3: Mosunetuzumab vs R^2 comparison – PFS extrapolations ***Mosunetuzumab vs R^2 - Overall survival*** Similar to the discussion regarding PFS, the company considers not accounting for any survival benefit (pooling OS) to be an overly conservative assumption even in light of the limited follow up, as it does not account for the several factors outlined above. To reflect reasonable estimates of patients being alive at 10 and 20 years, the company has selected distributions that offer clinically plausible estimates at these time points. The Weibull distributions have been selected for both mosunetuzumab and R^2 as these provide realistic estimates of patients alive at 20 years ([redacted] respectively)	

		compared to other distributions. The EAG preferred the log normal distribution to model OS estimates that over [redacted] of patients in both treatment arms will be alive after 20 years. Clinical experts considered this to be unrealistic. Please note however that in the absence of data, the Company has taken a conservative approach, although given the limitations and potential bias against mosun with the current ITC results [redacted], this may underestimate the true benefit of mosunetuzumab. **Figure 4: Mosunetuzumab vs R^2 comparison – OS extrapolations**	

		Mosunetuzumab vs RB - Progression-free survival The company has amended its previous submitted base case so the distributions used to model PFS for mosunetuzumab and RB reflect those selected by the EAG – the log normal model has been selected for both treatments. **Figure 5: Mosunetuzumab vs RB comparison – PFS extrapolations**	

		Mosunetuzumab vs RB - Overall survival The company has revised its base case so that the exponential model is applied to both the mosunetuzumab and RB curves, as per the EAGs preference; however the company has chosen not to assume equivalence between these survival curves by pooling the OS data as outlined above. In addition to these reasons, the company does not feel it is appropriate to assume equivalence between OS for mosunetuzumab and RB given the modelled difference in PFS between treatment arms (the EAG model estimates [redacted] of patents to be progression-free at 10 years on mosunetuzumab and RB, respectively). The company considers that a PFS benefit with any treatment would translate to a difference in OS curves, even if this was only reflected at earlier time points and the curves eventually converged later on. Together with the fact that mosunetuzumab offers patients an additional line of therapy, plus the fact that [redacted] the company considers it a reasonable assumption that there would be a separation in the curves throughout the duration of the model. The exponential distribution is deemed appropriate since it is not only the best fit to the KM data based on AIC and BIC statistics but also offers clinically plausible estimates of patients alive after 20 years after treatment with mosunetuzumab and RB ([redacted] respectively). While the Weibull distribution also offers reasonable estimates of patients alive at 20 years ([redacted] respectively), the exponential distribution is associated with less uncertainty compared to the Weibull distribution, given that it is only associated with a scale parameter (no shape parameter) and has a constant failure rate, therefore making it more suitable for estimations with very few events, as is the case with OS. Given the limitations and potential underestimation of the mosunetuzumab treatment effect from the ITC, fitting simpler parametric models with smaller variation to the available data is likely to be preferable. **Figure 6: Mosunetuzumab vs RB comparison – OS extrapolations**	

		Deterministic base case results The deterministic cost-effectiveness results based on the revised company base case is presented in **Error! Reference source not found.** (Appendix 3). For mosunetuzumab vs R^2, incremental costs were [redacted] and incremental QALYs were -[redacted], resulting in a cost saving ICER of £5,484 (NMB -£7,334). For mosunetuzumab vs RB, incremental costs were [redacted] and incremental QALYs were [redacted], resulting in an ICER of £37,821.	
		Probabilistic sensitivity analysis Probabilistic sensitivity analysis (PSA) was performed for 1,000 iterations, randomly drawn from the parametric assumptions (**Error! Reference source not found.**). The results from the probabilistic analysis are in line with those of the deterministic analysis	

		(ICERs of £3,995 (CS) and £35,235 vs R^2 and RB, respectively). Full results are presented in Appendix 3. **Scenario analysis** Scenario analyses have been run to help provide context to the deterministic cost-effectiveness results and aid decision making. Details of which are provided below and the outcome of each summarised in Table 3. ***Half cycle correction*** Removing half-cycle correction for TTOT has a minor impact on the cost-effectiveness results (**Error! Reference source not found.**). Incremental costs are slightly increased resulting in a small increase in the ICERs but not to an extent that changes the overall conclusions. ***Pooling OS*** Applying the EAGs preferred assumption to assume equivalence in OS to the revised company base case has varying effects on the cost-effectiveness estimates (**Error! Reference source not found.**). For the comparison with R^2, pooling OS results in mosunetuzumab being dominant over R^2. In contrast, pooling OS between mosunetuzumab and RB increases the ICER to £203,383. This scenario analysis demonstrates the uncertainty in the cost-effectiveness results when equivalent survival between mosunetuzumab and the comparators is assumed. For the mosunetuzumab vs R^2 comparison, [redacted] [redacted]	

		[redacted] ***Alternative parametric distributions*** The revised base case has applied a conservative approach to ensure modelled OS results in clinically plausible estimates of patients alive at 10 and 20 years in the model. This is despite of the fact that the clinical effectiveness of mosunetuzumab with respect to the comparators is anticipated to be underestimated given the imbalances in the matched ITC populations and underlying differences in the patient populations. However, given that the most recent [redacted] data cut demonstrates that [redacted], the actual benefit that relapsed/refractory patients can achieve from mosunetuzumab treatment is still yet to be realised, and only further data collection will enable this to be concluded. As such, a scenario analysis has been conducted to investigate the effect of applying alternative parametric distributions to the mosunetuzumab OS data that reflect the potential benefit offered by mosunetuzumab that may be realised with further follow up (**Error! Reference source not found.**). Distributions that estimated over 50% of patients alive at 20 years were disregarded in this scenario analysis. For mosunetuzumab vs R^2, applying a log logistic distribution (AIC/BIC ranking=3) to the mosunetuzumab OS data, which estimates [redacted] of patients alive at 20 years, results in increased incremental costs so that mosunetuzumab is no longer cost-saving, but also an increased in incremental QALYs ([redacted]), resulting in an ICER of £2,095. Similarly, applying a log logistic distribution (AIC/BIC ranking=3) to the mosunetuzumab OS data in the RB comparison ([redacted] of patients alive at 20 years), reduces the ICER to £29,490.	

In summary, while conservative base case estimates do not demonstrate cost-effective ICERs for mosunetuzumab vs the comparators, there are scenarios in which mosunetuzumab is demonstrated to be a cost-effective treatment option compared to both R^2 and RB if alternative distributions that estimate an increased OS benefit are selected. The current available data from the GO29781 study indicate that the patients continue to derive benefit from mosunetuzumab, plus comparative data with a RWD data set (Appendix 4) demonstrate [redacted], therefore further data collection will be required to confirm if these alternative scenarios can be supported with more robust data.

Regression adjustment method for matching mosunetuzumab and RB populations

In this scenario, the regression adjustment method previously described in our clarification question response and in the response above to Key Issue 1 has been applied to the comparison of mosunetuzumab and RB in order to estimate the missing potential outcomes and address the imbalance in the matched population and likely bias against mosunetuzumab compared to RB seen in the ITC.

The regression adjustment method involves first fitting a regression model in each treatment group based on a selection of baseline characteristics. These regression models can then be used to predict missing potential outcomes for a given treatment for each individual belonging to the other treatment group. Such unobserved outcomes can then be combined with the actual observed outcomes under each treatment, to obtain an estimate of the relative treatment effect in the overall patient population. This method combines a design-based method of bias reduction (matching) with an outcome model based method of bias reduction that is similar in spirit to 'doubly robust' methods, in that only one of the two models being estimated needs to be correctly specified to be able to properly identify the treatment effect.

Applying this method in the model with the revised company base case results in mosunetuzumab being cost-effective compared to RB, with an ICER of £14,042 (**Error! Reference source not found.**).

Table 3: Summary of scenario analyses

Scenario	ICER	Conclusion

		Half-cycle correction	vs R^2: £2,997 (CS) vs RB: 38,854	Minor increase to revised base case ICERs and no impact on overall conclusions
		Assuming equivalent OS	vs R^2: Mosun dominant vs RB: £203,383	Assuming equivalent OS and disregarding available evidence results in unstable and unreliable estimates.
		Alternative parametric distributions	vs R^2 (log logistic distribution for mosun OS): £2,095 vs RB (log logistic distribution for mosun OS): £29,490	Alternative distributions that reflect [redacted] demonstrate mosunetuzumab to be cost-effective
		Regression adjustment method for matching mosunetuzumab and RB populations	vs RB: £14,042	Implementing a combination of a design-based method of bias reduction with an outcome model based method of bias reduction (i.e. 'doubly robust' method) demonstrates mosunetuzumab to be cost-effective compared to RB

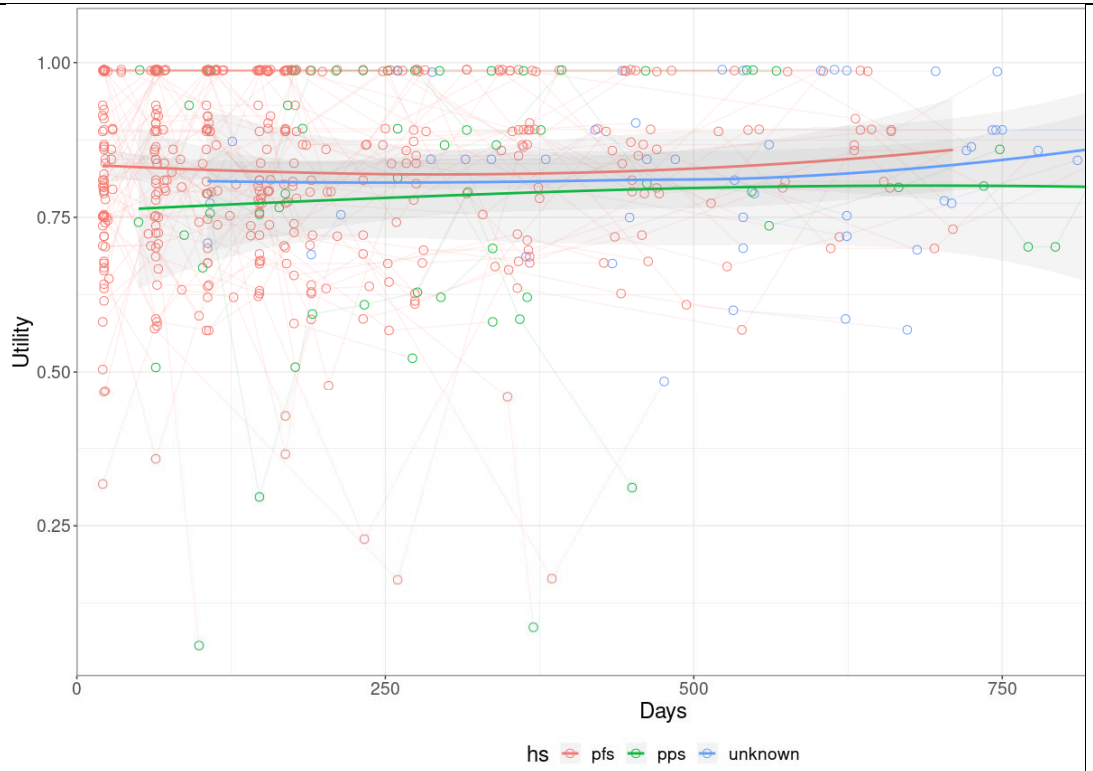
Conclusions

The base case has been revised to provide what the company believes to be conservative estimates of survival, given the available data. Parametric distributions to model OS were selected also based on the clinical plausibility of the degree of long-term survivors in accordance with clinical expert opinion. While this provides a more conservative view of the potential benefit of mosunetuzumab compared to the original company base case, it should be acknowledged that there is still uncertainty in the clinical effectiveness estimates. Given the imbalances in the matched populations and underlying differences in

		the populations used in the ITC, there is a potential for the data to be biased against mosunetuzumab, therefore this should be considered when interpreting the results. In light of this, and the updated data from the [redacted] data of GO29781 that demonstrates that patients [redacted] after [redacted] months follow up, scenarios that reflect the uncertainty in the data and the potential benefit offered by mosunetuzumab are provided (Appendix 3). The company base case has also been amended so that costs of subsequent treatment are applied at the point of progression, rather than treatment discontinuation as this is more reflective of when patients receive their next therapy in clinical practice. For the comparison vs R^2, mosunetuzumab is shown to be cost saving [redacted]. If the EAG's preferred assumption to assume equivalence in OS is applied, mosunetuzumab is dominant over R^2, [redacted], and if the log logistic distribution is applied to model OS, mosunetuzumab will be cost-effective with an ICER of £2,095. Overall, in the most conservative scenario, mosunetuzumab could provide cost savings to the NHS in comparison to R^2, while [redacted] should also be considered in the context of current uncertainty in the data, where the ITC is likely to underestimate the benefit of mosunetuzumab. Mosunetuzumab is marginally above the £30,000 per QALY gained threshold vs RB in the revised base case; however, there are plausible and potentially likely scenarios in which mosunetuzumab is demonstrated to be cost-effective. Again, it is important to consider the uncertainty in the clinical effectiveness estimates from the ITC and the potential benefit patients may actually derive from mosunetuzumab compared to another line of immunochemotherapy, especially since mosunetuzumab is the only treatment available that provides patients with a non-rituximab and non-chemotherapy based regimen. In summary, while the revised base case provides a more conservative view on the degree of survival benefit modelled for mosunetuzumab, the scenario analysis suggests the potential for mosunetuzumab to be cost-effective against R^2 and RB. The latest analysis of the GO29781 data demonstrates that [redacted], thereby supporting the rationale for being possibly an effective and cost effective treatment.	

		The company believes that mosunetuzumab could provide a very important and additional treatment for patients in need of further options, as cited by clinical experts who acknowledge that the complete response rate observed in GO29781 far exceeds what they see with current available treatments. Further data collection in the framework of the Cancer Drugs Fund would provide evidence to facilitate more robust cost-effectiveness estimates of mosunetuzumab.	
Key issue 4: Unnecessary half cycle correction applied in the model (section 4.2.2)	No	Despite the short cycle length, the company considers it appropriate to apply a half cycle correction for TTOT since mosunetuzumab is not administered at a specific time point in the cycle. For instance: • For Cycle 2, study drug infusion should occur on Day 1 of the cycle but may be given up to +/- 1 day from the scheduled date (with a minimum of 6 days after Cycle 1 Day 15 dosing) • For Cycle 3 and beyond study drug infusions should occur on Day 1 of each 21-day cycle but may be given up to +/- 2 days from the scheduled date (with a minimum of 19 days between doses) for logistic/scheduling reasons. • Other study visits starting in Cycle 2 should occur within +/- 2 days from the scheduled date, unless otherwise noted. Given that the model cycles are 7 days in length, half cycle correction is appropriate as +/- 2 days is a sizeable proportion of the model cycle. Furthermore, in TA604, the company did not consider it necessary to incorporate a half cycle correction given the cycle length was one week. However, the ERG deemed this to be necessary to ensure consistent application of total costs and QALY calculations. As such, the revised company base case retains the use of a half-cycle correction to align with previous appraisals, although a scenario in which this is not applied is provided (see Appendix). This scenario demonstrates that removing half-cycle correction for TTOT has a minor impact on the cost-effectiveness results; incremental costs are slightly increased	The EAG accepts the company's rationale for retaining half-cycle correction in the base case and exploring the impact in a scenario analysis. The EAG has also revised its base case and retained half-cycle correction.

		resulting in a small increase in the ICERs but not to an extent that changes the overall conclusions. Therefore, inclusion of the half-cycle correction is inconsequential to the cost-effectiveness estimates.	
Key issue 5: Immature data to model post-progression utilities (section 4.2.7.2)	No	All data up to the most recent follow-up was used in order to estimate the utilities in the model. Utilities included observations up to 2.5 years follow-up using the [redacted] data cut. 63 observations identified in the post-progression health state were used in the regression. Of these, 19 observations were recorded later than one year after the starting period. Observations that could not be identified due to censoring were treated as a different group. The figure below displays the utility for each patient in each of the health states identified (progression-free, post-progression, unknown). **Figure 7: Post-baseline utilities per health state through time fitted with LOESS**	The EAG's critique was made using the information supplied by the company in the CS and during clarification. This information gave no clear indication that data beyond cycle 8 was analysed. In their original submission the company had noted that *"PRO data was analysed up to Cycle 8 of treatment, due to the low number of evaluable patients at cycles beyond Cycle 8 (typically <25% of the patient population available at baseline)."* The EAG acknowledges that the additional information provided by the company during FAC and post technical engagement has clarified that utilities used to inform PPS were collected beyond treatment completion/discontinuation.

		 To conclude, the company believes there is no reason for concern as the number of observations in the progressed health state is enough to be estimated with precision and the values are stable with respect to time.	We note the following point in the company's TE response, *"All data up to the most recent follow-up was used in order to estimate the utilities in the model. Utilities included observations up to 2.5 years follow-up using the [redacted] data cut*" However, some uncertainty remains. We note that whilst the company states that *"63 observations identified in the post-progression health state were used in the regression*" it is very likely that these 63 observations came from a smaller number of patients. If that is the case, some uncertainty regarding PPS utilities still remains.
NICE technical team query: Please provide details of any new studies and further data collection,	No	Further analyses providing more robust, long-term data for mosunetuzumab from the GO29781 pivotal cohort is planned, with annual outcomes analyses conducted until at least [redacted]. In the context of the available evidence to date, an additional 2–3 years follow up of the GO29781 pivotal cohort would provide valuable data to inform the long term extrapolation which may help to resolve the current uncertainties in the degree of benefit derived from mosunetuzumab.	In the absence of head-to-head trials, longer follow-up and RWE would provide improved estimates of mosunetuzumab benefit and resolve some of the

<table>
<tr>
<td>including, start date, design, comparators and locations, that could support a managed access proposal.</td>
<td></td>
<td>In addition, [redacted]

Table 4: Summary of [redacted]
<table>
<tr><td>Study design</td><td>[redacted]</td></tr>
<tr><td>Population</td><td>[redacted]</td></tr>
<tr><td>Intervention(s)</td><td>[redacted]</td></tr>
<tr><td>Comparator(s)</td><td>[redacted]</td></tr>
<tr><td>Outcomes</td><td>[redacted]</td></tr>
<tr><td>Indicate if study used in the NICE economic model</td><td>[redacted]</td></tr>
<tr><td>Trial start date</td><td>[redacted]</td></tr>
<tr><td>Data cut submitted to NICE</td><td>[redacted]</td></tr>
<tr><td>Anticipated data cut after a period of managed access</td><td>[redacted]</td></tr>
</table>

Furthermore, the company plans to generate more robust comparator data for the control arms of the ITC through sponsored projects and supporting investigator-initiated analyses of real world data. In addition, data collection through SACT could help address EAG issues 7 and 8 below, by facilitating a retrospective cohort of patients receiving individual chemotherapy regimens. As highlighted by the comparative analysis in the Flatiron data set (Appendix 4),</td>
<td>uncertainty in the current submission.</td>
</tr>
</table>

		[redacted] therefore a UK-specific analysis would help demonstrate the clinical benefit of mosunetuzumab in a population representative of UK patients. The company believes that the proposed data collection methods are consistent with the evidence package submitted and appraised in recent NICE CDF reviews in haematological indications. For instance, daratumumab monotherapy was recommended as an option for multiple myeloma (TA783) after CDF review on the basis of further follow up of the single arm MMY2002 study, with comparator data sourced from the SACT dataset since this represented patients in UK clinical practice. Furthermore, venetoclax was recently recommended for the treatment of CLL following its CDF review in TA796, with this reappraisal conducted on the basis of SACT collected data. In summary, the company is committed to collecting further data for mosunetuzumab and the relevant comparators, thereby reducing the uncertainty, and it considers the proposed approach to be acceptable for future decision making based on the precedence of recent CDF reviews.	

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 2: 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
Other issue 6: Inclusion of RB as a comparator	Section 3.3.3, Table 4	No	As highlighted in the decision problem meeting, the only rituximab plus chemotherapy option that could be assessed in the ITCs was rituximab plus bendamustine. This is because the SLR could not identify any relevant studies for R-CVP, while a propensity score analysis against R-CHOP based on the EORTC 20981 trial was not feasible due to several limitations associated with small sample sizes, unavailability of information on some important prognostic factors, and important residual imbalances after adjusting for differences in the available factors. This issue has also been discussed in previous appraisals. In TA627, clinical experts confirmed that there were no available data for R-CHOP or R-CVP in previously treated	As stated in the EAG Report, we do not consider RB as a representative for the wider R+Chemo group. The EAG clinical advisor notes differences between patients who receive RB and those who receive other R+Chemo regimes. As stated, we recognised that this is a limitation of data availability. No additional comment to be made.

Other issue 7: Lack of suitable clinical effectiveness data for the comparison with R-CHOP	Section 3.3.4	No	follicular lymphoma, while the committee agreed that it was inappropriate to conduct a comparison with data from EORTC 20981 given the limitations in the evidence base and missing data. Moreover, in TA604, the manufacturer was unable to conduct a comparison with individual chemotherapy regimens having not identified trials or registries that could provide these data. The committee acknowledged that this evidence would be difficult to source. In the current appraisal, the Company has been able to conduct a propensity score analysis using individual patient data for R-bendamustine from two company-sponsored trials. This remains the best and only option available to facilitate a comparison between mosunetuzumab and R-chemotherapy. The company acknowledges the limitation of representing a comparison of mosunetuzumab to R-chemotherapy with one regimen that may not necessarily be commonly used in the third-line plus setting. However, clinical experts confirmed to Roche that if a patient was to receive R-bendamustine in the third-line setting, the observed data would reflect what they would expect to see with this regimen in clinical practice. The company is planning to generate further comparator data for the control arms of the ITC through company-sponsored projects and the support of investigator-initiated analyses of real world data. Moreover, retrospective data	As previously stated, the lack of clinical effectiveness data for the comparison with R-CHOP, and the subsequent lack of assessment of cost-effectiveness of mosunetuzumab against R-CHOP, represents a major gap in this assessment. The EAG consider R-CHOP a common treatment combination used in third line setting for R/R FL in the UK. • The company's propensity score analysis using IPD for R-bendamustine must be considered with caution given that the data was obtained from two company-sponsored trials. • The EAG are unable to verify the independence of the clinical advisors' statements that "*if a patient was to receive R-bendamustine in the third-line setting, the observed data would reflect what they would expect to see with this regimen in clinical practice.*"

			collected via SACT during a managed access period may also generate a cohort of patients receiving individual chemotherapy regimens. These robust comparator data will help to address the limitations and imbalances between patient populations that is a feature of the current ITC.	
Other issue 8: Generalisability of the patient cohort to the NHS	Section 3.2.1, 3.2.2	No	Roche has consulted with several UK clinical experts in the treatment of follicular lymphoma to gain feedback on whether the pivotal cohort population is representative of the UK patients that are eligible for treatment of relapsed or refractory follicular lymphoma following at least two prior therapies. There are some differences in the pivotal cohort population compared with the UK treatable population but the conclusion from the UK clinical experts is that the GO29781 pivotal cohort is representative of the patients they would treat in this indication in the context of the type of patients that are recruited for later line clinical trials in follicular lymphoma. The median age in the pivotal cohort population (60 years) is younger than the UK treatable population, where the median age at diagnosis of follicular lymphoma is reported as 65.6 by the Haematological Malignancy Research Network (HMRN; https://hmrn.org/statistics/quickstats). This is representative of the age of patients that are referred for clinical trials. The clinical experts highlighted that the UK median age is within the interquartile range of	The EAG are unable to verify the independence of the clinical experts' feedback on the issue of generalisability of the GO29781 cohort to the UK. The EAG reiterate that a small number (n=2) of participants in the trial cohort of 90 were from the UK. This uncertainty cannot be resolved without additional evidence from a UK population or real-world data.

			the age in the pivotal cohort population (53–67) years, with the range 29-90 years and the subgroup analysis demonstrated that the CR rate is consistent for patients aged <65 and >65. The ethnicity of the pivotal cohort population may not be representative of some areas of the UK. Again this reflects the types of patients that are entered into clinical trials. At a national level, the differential ethnicity categories is representative of the population; in the pivotal cohort population 82%, 9% and 4% were white, Asian and Black or African American, respectively. A between-census publication from the Office of National Statistics in 2019 reported that 85% of the England and Wales population were white, 8% were Asian/Asian British and 3.5% were Black / African / Caribbean / Black British (https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/populationestimatesbyethnicgroupandreligionenglandandwales/2019). Risk prognostic factors such as ECOG performance score, Ann Arbor stage, disease bulk, FLIPI score and refractoriness to last previous therapy in the pivotal cohort baseline characteristics are well representative of the UK treatable population according to the clinical experts. There is a higher proportion of POD24 patients in the pivotal cohort population than is expected in UK clinical practice; the proportion of patients classed as POD24 is approximately 20% of the patients completing first line	

			treatment for follicular lymphoma, however, this proportion tends to increase in the population treated with later lines of therapy. The previous lymphoma therapies in the pivotal cohort population is representative of UK patients treated in the third line and beyond, with all patients exposed to alkylator and anti-CD20 therapy. However, due to the limited access to PI3K inhibitors in the UK, fewer patients would have been treated with these agents.	

EAG Response to Key Issue 1

#1 Residual imbalances for the MAIC vs R^2

The company has presented output comparing the covariate balance after the MAIC analysis has been performed.

When covariates have been included in the matching analysis, they are well matched to the target value from the AUGMENT trial of R^2, with omitted covariates varying from the target value. Most of these imbalances suggest the resulting mosunetuzumab population is unhealthier than the target R^2 population (higher proportion of patients either with ECOG score 1, are refractory to aCD20, have had >1 prior therapies or have a presence of B symptoms). However, it has not been possible to quantify the potential effects of these differences and the company carries the resulting data from the MAIC analysis into the economic analyses without any adjustment.

The EAG's preferred analysis is to exclude the variable "Low Hgb level" as this value is unknown for the target population, and also increases the effective sample size but this was not implemented within the economic model.

The company imputed the value of this variable and used it in the matching analysis for reasons that remain unclear.

The small effective sample size of all the MAIC analyses means the true efficacy of mosunetuzumab is unlikely to be well-represented.

#2 Misinterpretation of approach taken for final propensity score model selection

The company clarifies the approach to the selection of interaction terms in their propensity score analyses. However, it remains unclear how relevant interaction terms were selected by the company's internal medical team and classed as clinically relevant.

The EAG previously noted that the company included interaction terms for one of their propensity score analyses (vs RB) but not the other (vs OB).

The company provides love plots to show the benefit of including the interaction terms in the OB analyses, and the lack of benefit of including them in the RB analysis. The EAG focuses on the IPTW analyses for both comparisons, as these are agreed as being most relevant by the company. For the OB comparison, the inclusion of interaction terms does not improve the analysis, and potentially weakens the balance of measure covariates. However, for the

RB IPTW analysis, the inclusion of interaction terms has a neutral effect on balance, and it cannot be concluded from the evidence provided that their inclusion improves the analysis.

The benefit of the inclusion of interaction terms is only clear in the matching analysis, which is not the preferred analysis of the EAG or the company.

#3 Relevance of double refractoriness (DBLRF) to both aCD20 and alkylating agent containing regimens as a confounder

In both the OB and RB propensity score analyses, the company include the "double refractory" which it classed as a variable of "unclear" priority, whilst other variables were of "high" or "low" priority. The company has provided a series of analyses where the inclusion of double refractory status has been varied. These analyses compare mosunetuzumab patients with a pooled population of OB and RB patients, and do not involve any indirect treatment comparison but control for covariate differences via their inclusion in the Cox model.

The first analysis of PFS and OS excludes double refractory status, the second includes it, and the third includes it along with its interaction with treatment arm.

From these analyses, the company conclude that double refractory status is an important confounder which should be included in the propensity score analysis. Whilst it appears influential on the treatment effect, it is difficult to say whether it should be included for the following reasons.

The magnitude of the standard errors across the analyses provided in Tables 8 to 13 of the company's TE response suggest little can be reliably inferred from them. The p-values show that almost no covariates included in the models are statistically significant. An examination of the company's analyses including the interaction term suggests that double refractory status only has an effect in the mosunetuzumab arm and has a neutral effect in the RB-OB arm.

It is unclear why this differing effect could be considered clinically plausible.

The effect observed in the Kaplan-Meier plots shown in the company's TE response Figure 10 and 11 may instead be demonstrating the effect of a variable correlated with double refractory status.

Furthermore, the company's tables of results do not appear to be consistent, with only the treatment effects reported for the analyses that exclude double refractory status for PFS and

OS (Table 8 and Table 11, TE response) matching what appears in the earlier tables 6 and 7.

It is unclear why the treatment effect size estimate is quite so sensitive to the inclusion of double refractory status or its interaction with treatment effect. This instability is an indicator that the model may not be reliable.

The EAG conclude that these analyses are not supportive of the classification of double refractory status as an important variable and do not justify its inclusion in the propensity score analyses. It is unclear whether its inclusion is sensible.

#4 "Inflation" of ESS in the number at risk tables for the KM plots

The EAG had concerns that the company had not reported how they had applied the weightings for the propensity score analyses. The company has provided an explanation and additional analyses where an alternative weighting approach is applied. A comparison of the two methods shows the point estimates are virtually identical. However, there are a couple of occasions where there is a minor difference in the confidence interval for some the analyses (e.g. the OB OS hazard ratio lower confidence interval has increased from 0.18 to 0.28 in the company's new scenario). Ideally, the company would have provided Kaplan-Meier plots with confidence intervals for each arm to allow a more meaningful comparison, however these concerns are very minor.

#5 Presence of individuals with potentially outlier weights

The EAG's comment related to the RB IPTW analysis, not the matching analysis, and remains a valid limitation of that analysis.

#6 Outcome model appropriateness in matching and IPTW analyses

The EAG's main original concerns on this issue were around the optimal matching chosen by the company as their preferred analysis for the RB comparison.

As the company have switched to the IPTW analysis during technical engagement, this original point is no longer relevant.

However, the company have also provided the coefficients for the covariates that were included in the post-hoc adjustment of the IPTW analysis for RB. The company have

selected ECOG (1 vs 0) and LDH as these covariates were outside the threshold set by the company for the absolute difference in standardised means after IPTW was performed.

Whilst the rationale for the post-hoc adjustment of this IPTW analysis is clear, it is not obvious whether the adjustments improve or detract from the reliability of the analysis, given the wide degree of uncertainty surrounding them (Table 5). ECOG varies between a substantial and a negligible effect size, whilst a higher ECOG score is associated with worse PFS, OS and CR outcomes, but better ORR outcomes. The confidence interval for LDH on the OS HR is extremely large.

Estimates from doubly robust regression models can be biased as reported by NICE TSD 17 when both the regression model and propensity model are misspecified. Also, as demonstrated by Kreif et al. 2013[1], where IPTW alone performs comparably to regression and IPTW when the propensity scores are specified correctly and the regression is misspecified, which may the case in the company's analysis.

The EAG repeats that the standard IPTW (without post-hoc covariate adjustment) analyses would be a valuable comparison for decision making purposes, but these have not been provided by the company.

Table 5: Comparison of coefficient effect sizes in post-hoc adjustments of the RB IPTW analyses.

	PFS HR (95% CI)	OS HR (95% CI)	ORR OR (95% CI)	CR OR (95% CI)
Treatment arm (Mosunetuzumab vs RB)	[redacted]	[redacted]	[redacted]	[redacted]
ECOG (1 vs 0)	[redacted]	[redacted]	[redacted]	[redacted]
High LDH (Yes vs No)	[redacted]	[redacted]	[redacted]	[redacted]

#7 Disagreement between “expected” results from unadjusted and adjusted KM curves and HRs

The EAG understands the reasons why the adjusted and unadjusted estimates differ. However, our critique highlights the difference to improve the transparency in the company’s reporting. This is beneficial to support the committee’s understanding.

#8 Handling of missing covariate values in ITCs

The company provided R code to show how they dealt with missing values across the patient level data that was used in their analyses. Missingness was generally low, ranging from 1-9 patients across the variables and datasets. The most commonly missing variable was bone marrow involvement. Presuming these values were based on populations that had already excluded patients based on trial inclusion criteria, these values represent 1% - 12% of patients in a single treatment group. All imputation used the simple approach of using the most common approach. For the variables with higher missingness, this may introduce slightly influential amounts of bias.

#9 Concerns about the implementation of the method described in Austin et al 2020

In their addendum submission, the company performed a scenario analysis whereby they estimated counterfactual event times in order to obtain an estimate of the average treatment effect. This was motivated by potential bias in the RB comparison that remained after propensity score matching had been performed.

However, given that their preferred method is now IPTW, which produced a much improved population balance, it is unclear whether this analysis is relevant to the appraisal.

One of the EAG’s concerns with the analysis was that the regression models used different covariates to estimate each set of unobserved outcomes.

Whilst such an approach can be appropriate to allow flexibility and for varying effects across arms, it should also be supported by clinical rationale. The EAG has compiled a table of the covariates used in each analysis. Across ORR and CR, the covariate choice is fairly consistent for both arms, as to be expected for correlated outcomes (Table 6). However, for PFS and OS, the covariates used show little overlap or consistency across outcomes or arms.

It is unclear why it is appropriate that the set of covariates for these treatments and outcomes should vary so much.

The approach does not seem to have considered clinical plausibility and has focused on statistical fit.

Table 6: Covariates included in company preferred models.

	Size of Largest Lesion	High LDH	Time since last therapy	Refrac to aCD20	POD24	FLIPI ≥ 3	Age	ECOG	Refrac to last line	Low Hgb	Bone Mar
PFS Mosun	Yes	Yes	Yes								
PFS RB				Yes	Yes						
OS Mosun	Yes					Yes					
OS RB		Yes					Yes	Yes	Yes	Yes	
ORR Mosun	Yes		Yes			Yes					Yes
ORR RB	Yes		Yes			Yes					Yes
CR Mosun			Yes			Yes				Yes	Yes
CR RB			Yes			Yes				Yes	Yes
AE Mosun	Yes										
AE RB	Yes										

In an attempt to address the covariate selection issue, the company implemented the models with the second, third and fourth lowest AIC and presented output in their TE response.

The company only present information for PFS and OS outcomes, omitting ORR and CR. The EAG has compiled this information in Table 7. The estimated hazard ratio for PFS noticeably increases in the third model, although remains statistically significant. It is possible that this estimate or the upper confidence interval is a typo by the company. The OS estimate appears more stable across the four models provided.

Table 7: Comparison of hazard ratios from the models with lowest AIC

	PFS Hazard Ratio (95% CI)	OS Hazard Ratio (95% CI)
Lowest AIC	[redacted]	[redacted]
Second Lowest AIC	[redacted]	[redacted]
Third Lowest AIC	[redacted]	[redacted]
Fourth Lowest AIC	[redacted]	[redacted]

The EAG's previous criticism of the generation and censoring of large event times remains valid. The company imply weaknesses of extrapolating beyond the observed data using these models, however we note that extrapolation of this data is then performed within the cost-effectiveness analysis.

A further limitation raised by the EAG was that the observed and estimated event times were assumed to follow a Weibull distribution.

The company have attempted to address this concern by using alternative parametric survival models. Each of these models changes both its functional form but are not fixed to have the same set of covariates, so it is difficult to attribute any change (or lack of) of effect size solely to the change in parametric model. The EAG present the different models in Table 8. It is not reported how the covariates for each model were chosen, but presumably these each had the lowest AIC within the set of candidate models of their respective parametric form. The company provided only PFS and OS outcomes for their preferred models, which are shown in Table 9, and not ORR or CR.

Table 8: Comparison of covariates included in the analysis using each parametric model.

	Model	Size of Largest Lesion	High LDH	Time since last ther	Refrac to aCD20	POD24	FLIPI ≥ 3	Age	ECOG	Refrac to last line	Low Hgb	Ann Arb	≥ 3 prev ther
PFS Mosun	Weib LNor Llog Exp	Yes Yes	Yes Yes Yes Yes	Yes Yes Yes Yes									
PFS RB	Weib LNor Llog Exp				Yes Yes Yes Yes	Yes							
OS Mosun	Weib LNor Llog Exp	Yes Yes Yes Yes					Yes Yes Yes			 Yes			
OS RB	Weib LNor		Yes Yes			Yes		Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes	

	Llog Exp	Yes	Yes Yes					Yes Yes	Yes	Yes Yes	Yes		Yes

Table 9: Estimates of PFS and OS hazard ratios for the different parametric survival models

	PFS Hazard Ratio (95% CI)	OS Hazard Ratio (95% CI)
Weibull model	[redacted]	[redacted]
Log normal model	[redacted]	[redacted]
Log logistic model	[redacted]	[redacted]
Exponential model	[redacted]	[redacted]

The EAG identified that tables 65-67 of the company TE response are labelled incorrectly as log-logistic analyses, when they presumably use the exponential distribution.

The EAG note additional potential limitations to the company's exploratory analysis.

Firstly, in their addendum submission, where this analysis was originally reported, the company only presented output from the PFS and OS analyses and failed to present output for the ORR and CR outcomes. The EAG compile all of the results in Table 10. Whilst PFS and OS appear to support a stronger effect of mosunetuzumab, the ORR and CR outcomes still show a negative effect of mosunetuzumab, albeit slightly weaker, relative to the original matching and IPTW analyses. The confidence intervals remain wide across the majority of analyses.

Table 10: Hazard Ratios and Odds Ratios for RB comparison, extracted from Company ITC Report

	Optimal Matching and Post-hoc adjustment	IPTW and Post-hoc adjustment	Matching + Regression (Austin)
PFS HR	[redacted]	[redacted]	[redacted]
OS HR	[redacted]	[redacted]	[redacted]
ORR OR	[redacted]	[redacted]	[redacted]
CR OR	[redacted]	[redacted]	[redacted]
AE OR	[redacted]	[redacted]	[redacted]

Secondly, it is unclear why the company used AIC values to compare the candidate models and inform covariate selection rather than likelihood ratio tests, when the models are all of the same parametric form. Neither is it clear what method of matching the company used in this analysis, whilst the Austin et al. study used two versions of nearest neighbour matching.

This exploratory analysis presented by the company may no longer be relevant to this appraisal. If it is deemed relevant, then the limitations outlined above should be considered carefully. As shown in Table 10, the exploratory analyses when considered in their entirety, reinforce the conflicting nature of the company's indirect treatment comparison results, and do little to address the uncertainty in this appraisal.

In summary, the EAG still has a number of concerns surrounding the implementation and reliability of the company's ITCs and supporting analyses, which are carried into the cost-effectiveness assessments.

EAG Response to Key Issue 3

The company have outlined their revised assumptions for their preferred base-case analysis. These include dropping the OB comparison, changing how the costs for subsequent treatment are applied, and some changes to the choice of parametric models used to extrapolate PFS and OS.

Regarding the removal of the OB comparison, the EAG accept it is the least relevant of the three original comparisons, but it is unclear why the company has made this decision at this stage at the appraisal. There does not appear to be any new motivating evidence for this change at technical engagement.

Regarding the application of subsequent treatment costs, previously these were applied at the point of stopping previous therapy, whereas the company have changed this to occur at the point of disease progression. This change will have most effect in the mosunetuzumab population as this is the only group where time on treatment is distinguished from PFS before the treatment regimens finish.

Whilst this makes the modelling of mosunetuzumab more realistic, it draws attention to the bias in favour of mosunetuzumab caused by the assumption of PFS and TTOT for the comparator treatments.

The company's decision to change the time of application of costs for later lines of therapy led the EAG to re-investigate this area. The company assumes that the proportions of patients who move onto different therapy classes (i.e., subsequent treatment) are the same across Mosunetuzumab and comparator groups. For R^2, that means 35% of patients will receive Rituximab-Lenalidomide as subsequent therapy, a figure that is deemed potentially too high by the EAG's clinical expert and unlikely to reflect clinical practice even when one considers that some patients may have initially been successfully treated with R-Len. Moreover, the company model assumes for R^2, that up till 12 months (treatment-specific maximum number of cycles), patients are treated until progression. It is unlikely that patients who progressed whilst on treatment specifically for the initial 12 months would receive R-Len as subsequent therapy.

The EAG revised its base case to reflect this and considers that no patients would receive R-Len as subsequent therapy in the R^2 comparator arm. As no data were available to inform the distributions for the rest of the subsequent treatments, the EAG maintains the company assumption that 10% and 20% of patients receive palliative care and are on Trials respectively. The rest of the patients, either receive R-chemo or other (non-Rituximab) chemo in a ratio of 2.5:1 as indicated in the company submission (Table 11).

Table 11: Proportions of patients receiving subsequent therapy by therapy class (assumed equivalent for all arms)

Therapy Class	Proportions assumed (Company submission)	Proportions assumed (EAG base case for R^2)
R-len	35.00%	0%
R-chemo	25.00%	50.00%
Other (non rituximab) chemo	10.00%	20.00%
Palliative care	10.00%	10.00%
Trials	20.00%	20.00%

The company disagree with some of the EAG's assumptions: 1) Removal of half-cycle correction, 2) Pooling of OS and 3) equivalent PFS in R^2 comparison after [redacted].

The first point is discussed in more detail under Key Issue 4.

For the other two points, the EAG provides a table of the company's preferred estimates of mosunetuzumab effect relative to each of the comparators (Table 12), before responding to the company's comments.

Table 12: Overview of results from preferred indirect treatment comparisons

	PFS HR (95%CI)	OS HR (95%CI)	ORR OR (95%CI)	CR OR (95%CI)
Mosun vs R^2 (MAIC)	[redacted]	[redacted]	[redacted]	[redacted]
Mosun vs RB (IPTW + adjusted)	[redacted]	[redacted]	[redacted]	[redacted]
Mosun vs OB (IPTW)	[redacted]	[redacted]	[redacted]	[redacted]

1. The company state that potential bias in the ITC analyses against mosunetuzumab is ignored and not factored into the analyses if the assumption of equivalence is made. The EAG note that the effect of this potential bias has not been quantified, and is unlikely to be uniform across each comparison.

Hence it is not possible to say whether the combination of models selected by the company are any way proportional to any bias that may be present in the analyses.

2. The company state that the assumptions of equivalence ignores evidence from GO29781, seemingly due to the high CR rate of mosunetuzumab in this trial. The single arm nature of the trial results in reliance on the ITCs to estimate the relative effects of mosunetuzumab.

The EAG struggle to make sense of the conflicting results across the indirect treatment comparisons where the effect of mosunetuzumab varies in direction and magnitude.

[redacted] When considering the entirety of the evidence the company's rebuttal is unsupported.

3. The company provided updated follow-up from GO29781, using a [redacted] data-cut, compared to the [redacted] data-cut used in the previous submission, which the company says supports an OS benefit of mosunetuzumab. **As information from this**

data-cut has not been carried into the indirect comparisons or economic analyses, it is impossible to draw any conclusions over relative effect. As ORR and CR information were not presented for the previous [redacted] data-cut, and Kaplan-Meier plots for PFS and OS of the latest data-cut were also not provided despite these being offered in the Technical Engagement Telephone Call, it is difficult to infer anything from this comparison. For reference, the median duration of response in the AUGMENT trial of R[2] was 36.6 months[2], which exceeds the current follow-up from GO29781 ([redacted] months).

In summary, the study design and sample size mean that this extension of follow-up is insufficient to make any meaningful reduction in the uncertainty.

4. The company state again that the CR benefit is not reflect in the assumptions of equivalence. This is a consequence of the company selecting a partitioned survival model which uses progression-free, post-progression and death health states, and does not directly model response rates. The company repeat results from an analysis originally presented in their clarification responses, showing a difference in OS for patients with any response or complete response, using follow-up from a range of trials (CONTRALTO + GO29365 trials of RB, GADOLIN trial of OB, EORTC trial of R-CHOP and GAUSS trial of rituximab monotherapy).

The analysis does suggest there is a link between response overall survival across the period of follow-up. The ORR vs non- ORR analysis shows a strong OS benefit but is based on a small number of patients in the non-ORR group. The benefit associated with CR is weaker and decreases from about halfway through the follow-up period and almost converge at the end. This suggests that a higher CR rate may not result a long-term OS benefit, and it may be appropriate to assume equal efficacy. This analysis does not include any mosunetuzumab patients, and has varying definitions of CR across the trials included in the analysis. Aside from having access to the data for the trials included in this analysis, there is no justification that pooling their results is a sensible idea, neither does the analysis include any adjustment for baseline differences in these groups that may affect OS independently of treatment response. Results were not provided for an equivalent analysis on PFS.

5. The company present a new analysis which uses data from the August 2021 data-cut of GO29781, and compares it to real-world data for third line+ patients from the Flatiron Health Database of patients in the USA. A summary of results is included in the TE response form, with the main report presented in a separate report.

[redacted]

[redacted]

The generalisability of this study to the present appraisal is also a concern as it is not clear whether the pooled comparator population is representative of UK care, either before or after the weighting was performed.

6. The company state that mosunetuzumab will add an additional option for patients, meaning patients can have an extra line of therapy before receiving best supportive care which should improve OS. Whilst this is plausible, the data are too immature to confirm this, and it may be affected by other causes of mortality in this ageing population, along with the order of treatment sequencing.

[redacted]

7. Finally, the company presents expert comments highlighting differences between lenalidomide and mosunetuzumab, to justify modelling them separately. However, these comments appear to neglect the fact that the comparator is lenalidomide with rituximab and not lenalidomide monotherapy, which would nullify the company's argument.

EAG's justification for pooling OS and PFS outcomes in some of the analyses.

1. The EAG preferred to use pooled estimates to extrapolate OS because the OS data from GO29781 is particularly immature and unlikely to allow any parametric model to reliably predict the long-term efficacy.
2. [redacted] All of the comparator trials had longer follow-up which meant their data would provide a more reliable extrapolation.
3. The EAG's and company's clinical experts agreed that there would be little difference in OS across the treatments considered, hence it seems a reasonable assumption to make.

Regarding the assumption of equivalent PFS after [redacted], the results of the MAIC analysis show

[redacted] The approach to assume beyond the point the curves cross maximises the use of the more mature R² follow-up, and removes the chance of putting too much weight on tail data.

The company have tried several approaches to obtain a reliable estimate of the effect of mosunetuzumab to R², RB and OB. Unfortunately, these analyses continue to give a conflicting account of the effect of mosunetuzumab, as it appears to be simultaneously better and worse depending on the outcome assessed.

The company report changing some of their preferred parametric curves and provide justification for these changes. An overview of model preferences is provided in Table 13.

Table 13: Overview of extrapolation preferences by company and EAG

	Company Preferred Addendum Submission	EAG Report Preferred	Company Preferred TE Response
PFS R² PFS Mosun (vs R²)	Log-normal Log-normal	Log-normal Same as R² from [redacted]	Log-normal Weibull
OS R² OS Mosun (vs R²)	Weibull Log-normal	Log-normal (pooled) Log-normal (pooled)	Weibull Weibull
PFS RB PFS Mosun (vs RB)	Log-logistic Log-normal	Log-normal Log-normal	Log-normal Log-normal
OS RB OS Mosun (vs RB)	Exponential Log-normal	Exponential (pooled) Exponential (pooled)	Exponential Exponential
PFS OB PFS Mosun (vs OB)	HR applied to Mosun Log-normal	Exponential Log-normal	-
OS OB OS Mosun (vs OB)	HR applied to Mosun Log-normal	Log-normal Exponential	-

R² Comparison

For the comparison to R², the company have changed to a Weibull model for both mosunetuzumab PFS and OS. As no changes have been made to the data at this stage, it is unclear why the company have decided at this stage of the appraisal process to alter their choice of preferred parametric model. However, the company suggests that these choices were made on plausibility grounds. A consequence of the company's modelling is that mosunetuzumab [redacted] to R² across the time horizon of the model.

The EAG's preference is to model PFS as in the previous EAG base-case, using a log-normal for R^2, and using a log-normal for mosunetuzumab until [redacted] before following the R^2 extrapolation.

The company question the EAG's selection of a log-normal distribution for OS. However, this choice matches the company's previous base case for mosunetuzumab in this comparison which also assumes over [redacted] of patients are alive at 20 years.

The EAG accept that the log-normal extrapolation could be considered optimistic, but as the EAG base case assumes equivalence between the treatments for OS and also PFS beyond [redacted], changing to a Weibull model will have little-to-no effect on the incremental benefit. The EAG has updated their base-case analyse to use a pooled Weibull model to allay the company's concerns.

RB Comparison

Here the company have changed their preferred PFS curves to match those preferred by the EAG (both log-normal). The EAG agrees with this decision.

For OS, the company have switched to an exponential model for both treatment groups but have not pooled the OS data, and model them separately.

The EAG maintains its preference to extrapolate using an exponential model fitted to pooled OS data, [redacted] (Figure 7 of the original EAG Report), and maximises the data being used in the model fitting.

OB Comparison

The company have removed this comparison from consideration so no longer present any survival extrapolations for this comparison.

Overall, the revised company analyses are an improvement over their previous base-cases. **However the EAG maintains that the relative effect of mosunetuzumab remains highly uncertain and this is best reflected in the EAG base case and previously presented scenario analyses.**

EAG Analyses

The following changes have been made to the EAG base case to account for new evidence/ data submitted by the company post TE and are summarised for R^2 and RB comparisons respectively. The impact of these changes is shown in Tables 14 - 16.

Mosunetuzumab vs R^2 comparison

EAG01: The EAG maintains its previous assumption and assumes equivalence in PFS between mosunetuzumab and R^2 at [redacted], whilst maintaining log-normal extrapolation for R^2 preferred by the company.

EAG02: For the OS, the EAG prefers a pooled OS, with the Weibull extrapolation preferred by the company in its new (post TE) base case.

EAG03: The EAG maintains the company's new assumption to apply costs of subsequent therapy at progression but revises the assumption that 35% of patients on R^2 receive Rituximab-Lenalidomide as retreatment therapy as advice from the EAG's clinical expert indicates that this does not reflect clinical practice (see Key Issue 3).

In contrast, the following assumptions **have not been applied** in the revised EAG base case

- **Half cycle correction:** Please refer to the EAG response to Key Issue 4

Mosunetuzumab vs RB comparison

The EAG maintains its previous assumptions for the below:

EAG04: The EAG prefers the log-normal distribution to extrapolate PFS in the RB arm as was the case in the original company submission whilst maintaining log-normal extrapolation for mosunetuzumab. This is also the preferred distribution in the revised (post TE) company base case.

EAG05: The EAG prefers an exponential model, fitted to pooled OS data from both treatments.

In contrast, the following assumptions **have not been applied** in the revised EAG base case

- **Half cycle correction:** Please refer to the EAG response to Key Issue 4

Additional comments

Across all analyses, the company and EAG the starting age of the modelled population is 60 years. In their TE response, the company have provided a reference which states that the median age at diagnosis of FL is 66.6.[3] Whilst this has negligible effect on the EAG base case due to the assumptions of equivalence for OS, the age parameter may become influential on future analyses if there is evidence to support separate modelling of OS. The NICE single technology appraisal of lenalidomide with rituximab for treated FL (TA627) explored a starting population age of 63 - 65. This parameter should be reviewed carefully in future reviews of this appraisal.

Table 14: Impact on ICER of individual assumptions in EAG base case

Preferred assumption	**Impact on ICER (£/QALY)**
Mosunetuzumab vs R^2	
Company base case	5,484 (cost-saving ICER)
EAG 01: PFS for R^2 set equal to Mosunetuzumab beyond [redacted]	3,030
EAG 02: Revised proportion of patients receiving subsequent therapy	Mosunetuzumab dominated
EAG 03: Weibull extrapolation for R^2 using OS from both treatments	Mosunetuzumab dominant
Mosunetuzumab vs RB	
Company base case	£37,821
EAG 05: Exponential model for OS fitted to pooled OS data from both treatments	£203,383

Table 15: Deterministic EAG base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosunetuzumab vs R²							
Mosunetuzumab	[redacted]	10.51	[redacted]				
R²	[redacted]	10.51	[redacted]	[redacted]	0	[redacted]	Mosun dominated
Mosunetuzumab vs RB							
Mosunetuzumab	[redacted]	9.23	[redacted]				
RB	[redacted]	9.23	[redacted]	[redacted]	0	[redacted]	£203,383

Table 16: Probabilistic EAG base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
Mosunetuzumab vs R²							
Mosunetuzumab	[redacted]	10.06	[redacted]				
R²	[redacted]	10.15	[redacted]	[redacted]	-0.09	[redacted]	Mosun dominated
Mosunetuzumab vs RB							
Mosunetuzumab	[redacted]	9.19	[redacted]				
RB	[redacted]	9.20	[redacted]	[redacted]	-0.01	[redacted]	£219,489

References

1. Kreif N, Grieve R, Radice R, Sekhon JS. Regression-adjusted matching and double-robust methods for estimating average treatment effects in health economic evaluation. *Health Services and Outcomes Research Methodology* 2013;**13**(2):174-202. http://dx.doi.org/10.1007/s10742-013-0109-2

2. Leonard JP, Trneny M, Izutsu K, Fowler NH, Hong X, Zhu J, *et al.* AUGMENT: A phase III study of lenalidomide plus rituximab versus placebo plus rituximab in relapsed or refractory indolent lymphoma. *J Clin Oncol* 2019;**37**(14):1188-99. http://dx.doi.org/10.1200/JCO.19.00010

3. Haematological Malignancy Research Network. *Quickstats*. URL: https://hmrn.org/statistics/quickstats (Accessed 26 October 2022).

Table 1: Deterministic EAG base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB_20K (£)	INMB_30K (£)
Mosunetuzumab vs R²									
Mosunetuzumab	[redacted]	10.51	[redacted]						
R²	[redacted]	10.51	[redacted]	[redacted]	0	[redacted]	Mosun dominated	[redacted]	[redacted]
Mosunetuzumab vs RB									
Mosunetuzumab	[redacted]	9.23	[redacted]						
RB	[redacted]	9.23	[redacted]	[redacted]	0	[redacted]	£203,383	[redacted]	[redacted]

Table 2: Probabilistic EAG base case cost-effectiveness results with revised assumptions (PAS discount)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)	INMB_20K (£)	INMB_30K (£)
Mosunetuzumab vs R²									
Mosunetuzumab	[redacted]	10.06	[redacted]						
R²	[redacted]	10.15	[redacted]	[redacted]	-0.09	[redacted]	Mosun dominated	[redacted]	[redacted]
Mosunetuzumab vs RB									
Mosunetuzumab	[redacted]	9.19	[redacted]						
RB	[redacted]	9.20	[redacted]	[redacted]	-0.01	[redacted]	£219,489	[redacted]	[redacted]

Additional EAG Submission as requested following the PMB

EAG Preferred OS extrapolation for RB comparison (the KM data in the economic model looks to contain an error as both arms are very similar, but the extrapolation is accurate)

EAG Preferred OS extrapolation for R^2 comparison (again, the KM data in the economic model looks to contain an error as both arms are very similar, but the extrapolation is accurate)

EAG Preferred PFS extrapolation for R^2 comparison

Table 1: Impact on ICER of individual assumptions in EAG base case

Preferred assumption	**Impact on ICER (£/QALY)**
Mosunetuzumab vs R^2	
Revised company base case (new PAS)	£27,422 (cost-saving)
EAG 01: For PFS, log-normal extrapolation for mosunetuzumab switched to follow R^2 extrapolation beyond [redacted]	£22,683 (cost-saving)
EAG 02: Revised proportion of patients receiving subsequent therapy	Mosunetuzumab dominated
EAG 03: Weibull extrapolation for R^2 using OS from both treatments	Mosunetuzumab dominant
Mosunetuzumab vs RB	
Company base case	£28,333
EAG 05: Exponential model for OS fitted to pooled OS data from both treatments	£143,234