NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Contents:

The following documents are made available to consultees and commentators:

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

1. **Company evidence submission** from **Celgene UK**

2. **Company response to NICE's request for clarification**

3. **Patient group, professional group and NHS organisation submission** from:
   a. Lymphoma Action
   b. Royal College of Radiologists-Royal College of Pathologists

4. **Expert personal perspectives** from:
   a. Dr Kim Linton – clinical expert, nominated by Celgene Ltd
   b. Professor Andrew Pettitt – clinical expert, nominated by RCR-RCPath
   c. Susan Jones – patient expert, nominated by Lymphoma Action
   d. Peter Loftus – patient expert, nominated by Lymphoma Action

5. **Evidence Review Group report** prepared by **Kleijnen Systematic Reviews**

6. **Company evidence submission addendum for FL population only** from **Celgene UK**

7. **Evidence Review Group addendum to the report** prepared by **Kleijnen Systematic Reviews**
   a. ERG addendum in response to the company addendum for FL population only
   b. ERG erratum to addendum in response to company addendum for FL population only

8. **Evidence Review Group – factual accuracy check**

9. **Draft Technical Report sent out for Technical Engagement**

10. **Technical engagement response** from **Celgene UK**
    a. Technical engagement response form
    b. State transition model addendum
    c. Clinical Validation summary

Technical engagement responses from experts:
None received

11. Technical engagement response from consultees and commentators:
a. Janssen

12. Evidence Review Group critique of company response to technical engagement prepared by **Kleijnen Systematic Reviews**
a. ERG critique of the company response to technical engagement
b. ERG critique of the company's STM model

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Document B

Company evidence submission

22 July 2019

File name	Version	Contains confidential information	Date
		Yes/no	

Instructions for companies

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

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

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

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

Highlighting in the template (excluding the contents list)

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

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

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

Contents

Tables and figures

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

B.1.1. Decision problem

The submission covers the technology's full marketing authorisation for this indication. The decision problem that this submission addresses is presented in Table 1.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with treated follicular lymphoma or marginal zone lymphoma	Adults with treated follicular lymphoma or marginal zone lymphoma	N/A
Intervention	Lenalidomide with rituximab (R^2)	Lenalidomide with rituximab (R^2)	N/A
Comparator(s)	• Rituximab monotherapy (R mono) • Rituximab in combination with chemotherapy • Established clinical management without lenalidomide (including but not limited to bendamustine)	For non-rituximab refractory patients: • Rituximab in combination with chemotherapy – Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP) – Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP) For rituximab refractory patients: • Established clinical management without lenalidomide – Obinutuzumab in combination with bendamustine (O-Benda)	For non-rituximab refractory patients: • R mono is not considered a relevant comparator as clinical expert opinion confirmed it is rarely used in the relapsed/refractory setting in the UK[1, 2] For rituximab refractory patients: • O-Benda is included as an option for rituximab-refractory patients under the category 'Established clinical management without lenalidomide'. This is the only NICE-recommended option for this patient group (via the CDF) and clinical experts stated this is the likely treatment choice for FL patients refractory to rituximab[1] • Bendamustine monotherapy (Benda mono) is not considered a comparator in this population given that clinical experts believe O-Benda has largely replaced use of Benda mono in rituximab refractory patients[1]

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Outcomes	• The outcome measures to be considered include: • Overall survival • Progression-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life	The outcome measures to be considered include: • Overall survival • Progression-free survival • Event-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life • Time to next antilymphoma treatment • Time to next chemotherapy treatment • Response rate to next antilymphoma treatment	Several efficacy outcomes have been presented in addition to those in the scope as several secondary and exploratory outcomes were reported in the AUGMENT and MAGNIFY studies that provide additional insight into the efficacy of R[2]
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 estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes	Adhering to the reference case, the cost-effectiveness of treatments is expressed in terms of incremental cost per quality adjusted life year. Adhering to the reference case, a lifetime horizon is used. Adhering to the reference case the economic analyses has been conducted from an NHS and	Confidential PAS schemes that apply to relevant subsequent comparator therapies are not included in these analyses as Celgene is not privy to such information

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any PAS for the intervention or comparator technologies will be taken into account.	Personal Social Services perspective Adhering to the reference case, the PAS has been applied in all economic analysis for all Celgene products.	
Subgroups to be considered	None listed in scope	• No specific subgroups	• N/A
Key: CDF, Cancer Drugs Fund; FL, follicular lymphoma; MZL, marginal zone lymphoma, NICE, National Institute for Health and Care Excellence.			

B.1.2. Description of the technology being appraised

In appendix C include the summary of product characteristics or information for use, and the European public assessment report, scientific discussion or drafts.

A description of lenalidomide in combination with rituximab (R^2) is presented in Table 2. The draft summary of product characteristics (SmPC) is presented in Appendix C.

Table 2: Technology being appraised

UK approved name and brand name	Lenalidomide (Revlimid®) in combination with rituximab (MabThera®) (R^2)
Mechanism of action	Lenalidomide binds to cereblon in the Cullin-4 RING E3 ubiquitin ligase and promotes degradation of the haematopoietic transcription factors Ikaros and Aiolos.[3, 4] Degradation of these substrates results in direct cytotoxic and immunomodulatory effects. Specifically, lenalidomide inhibits proliferation and enhances apoptosis of certain haematopoietic tumour cells (including FL and MZL tumour cells), enhances T cell- and NK cell-mediated immunity and increases the number of NK, T and NK T cells. Single agent lenalidomide reactivates dysfunctional T and NK cells from FL patients.[3] Rituximab is an anti-CD20 antibody; its mechanisms of action are to augment NK cell-mediated killing of malignant B cells via ADCC, to enhance ADCP and to induce complement-mediated killing.[3] The combination immunotherapy of lenalidomide and rituximab acts by complementary mechanisms including direct tumour apoptosis in FL and MZL and immune-mediated activities, such as activation of NK cells and immune synapse formation, resulting in increased ADCC in vitro.[4] While single-agent lenalidomide and rituximab increased formation of lytic NK cell immunological synapses with primary FL tumour cells, the combination was superior and correlated with enhanced cytotoxicity.[3] The combination of lenalidomide and rituximab is characterized by immune enhancement, and not the immunosuppression that has been observed with immunochemotherapy.[3] Immunophenotyping of FL patient samples from the first-line Phase 3 RCT, RELEVANCE,[5] revealed that R^2 treatment increased circulating T and NK cell counts, while R-chemo was associated with reduced numbers of these cells.[3] The effects of lenalidomide and chemotherapy on neutrophil maturation have been compared using an in vitro model of myeloid maturation. Bendamustine was shown to be cytotoxic to neutrophil progenitors while lenalidomide caused a reversible block in neutrophil maturation without loss of cell viability. This may explain the lower rates of neutropenia observed with R^2 versus R-chemo.

UK approved name and brand name	Lenalidomide (Revlimid®) in combination with rituximab (MabThera®) (R^2)
Marketing authorization/CE mark status	R^2 does not currently have a UK marketing authorization, although CHMP opinion is anticipated on [redacted], and marketing authorization is expected in [redacted] Lenalidomide currently has a UK marketing authorization and was first approved by the EMA in June 2007. It is currently approved for the following indications: • As monotherapy for the maintenance treatment of adult patients with newly diagnosed multiple myeloma who have undergone autologous stem cell transplantation • As combination therapy with dexamethasone, or bortezomib and dexamethasone, or melphalan and prednisone, for the treatment of adult patients with previously untreated multiple myeloma who are not eligible for a transplant • In combination with dexamethasone, for the treatment of multiple myeloma in adult patients who have received at least one previous therapy • As monotherapy, for the treatment of adult patients with transfusion-dependent anaemia due to low- or intermediate-1-risk myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality when other therapeutic options are insufficient or inadequate • As monotherapy, for the treatment of adult patients with relapsed or refractory mantle cell lymphoma
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Lenalidomide in combination with rituximab (anti-CD20 antibody) is indicated for the treatment of adult patients with previously treated FL or MZL.[4] **Restrictions:** Treatment should not be initiated in the following patients: • Those with hypersensitivity to the active substance or to any of the excipients • Women who are pregnant • Women of childbearing potential unless all of the conditions of the Pregnancy Prevention Programme are met • Children and adolescents from birth to less than 18 years
Method of administration and dosage	Lenalidomide is administered orally and rituximab is administered by (IV) infusion. Lenalidomide capsules should be taken orally at about the same time on the scheduled days.[4] The recommended starting dose of lenalidomide is 20 mg, orally once daily on Days 1 to 21 of repeated 28-day cycles for up to 12 cycles of treatment. The recommended starting dose of rituximab is 375 mg/m^2 IV every week in Cycle 1 (Days 1, 8, 15, and 22) and Day 1 of every 28-day cycle for Cycles 2 through 5.[4]
Additional tests or investigations	The following tests/investigations are recommended when administering R^2:[4] • Medically supervised pregnancy tests with a minimum sensitivity of 25 mIU/mL must be performed for women of childbearing potential, including those who practice

UK approved name and brand name	Lenalidomide (Revlimid®) in combination with rituximab (MabThera®) (R^2)
	abstinence, before treatment, every four weeks during treatment, and 4 weeks after the end of treatment (except in the case of confirmed tubal sterilization) • Patients with known risk factors for myocardial infarction (including prior thrombosis) should be closely monitored, and action should be taken to try to minimize all modifiable risk factors (e.g. smoking, hypertension, and hyperlipidaemia) • A complete blood cell count should be performed at baseline and then weekly for the first 3 weeks of Cycle 1 (28 days), every 2 weeks during Cycles 2 through 4, and then at the start of each cycle thereafter • Careful monitoring and evaluation for TFR is recommended. • Careful monitoring and evaluation for TLS is recommended. Patients should be well hydrated and receive TLS prophylaxis, in addition to weekly chemistry panels during the first cycle or longer, as clinically indicated
List price and average cost of a course of treatment	The list price of lenalidomide is £3,426.00 per 21 x 2.5 mg pack, £3,570.00 per 21 x 5 mg pack, £3,780.00 per 21 x 10 mg pack, £3,969.00 per 21 x 15 mg pack and £4,168.50 per 21 x 20 mg pack. The list price of rituximab is £349.25 per 2 x 100 mg vials and £873.15 per 1 x 500 mg vial (MabThera®). The prices of biosimilar rituximab also used in the economic analyses are £314.33 per 2 x 100 mg vials and £785.84 per 1 x 500 mg vial. Assuming the starting dose of 20 mg lenalidomide and the AUGMENT mean patient BSA of 1.85 m^2, the per 28-day cycle costs are £4,168.50 (list price) or [redacted] (with PAS) for lenalidomide and £4,845.98 (cycle 1) and £1,211.50 (cycles 2–5) for rituximab.[39] Based on the median treatment durations for R^2 patients in AUGMENT ([redacted] months for lenalidomide and [redacted] months for rituximab)[39], the average cost of a course of R^2 treatment is: List price: £60,438 Patient access scheme price: [redacted]
Patient access scheme (if applicable)	There is a confidential simple discount PAS for lenalidomide which applies to all current and future indications

Key: ADCC, antibody-dependent cellular cytotoxicity; ADCP, antibody-dependent cellular phagocytosis; CHMP, Committee for Medicinal Products for Human Use; EMA, European Medicines Agency; FL, follicular lymphoma; IV, intravenous; MZL, marginal zone lymphoma; NSAID, non-steroidal anti-inflammatory drug; PD, progressive disease; NK, natural killer; R-chemo, rituximab plus chemotherapy; RCT, randomized controlled trial; TFR, tumour flare reaction; TLS, tumour lysis syndrome.

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

Disease background

Lymphoma is a cancer of the lymphatic system of which there are two main types: Hodgkin's lymphoma (HL) and non-Hodgkin's lymphoma (NHL). NHLs are a heterogeneous group of lymphoproliferative cancers, most of which (80%–95%) arise from B-cells and the remainder from T-cells.[6] There are many types of NHL, which can be divided into indolent (iNHL; the incurable, usually slower-growing and 'low-grade' type) and aggressive (the faster-growing, 'high-grade' type). Of iNHL, the two most common types are follicular lymphoma (FL), and marginal zone lymphoma (MZL).[7, 8] In the UK, FL and MZL account for 18% and 12% respectively of all iNHL.[9-12]

The World Health Organization (WHO) system is used to grade FL from 1 to 3 according to the proportion of centroblasts (large cells) found amongst centrocytes (small/medium sized cells).[13] Grades 1, 2 and 3A are considered indolent disease, whereas Grade 3B is considered an aggressive lymphoma.[14] The extent of disease is classified according to the Ann Arbor system with stages III, IV and bulky Stage II representing advanced disease.

The Follicular Lymphoma-specific International Prognostic Index (FLIPI) is used to assess the prognostic factors in FL.[14] The FLIPI risk factors are: disease in ≥4 lymph node regions; age >60 years; elevated lactate dehydrogenase (LDH); Ann Arbor Stage III–IV; and haemoglobin level <12 g/dl. Patients are considered high risk if they have three or more factors, intermediate risk if they have two risk factors, and low risk if they have 0 or 1 risk factor.[14] A revised FLIPI 2 index incorporates β2-microglobulin, diameter of largest lymph node and bone marrow involvement. Additional prognostic information is derived from the Groupe d'Etude des Lymphomes Folliculaires (GELF) criteria used to define tumour burden.[14]

For MZL, there are three subtypes:[8, 15]

- Mucosa-associated lymphoid tissue (MALT) lymphoma – the most common type of MZL, usually categorized as gastric MALT and non-gastric MALT lymphoma

- Nodal marginal zone lymphoma (NZML)
- Splenic marginal zone lymphoma (SMZL)

Aetiology, course and prognosis

FL is an incurable disease of immune dysfunction[3] with a substantial symptom burden, including B symptoms (night sweats; unexplained fevers and weight loss), fatigue and bone marrow failure.[16, 17] In addition, the local mass effect of lymph node enlargement may lead to restricted movement, cosmetic disfiguration and pain.[18-23] [24]

FL is typically characterized by an indolent clinical course, with recurrent remissions and relapses; with each relapse, the disease becomes more resistant and/or refractory to treatment and each remission becomes shorter than the preceding one. Recently published US data for a large FL population illustrate this point;[25] median PFS (in years) decreased from first- to fifth-line treatment: first: 6.6; second: 1.5; third: 0.83, fourth: 0.69; fifth: 0.68. Notably, for patients considered refractory to rituximab-containing regimens, median PFS at second-line treatment and beyond was shorter compared with non-R-refractory patients (0.47 vs. 1.58 years).[25]

The incidence of FL increases with age, with a median presentation between 60 and 65 years, and a slightly higher incidence in females.[18] At diagnosis, most patients have advanced disease (Stage III: 18.4%; Stage IV: 50.5%).[19] The overall 5-year relative survival rate for patients with FL in the UK is 89% and specifically for Stages III and IV, is approximately 80%.[20, 21] Since the introduction of rituximab, the median OS of patients with FL has extended to 20 years in some studies,[22] compared with 9 years previously reported.[23] Despite the available treatment options, most patients eventually die from this disease.[24]

Patients with MZL represent a generally older (median age at diagnosis is 70–73 years)[18] and more advanced population compared with those with FL.[18, 19, 26] The primary organ of origin is the most significant prognostic factor and dictates organ-specific management strategies.[26]

Patients with MZL have a similar prognosis to those with FL. In the UK, the overall 5-year survival for extranodal (MALT) MZL is 89.5% and 77.3% for systemic MZL (i.e.,

SMZL/NMZL).[20] The median OS for UK patients with MALT MZL has been reported as 12.6 years, compared with 8–10 years for SMZL, and 8.3 years for NMZL.[27, 28]

In nine articles identified in a targeted literature review of clinical studies investigating various therapies in relapsed/refractory NHL, and presenting data split by MZL and FL, outcomes tended to be similar between these two histologies.[29-37]

Epidemiology

For FL, the Office of National Statistics (ONS) estimates 2,168 patients were diagnosed with FL in 2017 in England.[9] Of these, [redacted] (n=[redacted]) have first-line chemotherapy, while [redacted] (n=[redacted]) undergo a 'watch and wait' approach[19], of which [redacted] (n=[redacted]) go on to receive chemotherapy.[38] Therefore, the total number of FL patients on first-line chemotherapy is [redacted]. Of these, [redacted] (n=[redacted]) are expected to receive second-line chemotherapy or beyond.[38]

For MZL, of 12,065 patients diagnosed with NHL in England in 2017[9], the number of patients with the different MZL types is anticipated to be:

- MALT lymphoma: 7.7% = 928 patients[10]
- Splenic MZL: 2.0% = 241 patients[11]
- Nodal MZL: 2.0% = 241 patients[12]

Therefore, the total number of MZL patients in England in 2017 is estimated at 1,411 patients. Of these, 34.9% (n=492) have first-line chemotherapy, while 49.9% (n=704) undergo a 'watch and wait' approach[19] of which [redacted] (n=[redacted]) go on to receive chemotherapy.[38] Therefore, the total number of MZL patients on first-line chemotherapy is [redacted]. Of these, [redacted] (n=[redacted]) are expected to receive second-line chemotherapy or beyond.[38]

Burden of disease

FL is associated with a substantial symptom burden, including lymph node enlargement (leading to swellings in the neck, armpit and/or groin), B symptoms, fatigue and cytopenias.[16, 17, 39, 40] The symptoms of MZL are similar to those of FL, although some site-specific complications may also be present such as gastric and bowel involvement in MALT lymphoma, and enlarged spleen in splenic MZL.[10-12]

The symptomatic burden of FL and MZL leads to a detrimental impact on patient health-related quality of life (HRQL). One study of 148 patients with FL in the Netherlands reported the significant impact of FL symptoms compared with a normative population, particularly in the areas of fatigue, dyspnoea, sleeping problems, appetite loss, constipation, diarrhoea and financial issues, as measured on the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire, Core 30 (EORTC QLQ-C30).[41] Similar QoL impacts were reported in a mixed population of 97 iNHL survivors (including 67 with MZL and 27 with FL).[42]

Patients whose disease has relapsed have been reported to experience a statistically significant decrement in HRQL compared with those who are newly diagnosed.[24] One UK cross-sectional study reported decreasing HRQL with each relapse as measured on the functional assessment of chronic illness therapy–lymphoma (FACT-Lym), as well as significantly higher anxiety and depression scores, and higher work activity impairment, in comparison to patients newly diagnosed with FL.[24] Fear of relapse also has a detrimental effect on HRQL.[42]

The toxicity of chemotherapy (e.g. nausea, vomiting, hair loss, skin irritation, sore mouth, dysphagia, and gastrointestinal problems), can have an even more negative HRQL impact than the disease itself.[24] This is, characterised by significantly worsening health functioning (p=0.004), depressive symptoms (p=0.005) and activity impairment (p=0.009) compared with FL patients in remission but not on treatment.[24] Chemotherapy treatment also had a detrimental impact on patients ability to work.[42, 43] Clinical experts at an advisory board agreed that patients regard remission as having substantial positive impact on their QoL.[1 42, 43]

FL and MZL also incur a substantial caregiver burden in terms of both time commitment and psychological burden.[42, 43] Of 84 patients with iNHL, 23% required caregiver assistance, with 33% of caregivers' working days missed as a consequence.[43] Unpaid caregivers provided a mean of 9.8 days of care in the 30 days before data collection, with a mean of 11.3 days of absenteeism, implying substantial social and economic burden.[43]

FL also incurred substantial direct costs on the healthcare system, as reported in a UK economic analysis of patients with FL enrolled in the Haematological Malignancy

Research Network (HMRN) between 2004 and 2011, and followed until 2015.[38] Average lifetime costs for FL ranged from £6,165 to £63,864 per patient, with an overall estimated cost to the UK healthcare system of £60–65 million, representing around 10% of the NHS budget for haematological cancers at the time of publication. No studies have been identified specifically reporting the economic burden of MZL.

Current pathway of care

A summary of clinical guidelines relevant to this submission is presented in Appendix L. NICE has issued a clinical pathway for the management of FL that encompasses guidance from the NICE guideline for management of NHL, and relevant technology appraisals.[44] NICE has also issued a clinical pathway for MALT lymphoma but not nodal or splenic MZL.[45] In addition, the British Committee for Standards in Haematology (BCSH) issued guidance in 2012 for the management of FL.[17] The European Society for Medical Oncology (ESMO) has also provided guidance as to the management of FL[14], and the subtypes of MZL.[15] The National Comprehensive Cancer Network (NCCN) also publishes guidelines on B-cell lymphomas including FL and MZL.[46].

For first-line treatment of advanced stage FL, these guidelines recommend R-chemo for patients requiring systemic therapy.[14, 17, 44, 47] Similarly, guidelines agree on offering further R-chemo in the setting of relapsed/refractory disease,[44, 47 14] with consideration given to patients' rituximab-refractory (R-refractory) status.[1] The definition of '*rituximab-refractory*' published in NICE TA472 and used in the UK is '*best response of progressive disease (PD) or stable disease (SD) to treatment with a rituximab-containing regimen (single agent or combination) or response lasting less than 6 months following last rituximab dose*.'[48 14, 44, 47] For R-refractory patients, NICE recommends obinutuzumab plus bendamustine (O-Benda), currently available via the Cancer Drugs Fund (CDF). NICE does not recommend bendamustine monotherapy (Benda mono) for R-refractory patients.

Lenalidomide plus rituximab (R^2) was described as an innovative treatment in the 2016 ESMO FL guidelines which noted the promise seen in Phase II studies.[14] In the US, where R^2 has recently been approved, the National Comprehensive Cancer

Network (NCCN) includes R^2 as a preferred regimen for second-line and subsequent treatment of FL and MZL.[46]

For MZL, limited specific guidance is available. NICE provides recommendations only for MALT MZL, and not for NMZL or SMZL,[45] noting treatment with R-chemo as an option for patients with relapsed/refractory disease requiring systemic therapy. Guidelines from ESMO highlight that management of MZL is similar to that of FL for patients in the relapsed/refractory setting.[14]

To understand the current management of relapsed/refractory FL and MZL in UK clinical practice, an advisory board was conducted in March 2019, involving six UK clinical experts in NHL and two health economics experts.[1]

Advisors agreed that R-Benda is the primary first-line intervention for treatment of FL in the UK, with a smaller proportion of patients treated with R-CVP or R-CHOP. Prior therapy is an important factor in selecting subsequent therapies when patients progress.[1] Importantly, after treating with R-Benda in first-line, clinicians are reluctant to re-challenge with further bendamustine due to concerns regarding cumulative toxicity. Accordingly, the predominant treatments for relapsed/refractory disease in the UK are R-CHOP and R-CVP. Patients considered R-refractory will likely receive O-Benda.[1] For third-line treatment and beyond, clinical experts agreed there is no standard approach to management and new treatment options are needed in this area of unmet need.[1]

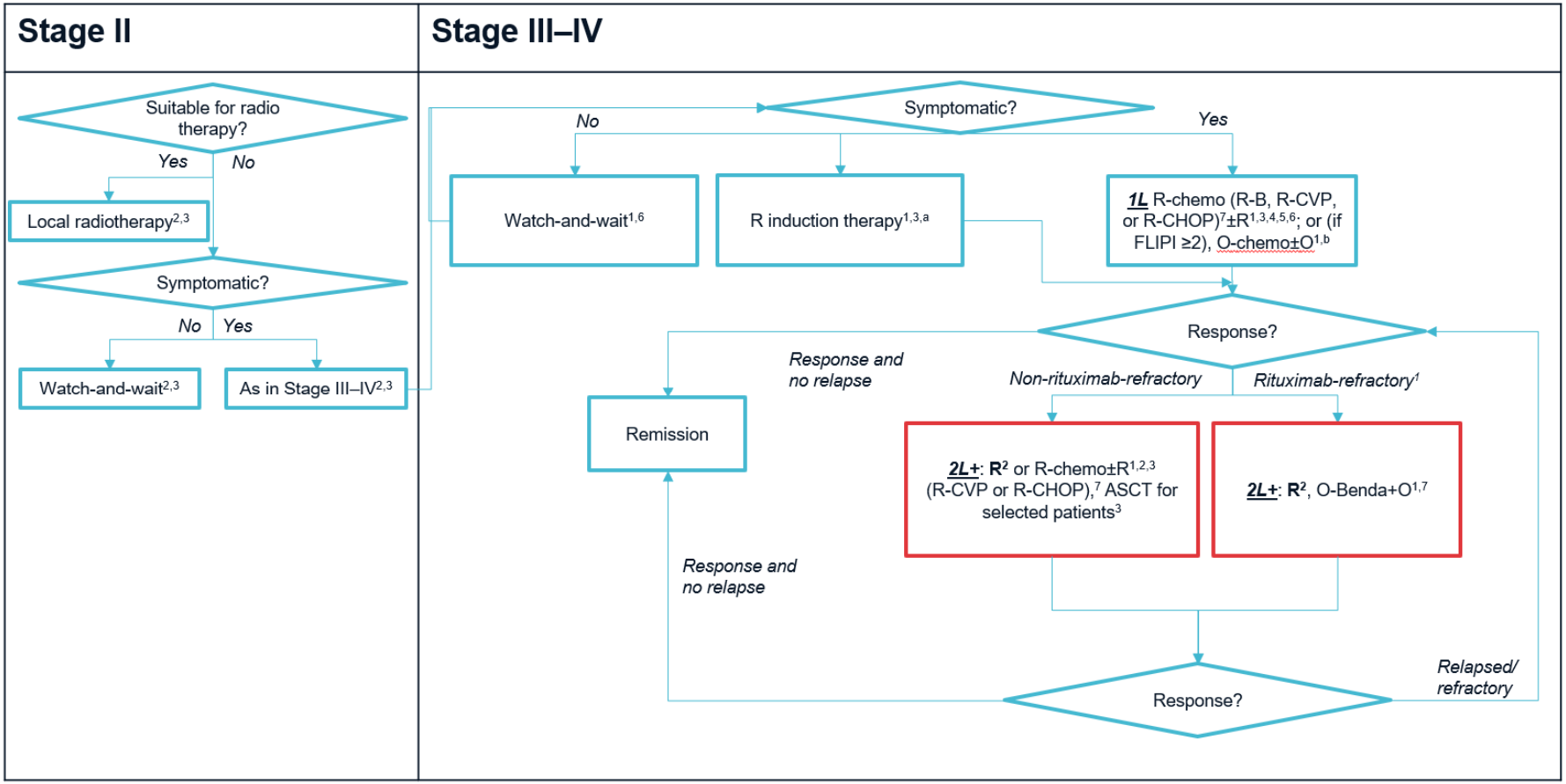
For patients with relapsed/refractory MZL, clinical advisors agreed that management and treatment outcomes were broadly similar to those for patients with relapsed/refractory FL,[1] albeit noting that treatment options are further limited for R-refractory MZL patients given obinutuzumab lacks an indication in MZL.[49]

Using the current treatment guidelines, and the insights gained from the advisory board, current treatment pathways in FL and MZL have been developed, also incorporating the proposed positioning of R^2 (Figure 1 and Figure 2).

R^2 is considered an important additional therapeutic intervention across the spectrum of relapsed/refractory FL and MZL as an alternative to existing chemotherapy-based regimens.[1] Clinical experts agreed that R^2 should be broadly

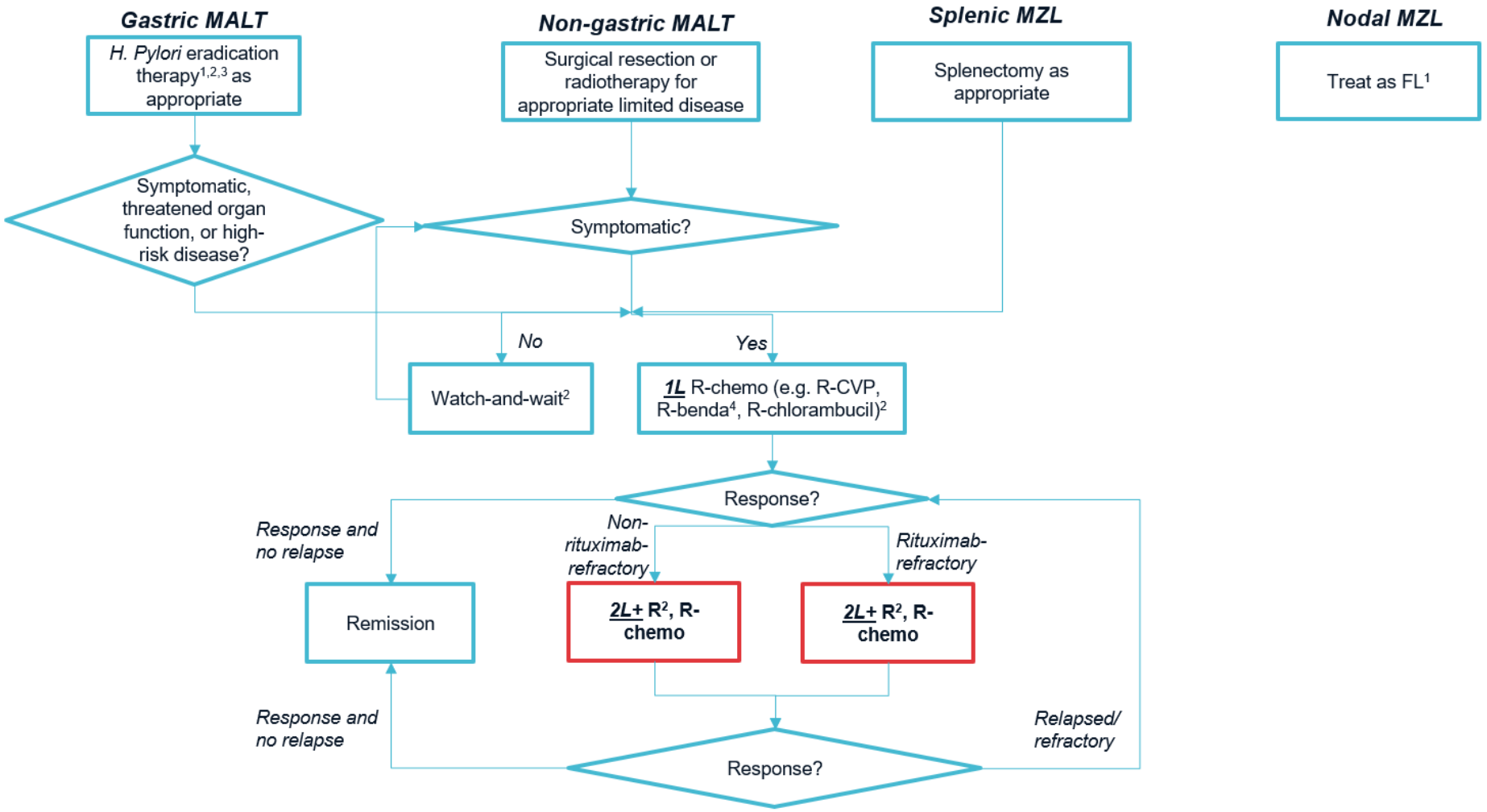
available for the relapsed/refractory FL and MZL populations, including those defined as R-refractory. Anticipating widespread application of R^2 across different relapsed/refractory FL/MZL patient types and lines of therapy, clinical experts believe its introduction would likely lead to altered sequencing of current immunochemotherapies, but with no specific treatment anticipated to be entirely displaced.[1]

Figure 1: Current clinical pathway for treatment of follicular lymphoma, with proposed positioning of R²



Key: 1L, first-line; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisolone; CVP, cyclophosphamide, vincristine, prednisolone; O, obinutuzumab; R, rituximab; R-chemo+R, rituximab and chemotherapy induction followed by rituximab maintenance therapy; SCT, stem cell transplant.
Notes: [a], R induction therapy is not mentioned as an option for asymptomatic patients in guidance more recent than the NICE guidelines (e.g. TA472, ID1379); [b], In TA513, it was noted that R mono and Benda-mono are occasionally used here, but the Committee concluded that R-chemo+R was the relevant treatment.
Sources: 1. NICE TA472, 2017;[48] 2. NICE, 2016;[44] 3. NICE, 2016;[47]; 4. NICE, 2012;[50] 5. NICE, 2011;[51] 6. NICE, 2018;[52] 7. Celgene, 2019.[1]

Figure 2: Current clinical pathway for treatment of marginal zone lymphoma and proposed positioning for R²



Key: benda, bendamustine; CHOP, cyclophosphamide doxorubicin vincristine prednisolone; CVP, cyclophosphamide vincristine prednisolone; FL, follicular lymphoma; H, Helicobacter; MALT, Mucosa-associated lymphoid tissue; MZL, marginal zone lymphoma; R, rituximab.
Sources: 1. Dreyling 2013;[15] 2. NICE, 2016;[47] 3. Lymphoma Research Foundation, 2018;[53] 4. Celgene, 2019.[1]

In line with the NICE scope for this appraisal[54], relevant comparators for R^2 can be split into those for a non-R-refractory population, and those for a R-refractory population. As such, the comparators for R^2 are as follows:

- Non-R-refractory:
  - Rituximab in combination with chemotherapy; predominantly:
    - R-CHOP
    - R-CVP

Note: According to clinical experts, R mono is rarely used in the relapsed/refractory setting in UK clinical practice.[1, 2] R-Benda is primarily used in a first-line setting and clinicians are reluctant to re-challenge relapsed/refractory patients with bendamustine in subsequent lines of therapy.[1] Therefore, R-mono and R-Benda are not considered relevant comparators for this appraisal.

- R-refractory:
  - Established clinical management without lenalidomide, specifically:
    - O-Benda; however, obinutuzumab does not have a marketing authorization for MZL so this is a relevant comparator for the FL population only[49]

Note: Benda mono is not considered a comparator in this population given that clinical experts believe that O-Benda has largely replaced use of Benda mono in R-refractory patients.[1]

Existing treatment options leave substantial unmet need in the management of relapsed/refractory FL and MZL.[1] The predominant regimens, R-CHOP and R-CVP, used to treat these populations represent the same chemotherapy-based modality as used in front-line treatment, associated with well-documented toxicities including immune suppression. These interventions do not fully address the immune dysfunction associated with FL, and novel biological agents are needed that target the immune microenvironment and cellular pathways leading to a more durable response.[55] In addition to the above, clinical experts identified the following specific populations with high unmet needs:[1]

- Patients who cannot tolerate aggressive treatments have very limited options in second-line and beyond
- Patients who are refractory to first-line therapy or who relapse early, particularly the POD24 population (those progressing within 24 months of commencement of

initial therapy; approximately 20% of first-line patients) are documented to have a poor prognosis[56 57]

- Patients receiving Benda-based treatment in first-line and who are R-refractory; physicians are reluctant to re-challenge with Benda, thus an alternative to O-Benda is sought for this group
- Patients requiring third-line treatment and beyond, where there is no standard approach to management
- Patients with relapsed/refractory MZL, for whom there are currently no licensed treatments specifically indicated for their treatment

R^2 addresses these unmet needs and limitations. Firstly, R^2 represents a new immunotherapy treatment modality not reliant on standard chemotherapeutic agents, with a different toxicity profile. In the head-to-head comparison of R^2 and R-chemo in the RELEVANCE RCT involving patients with treatment-naive FL, R^2 was associated with a lower incidence of any grade TEAEs compared with R-chemo.[5] The R^2 components, lenalidomide and rituximab, are both well-characterized and familiar to haematologists and oncologists who have long-term experience with these agents. Consequently, the R^2 toxicity profile should be familiar and manageable in clinical practice. Secondly, R^2 can restore and enhance patients' own immune function through the complementary and synergistic activity of lenalidomide and rituximab (described in Table 2).[3, 4] Taking together its mechanism of action and clinical data,[3, 58, 59] R^2 is anticipated to provide durable benefit for a wide range of relapsed/refractory FL and MZL patients, providing an alternative to current chemotherapy-based treatment options. Notably, once approved by the EMA, R^2 would be the only treatment specifically licensed for relapsed/refractory MZL in Europe.

B.1.4. Equality considerations

No equity or equality issues are anticipated for the appraisal of R^2.

B.2. Clinical effectiveness

B.2.1. Identification and selection of relevant studies

See Appendix D1 for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised. A systematic literature review (SLR) was conducted in September 2017 and subsequently updated in April 2019, to identify published clinical trial data on the efficacy and safety of R^2 in patients with relapsed/refractory FL or MZL.[60] In total, the SLR identified 45 studies (13 RCTs, 32 non-RCTs). Of these, 39 studies are not relevant for the submission because they did not investigate comparators of interest (1 for lenalidomide, 1 for obinutuzumab+lenalidomide, 4 for idelalisib, 2 for copanlisib, 3 for ibrutinib, 6 for R-Benda, 1 for other bendamustine-containing regimens, 15 for R mono, 5 for Benda mono, and 1 for tazemetostat). Therefore, there were a total of 6 relevant studies.

Of these, there were 5 relevant RCTs (AUGMENT (R^2),[58] MAGNIFY (R^2),[61] ALLIANCE (R^2),[62] Van Oers (R-CHOP)[63] and GADOLIN (O-Benda)[64]) and 1 relevant non-RCT (Tuscano 2014 (R^2)).[65] Of note, the SLR found no studies for the relevant comparator R-CVP.

Given that Phase III data are available for R^2, the two identified Phase II R^2 studies (the RCT, ALLIANCE and the non-RCT, Tuscano, 2014) have not been discussed in detail in this submission. The AUGMENT and MAGNIFY studies are described in detail in the following sections (Sections B.2.3-B.2.7), and the two relevant comparator studies are discussed in the indirect comparison section (Section B.2.9).

B.2.2. List of relevant clinical effectiveness evidence

This submission focuses primarily on one randomized Phase III study (AUGMENT) and supporting data from the induction phase of a randomized Phase IIIb open-label study (MAGNIFY) that provide evidence on the clinical benefits of R^2. One study (ALLIANCE) covered briefly in this submission also provides supporting evidence. All three studies are summarized in Table 3:

- AUGMENT is a randomized, double-blind, multicentre, Phase III study of R^2 versus rituximab plus placebo (R mono) in non-R-refractory patients with FL Grade 1–3a or MZL
- MAGNIFY is a randomized, open-label, multicentre, Phase IIIb study of R^2 induction therapy followed by either R^2 maintenance therapy or R mono maintenance therapy in patients with FL Grade 1–3b, MZL, or mantle cell lymphoma
  - Data for FL and MZL patients from the induction phase only are presented in this submission, to provide supportive data for the AUGMENT study population, and to provide additional data on the R-refractory population excluded from the AUGMENT study
- ALLIANCE is a randomized, multicentre, Phase II study of R^2 versus lenalidomide monotherapy in patients with previously treated FL and prior rituximab
  - Although the results of this study support the outcomes of the AUGMENT study, it is not included in the economic model due to the availability of phase III studies.

In total, AUGMENT and MAGNIFY enrolled 728 patients with relapsed/refractory FL or MZL. In AUGMENT all patients were non-R-refractory and in MAGNIFY, 41% of patients were R-refractory.

Table 3: Summary of RCTs in support of R^2

Study	AUGMENT					MAGNIFY					ALLIANCE				
Study design	Phase III, multicentre, double-blind, randomized study					Phase IIIb, multicentre, open-label, randomized study					Phase II, multicentre, randomized study				
Population	358 patients with relapsed/refractory FL or MZL All patients non-R-refractory					370 patients with relapsed/refractory FL or MZL 41% of patients R-refractory					66 patients with recurrent FL, previously treated with rituximab, either as monotherapy or in combination with chemotherapy				
Intervention(s)	Lenalidomide 20 mg or 10 mg once daily on Days 1 to 21 of every 28-day Cycle up to 12 cycles combined with rituximab 375 mg/m^2 every week in Cycle 1 and on Day 1 of every 28-day Cycle from Cycles 2 through 5					**Induction phase**: Lenalidomide 20 mg or 10 mg once daily on Days 1 to 21 of every 28-day Cycle up to 12 cycles combined with rituximab 375 mg/m^2 every week in Cycle 1 and on Day 1 of every other 28-day Cycle from Cycles 3 through 11 **Maintenance phase**: Lenalidomide 10 mg for 18 cycles and rituximab 375 mg/m^2 on Day 1 of every other cycle from Cycles 13 through 29, followed by optional lenalidomide 10 mg alone					Lenalidomide 15 mg once daily on Days 1 to 21, of a 28-day Cycle 1 and then 20 mg once daily on Days 1 to 21, of 28-day Cycles 2 to 12 combined with rituximab 375 mg/m^2 weekly for 4 weeks of Cycle 1				
Comparator(s)	Rituximab plus placebo (R mono)					**Induction phase:** No comparator **Maintenance phase:** R mono					Lenalidomide monotherapy				
Indicate if trial supports application for marketing authorization	Yes	X	**Indicate if trial used in the economic model**	Yes	X	Yes	X	**Indicate if trial used in the economic model**	Yes	X	Yes	X	**Indicate if trial used in the economic model**	Yes	
	No			No		No			No		No			No	X
Rationale for use/non-use in the model	Pivotal study supporting this indication					Study supporting this indication					Phase II study, small patient numbers, only 1 year of extra follow-up				

Study	AUGMENT	MAGNIFY	ALLIANCE
Reported outcomes specified in the decision problem	PFS OS ORR AEs QoL, assessed by QLQ-C30 and EQ-5D-3L	PFS OS ORR AEs **Note:** QoL is not assessed in the induction phase, only for the overall study	PFS ORR CR OS TTP
All other reported outcomes	TTNLT RTNLT DOR DOCR HT PFS2	TTNLT DOR DOCR TTR	
Key: AE, adverse event; CR, complete response; DOCR, duration of complete response; DOR, duration of response; FL, follicular lymphoma; HT, histological transformation; MCL, mantle cell lymphoma; MZL, marginal zone lymphoma; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; PFS2, PFS on next antilymphoma treatment; QoL, quality of life; R, rituximab; R mono, rituximab monotherapy; RTNLT, response to next antilymphoma treatment; TTNLT, time to next antilymphoma treatment; TTP, time to progression; TTR, time to response.			

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

AUGMENT

Study design

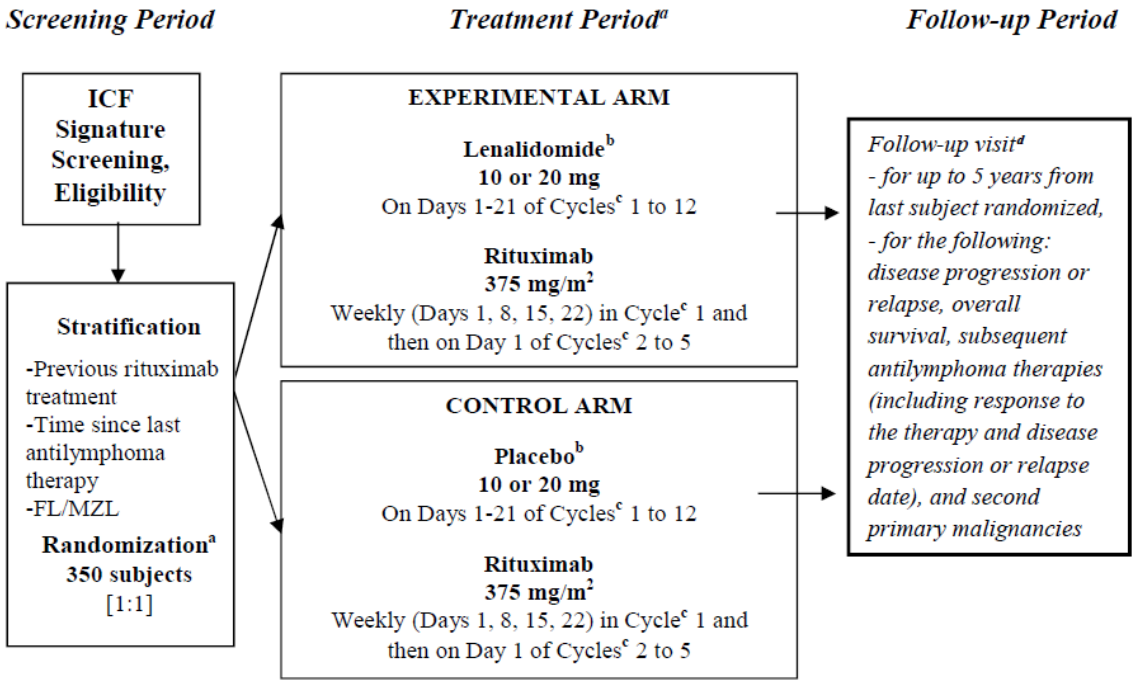
AUGMENT is a multicentre, double-blind, randomized Phase III study, that provides evidence of the clinical benefits of R^2 compared with rituximab plus placebo (R-mono). AUGMENT is the pivotal study supporting this indication and was the key study used in the European Medicines Agency (EMA) regulatory submission. The study was conducted across 96 sites in 17 countries.

To be eligible for inclusion in the study, patients had to be aged ≥18 years, with histologically confirmed MZL or Grade 1, 2, or 3a FL (Grade 3b FL patients were excluded). Patients were required to have been previously treated with at least one systemic chemotherapy, immunotherapy or R-chemo. Initially, rituximab-naïve patients were included in the study; however, a protocol change required patients to have received at least two previous doses of rituximab. This change was carried out to ensure a study population that aligned with a population commonly seen in clinical practice. Furthermore, patients had to have documented relapsed/refractory FL or MZL; however, R-refractory patients were excluded (inclusion and exclusion criteria are summarized in Table 4 and presented in full in Appendix M1).

During the treatment period, patients underwent efficacy and safety assessments for a maximum of 12 cycles. Patients received oral lenalidomide or placebo at a starting dose of 10 mg (if CrCl ≥30 mL/min and <60 mL/min) or 20 mg (if CrCl ≥60 mL/min) once daily on Days 1 to 21 in each 28-day cycle, combined with four-weekly infusions of rituximab intravenously (IV) at a dose of 375 mg/m^2, followed by four additional doses on Day 1 of Cycles 2, 3, 4, and 5. Patients were stratified by prior rituximab treatment (yes vs. no), time since last antilymphoma therapy (≤2 vs. >2 years), and histology (FL vs. MZL), and then randomized 1:1 to R^2 or R-mono for 12 cycles. Treatment was terminated upon relapse or progression of disease, withdrawal of consent, or unacceptable toxicity.

Figure 3 presents a study design schematic for AUGMENT.

Figure 3: Study design for AUGMENT



Key: CrCl, creatinine clearance; FL, follicular lymphoma; ICF, informed consent form; MZL, marginal zone lymphoma.
Notes: [a], treatment had to begin as soon as possible after randomization but no later than one week after randomization; [b], 10 mg if CrCl ≥30 mL/min but <60 mL/min; 20 mg if CrCl ≥60 mL/min; [c], cycle defined as lenalidomide or placebo Cycle of 28 days (21-day treatment and 7-day rest period); [d], all randomized patients were followed for disease progression and overall survival using the same schedule. This included patients who discontinued the protocol-specified treatment or the study early for any reason without documented evidence of disease progression or relapse.
Source: Celgene, 2018.[66]

Three analysis populations were evaluated in the AUGMENT study:

- The intention-to-treat (ITT) population, defined as all randomized patients
- The modified intention-to-treat (mITT) population, defined as all randomized patients who received at least one dose of study medication, had a confirmed diagnosis of relapsed/refractory FL or MZL by central pathology review (except splenic MZL which was based on local pathology assessment), and had baseline (screening) and at least one post-baseline tumour assessment
- The safety population, defined as all patients who received at least one dose of study treatment

Primary efficacy analyses were conducted on the ITT population. The primary endpoint of the study was PFS, as assessed by the Independent Review Committee (IRC) using a modification of the 2007 International Working Group Response

Criteria (IWGRC [i.e. without a positron emission tomography scan]). Efficacy was assessed further in the ITT population through a number of secondary endpoints, including overall response rate (ORR), complete response (CR) rate, time to next antilymphoma treatment (TTNLT), duration of response (DOR), durable complete response rate (DCRR; defined as the proportion of patients that stayed in complete response for at least one year) and duration of complete response (DOCR). Efficacy analyses were also conducted on the mITT population to be used as supportive data and to evaluate the robustness of efficacy findings. Safety analyses were conducted on the safety population.

Pre-defined subgroup efficacy analyses were performed to compare treatments within the stratification factors, and between demographic and baseline characteristics. Table 4 presents a summary and methodology for AUGMENT and planned subgroup analyses (Appendix M1 presents the full methodology).

Table 4: Summary of AUGMENT methodology

Trial Name	**AUGMENT**
Location	96 sites across 17 countries across North America, Europe, China and Brazil
Trial design	A multinational, randomized, double-blind, Phase III study Patients were randomized in a 1:1 ratio through an IVRS Randomization was stratified by previous rituximab treatment (yes, no), time since last antilymphoma therapy (≤2, >2 years) and disease histology (FL, MZL)
Eligibility criteria for participants	**Inclusion criteria** • Aged ≥18 years • Histologically confirmed MZL or Grade 1, 2, or 3a FL (CD20+ by flow cytometry or histochemistry) as assessed by investigator or local pathologist • Had to have been previously treated with at least one prior systemic chemotherapy, immunotherapy or R-chemo and had to have received at least two previous doses of rituximab: – Systemic therapy did not include local involved field radiotherapy for limited stage disease or *Helicobacter pylori* eradication – Prior investigational therapies were allowed provided the patient had received at least one prior systemic therapy • Had to have documented relapsed, refractory, or PD after treatment with systemic therapy, and not be R-refractory – Rituximab-refractoriness was defined as did not respond (at least a PR) to rituximab or R-chemo therapy and/or time to disease progression <6 months after last rituximab dose

Trial Name	AUGMENT
	– Rituximab-sensitive MZL or FL was defined as responded (at least a PR) to rituximab or R-chemo regimen therapy and time to disease progression ≥6 months after last rituximab dose • Must have needed treatment for relapsed, progressed, or refractory disease as assessed by the investigator • Performance status ≤2 on the ECOG scale **Exclusion criteria** • Life expectancy <6 months • Prior use of lenalidomide • Presence or history of central nervous system (CNS) involvement by lymphoma • Patients who were at a risk for a thromboembolic event and were not willing to take venous thromboembolism (VTE) prophylaxis
Settings and locations where the data were collected	An independent external DMC assessed ongoing safety throughout the study. The DMC conducted the planned interim futility analysis when an estimated 96 events per IRC review were reported. Response-related efficacy assessments were based on central review, including central radiology and clinical review by the IRC. Images received from investigators' sites were sent to the IRC, as well as relevant clinical information for haemato-oncology review.
Trial drugs	Lenalidomide 10 mg or 20 mg oral capsules[a] once daily on Days 1 to 21 of every 28-day Cycle up to 12 cycles combined with rituximab 375 mg/m^2 IV every week in Cycle 1 and on Day 1 of every 28-day Cycle from Cycles 2 through 5. Treatment continued until progression or unacceptable toxicity.
Permitted and disallowed medication	The following medications are prohibited during the study: • Systemic chronic corticosteroid at doses above 20 mg/day (prednisone/prednisolone or equivalent) during treatment phase. A seven-day washout period before Cycle 1 Day 1 study drug dosing was required for these patients • All investigational therapies (drug or otherwise) and anticancer therapies, other than lenalidomide or rituximab were prohibited during the entire Treatment Period of the study
Primary outcomes (including scoring methods and timings of assessments)	• PFS in relapsed/refractory indolent lymphoma patients, defined as the time from randomization to the first observation of disease progression, based on the modified 2007 IWGRC, or death due to any cause • Analysis was based on the IRC determination of disease progression
Other outcomes used in the economic model/specified in the scope	**Secondary endpoints** • To compare the safety of R^2 versus rituximab plus placebo • To compare the efficacy of R^2 versus rituximab plus placebo using other parameters of efficacy: – DCRR, ORR, CR rate, DOR, and DOCR by the 2007 IWGRC without PET – OS, EFS, and TTNLT

Trial Name	AUGMENT
	Exploratory endpoints • To compare the effects of R^2 versus R mono on: • TTNCT and RTNLT • CR/CRu rate in patients with FL based on the 1999 IWGRC • PFS on next antilymphoma treatment (PFS2) • HRQL as measured by the EORTC Quality of Life Questionnaire, Core 30 (QLQ-C30) and EuroQol Group's questionnaire 5 dimensions (EQ-5D-3L)
Pre-planned subgroups	Efficacy analyses were performed within a number of patient subgroups. These are described in Appendix M

Key: CR, complete response; CT, computerized tomography; DCRR, durable complete response rate; DMC, data monitoring committee; DOCR, duration of complete response; DOR, duration of response; ECOG, Eastern Cooperative Oncology Group; EFS, event-free survival; EORTC, European Organisation for Research and Treatment of Cancer, FL, follicular lymphoma; HRQL, health-related quality of life; IRC, Independent Review Committee; IVRS, interactive voice response system; IWGRC, International Working Group Response Criteria; MALT, mucosa-associated lymphoid tissue; MRI, magnetic resonance imaging; MZL, marginal zone lymphoma; ORR, overall response rate; PD, progressive disease; PR, partial response; R^2, rituximab plus lenalidomide; R-chemo, rituximab-containing chemotherapy; R mono, rituximab monotherapy; RTNLT, response rate to next antilymphoma treatment; TTNLT, time to next antilymphoma treatment; TTNCT, time to next chemotherapy treatment.
Notes: [a] dose modification rules allowed for dosing down to 2.5 mg with Celgene supplying lenalidomide 2.5 mg, 5 mg, 10 mg, 15 mg, and 20 mg capsules.
Source: Celgene, 2018.[66]

Baseline characteristics

Patient disposition for ITT population is presented in Appendix D4.

For the ITT population, baseline demographics between the R^2 arm and the R mono arm were generally similar. Overall, 261 patients (73%) had Ann Arbor Stage III to IV disease; 123 patients (34%) had a FLIPI score ≥3; and 183 patients (51%) had high tumour burden per GELF criteria. The majority of enrolled patients were at second-line (56%) and the remainder were at third-line (20%) or at fourth-line or greater (24%). A total of 117 patients (33%) were enrolled with POD24 status.

Of note, more patients in the R^2 arm than in the R mono arm were female (58% vs. 46%), aged ≥65 years (46% vs. 41%) had Ann Arbor Stage III to IV disease (77% vs. 69%), FLIPI score ≥3 (39% vs. 30%), had an ECOG score of 1 or 2 (35% vs. 29%) and were refractory to the last prior regimen (17% vs. 14%). The distribution of POD24 patients was similar between treatment arms. Overall, the patient baseline characteristics of the study population were reflective of a patient population with previously treated FL or MZL.

For FL patients, baseline disease characteristics were similar to the overall population; however, in patients with MZL, baseline disease characteristics were imbalanced, favouring the R mono arm (R^2 arm vs. R mono arm): ECOG 0 (55% vs 72%); Ann Arbor Stage III to IV disease (77% vs. 56%); Ann Arbor Stage IV (65% vs. 41%); FLIPI score ≥3 (48% vs. 25%); B symptoms (13% vs. 3%); elevated LDH (29% vs. 19%); and high tumour burden per GELF criteria (65% vs. 56%). Table 5 presents baseline demographic and disease characteristics for the ITT population.

Table 5: Baseline demographic and disease characteristics, AUGMENT – ITT population

	FL		MZL		Total		Overall (N=358)
	R² (N=147)	R mono (N=148)	R² (N=31)	R mono (N=32)	R² (N=178)	R mono (N=180)	
Male, n (%)	61 (41.5)	80 (54.1)	14 (45.2)	17 (53.1)	75 (42.1)	97 (53.9)	172 (48.0)
Median age, years (range)	62.0 (26.0-86.0)	61.0 (35.0-88.0)	68.0 (37.0-80.0)	66.0 (36.0-82.0)	64.0 (26.0-86.0)	62.0 (35.0-88.0)	62.5 (26.0-88.0)
Age distribution, n (%)							
<65	86 (58.5)	94 (63.5)	10 (32.3)	13 (40.6)	96 (53.9)	107 (59.4)	203 (56.7)
≥65	61 (41.5)	54 (36.5)	21 (67.7)	19 (59.4)	82 (46.1)	73 (40.6)	155 (43.3)
≥70	34 (23.1)	32 (21.6)	13 (41.9)	12 (37.5)	47 (26.4)	44 (24.4)	91 (25.4)
Race, white (%)	[redacted]	[redacted]	[redacted]	[redacted]	118 (66.3)	115 (63.9)	233 (65.1)
Histology (investigator review), n (%)							
FL	[redacted]	[redacted]	[redacted]	[redacted]	147 (82.6)	148 (82.2)	295 (82.4)
Grade 1	[redacted]	[redacted]	[redacted]	[redacted]	50 (28.1)	62 (34.4)	112 (31.3)
Grade 2	[redacted]	[redacted]	[redacted]	[redacted]	75 (42.1)	61 (33.9)	136 (38.0)
Grade 3a	[redacted]	[redacted]	[redacted]	[redacted]	22 (12.4)	25 (13.9)	47 (13.1)
MZL	N/A	N/A	31 (100.0)	32 (100.0)	31 (17.4)	32 (17.8)	63 (17.6)
MALT	N/A	N/A	14 (45.2)	16 (50.0)	14 (7.9)	16 (8.9)	30 (8.4)
Nodal	N/A	N/A	8 (25.8)	10 (31.3)	8 (4.5)	10 (5.6)	18 (5.0)
Splenic	N/A	N/A	9 (29.0)	6 (18.8)	9 (5.1)	6 (3.3)	15 (4.2)
Ann Arbor stage, n (%)							
I	13 (8.8)	13 (8.8)	2 (6.5)	5 (15.6)	15 (8.4)	18 (10.0)	33 (9.2)
II	21 (14.3)	29 (19.6)	5 (16.1)	9 (28.1)	26 (14.6)	38 (21.1)	64 (17.9)
III	69 (46.9)	60 (40.5)	4 (12.9)	5 (15.6)	73 (41.0)	65 (36.1)	138 (38.5)
IV	44 (29.9)	46 (31.1)	20 (64.5)	13 (40.6)	64 (36.0)	59 (32.8)	123 (34.4)
FLIPI category (derived), n (%)							

	FL		MZL		Total		Overall (N=358)
	R^2 (N=147)	R mono (N=148)	R^2 (N=31)	R mono (N=32)	R^2 (N=178)	R mono (N=180)	
Low (0,1)	[redacted]	[redacted]	[redacted]	[redacted]	52 (29.2)	67 (37.2)	119 (33.2)
Intermediate (2)	[redacted]	[redacted]	[redacted]	[redacted]	55 (30.9)	58 (32.2)	113 (31.6)
High (≥3)	[redacted]	[redacted]	[redacted]	[redacted]	69 (38.8)	54 (30.0)	123 (34.4)
Baseline ECOG score, n (%)							
0	99 (67.3)	105 (70.9)	17 (54.8)	23 (71.9)	116 (65.2)	128 (71.1)	244 (68.2)
1	47 (32.0)	42 (28.4)	13 (41.9)	8 (25.0)	60 (33.7)	50 (27.8)	110 (30.7)
2	[redacted]	[redacted]	[redacted]	[redacted]	2 (1.1)	2 (1.1)	4 (1.1)
LDH elevated, n (%)							
Yes	39 (26.5)	43 (29.1)	6 (19.4)	6 (18.8)	45 (25.3)	49 (27.2)	94 (26.3)
No	107 (72.8)	105 (70.9)	25 (80.6)	26 (81.3)	132 (74.2)	131 (72.8)	263 (73.5)
High tumour burden (GELF criteria)							
Yes	77 (52.4)	68 (45.9)	20 (64.5)	18 (56.3)	97 (54.5)	86 (47.8)	183 (51.1)
No	70 (47.6)	80 (54.1)	11 (35.5)	14 (43.8)	81 (45.5)	94 (52.2)	175 (48.9)
Prior antilymphoma regimens							
1	[redacted]	[redacted]	[redacted]	[redacted]	102 (57.3)	97 (53.9)	199 (55.6)
>1	[redacted]	[redacted]	[redacted]	[redacted]	76 (42.7)	83 (46.1)	159 (44.4)
Refractory to last prior regimen							
Yes	26 (17.7)	25 (16.9)	4 (12.9)	1 (3.1)	30 (16.9)	26 (14.4)	56 (15.6)
No	121 (82.3)	123 (83.1)	27 (87.1)	31 (96.9)	148 (83.1)	154 (85.6)	302 (84.4)
POD24[a], n (%)							
Yes	[redacted]	[redacted]	[redacted]	[redacted]	56 (31.5)	61 (33.9)	117 (32.7)
No	[redacted]	[redacted]	[redacted]	[redacted]	122 (68.5)	118 (65.6)	240 (67.0)

Key: ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; FLIPI, follicular lymphoma international prognostic index; GELF, Groupe d'Etude des Lymphomes Folliculaires; LDH, lactate dehydrogenase; MZL, marginal zone lymphoma; MALT, mucosa associated lymphatic tissue; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo. **Notes**: [a], POD24 is defined as relapse within two years of initial chemoimmunotherapy. **Source**: Celgene, 2018.[66]

MAGNIFY

Study design

MAGNIFY is an ongoing multicentre, open-label, randomized, Phase IIIb study comparing the efficacy and safety of R^2 maintenance therapy with rituximab monotherapy maintenance therapy, following a 12-cycle induction therapy with R^2. The study is being conducted across 114 sites in three countries. MAGNIFY is one of the key studies used in EMA regulatory submission.

To be eligible for inclusion in the study, patients had to be aged ≥18 years with histologically confirmed Grade 1, 2, 3a or 3b FL, transformed FL, MZL or mantle cell lymphoma (MCL) (WHO 2008 classification)[68]. Patients were required to have received at least one prior treatment for lymphoma including radiotherapy, chemotherapy, immunotherapy, chemoimmunotherapy, or novel agent. Furthermore, patients had to have documented relapsed/refractory or PD after their last treatment; and R-refractory patients were eligible for the study. Table 6 summarizes inclusion and exclusion criteria; see Appendix M2 for full details.

Following a 28-day screening period, patients entered a 12-cycle induction treatment period during which they received R^2. Patients received oral lenalidomide at a dose of 10 mg or 20 mg (based on CrCl) once daily on Days 1 to 21 in each 28-day cycle, combined with four-weekly infusions of rituximab by IV at a dose of 375 mg/m^2, followed by five additional doses on Day 1 of Cycles 3, 5, 7, 9, and 11. Patients were stratified by lines of antilymphoma therapy (≤2 vs. >2 lines), and histology (FL Grades 1–3b and transformed FL [tFL] vs. MZL vs. MCL) and age (65 vs. ≥65 years).

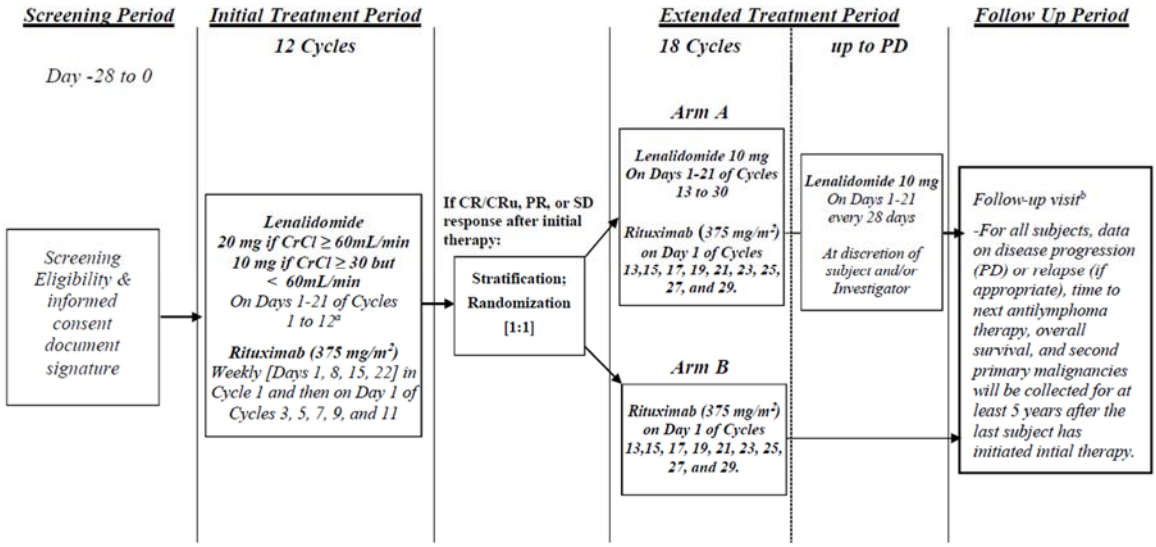
Of note, the induction phase dosing schedule of R^2 used in MAGNIFY was similar to but not identical to AUGMENT, with a subtle difference in rituximab scheduling that literature and clinical expert opinion confirm is not expected to have a meaningful impact on the therapeutic benefit of R^2. [69-73] [1]

Following the induction phase, patients who had SD, partial response (PR), CR or complete response unconfirmed (CRu) at the end of 12 cycles of initial therapy were

randomized 1:1 to receive R^2 maintenance therapy or rituximab maintenance therapy.

Figure 4 presents a schematic of the MAGNIFY study.

Figure 4: Study design for MAGNIFY



Key: CR, complete response; CrCl, creatinine clearance; CRu, complete response unconfirmed; IVRS, interactive voice response system; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: a, treatment must begin no later than one week after completing enrolment in IVRS to receive initial therapy. 1 cycle = 28 days. b, all enrolled patients are followed for disease progression. This includes patients who discontinue the protocol-specified treatment or the study early for any reason without documented evidence of PD or relapse.
Source: Celgene, 2017.[74]

The primary endpoint for the extended treatment period (final analysis) is PFS, which will be assessed using a modification of the 1999 IWGRC (i.e. allowing the inclusion of extranodal disease as measurable disease). The study is ongoing and the primary analysis is yet to be conducted.

This current submission presents data from interim analyses. Three analysis populations are presented using data from the interim and safety data cuts:

- The induction efficacy evaluable (IEE) population, defined as all patients in the induction intent-to-treat (IITT) population who received ≥1 dose of initial therapy, who had baseline and at least one post-baseline efficacy assessment (including patients who died or progressed before first on-study assessment)

- The IITT population, defined as all enrolled patients with FL Grade 1 to 3a or MZL who met all the eligibility criteria for the study
- The induction safety population, defined as patients with FL Grade 1 to 3a or MZL who had received ≥1 dose of initial therapy (either lenalidomide or rituximab)

For the induction treatment period (interim analysis), the primary endpoint was ORR as observed for the IEE population, assessed using the modified 1999 IWGRC. CR rate for the IEE population was considered a secondary endpoint; however, for the purpose of this submission we present it alongside the breakdown of ORR. Other secondary endpoints included response rate analyses performed on the IITT population, with the addition of DOR, DOCR and time to response (TTR). Exploratory endpoints included PFS (analysed in the induction safety population). Safety analyses were also carried out in the induction safety population.

For this submission, PFS and OS data have been censored at 12 months because survival outcomes beyond the 12-month induction therapy phase become increasingly confounded by maintenance therapy. It should also be noted that response durations (e.g. DOR, DOCR) are potentially confounded by maintenance and therefore not discussed in detail in this submission.

Table 6 summarizes the methodology for MAGNIFY; the full methodology is presented in Appendix M2.

Table 6: Summary of MAGNIFY methodology

Trial Name	**MAGNIFY**
Location	114 sites in 3 countries including the US, Germany and Puerto Rico.
Trial design	A multinational, randomized, open-label, Phase IIIb study Following an initial treatment period where patients received 12 cycles of R^2, patients were randomized in a 1:1 ratio through an IVRS to R^2 or R-mono. Randomization was stratified by: • Histology (Grade 1-3b FL, transformed FL, MZL, MCL) • Lines of antilymphoma therapy (≤2 vs. >2 lines) • Age (<65 vs. ≥65 years)
Eligibility criteria for participants	**Inclusion criteria** • Aged ≥18 years • Histologically confirmed FL (Grade 1, 2, 3a, or 3b), transformed FL, MZL, or MCL (WHO 2008 classification)

Trial Name	MAGNIFY
	• Must have been previously treated with at least one prior treatment for lymphoma including radiotherapy, chemotherapy, immunotherapy, chemoimmunotherapy, or novel agent • Must have had documented relapsed, refractory or PD after last treatment – Patients who were R-refractory (defined as a patient who experienced a best response of PD or SD to treatment with rituximab or a rituximab-containing regimen OR a response (PR or CR) lasting fewer than six months following the last rituximab dose) were eligible • Must have needed treatment for relapsed or refractory disease as assessed by the investigator • Performance status <2 on the Eastern Cooperative Oncology Group (ECOG) scale **Exclusion criteria** • Presence or history of central nervous system involvement by lymphoma • Life expectancy <6 months • Prior use of lenalidomide
Settings and locations where the data were collected	As of the 10 August 2018 database cut-off, 370 patients were enrolled across 114 sites in three countries as follows: • 99 clinical sites in the US • 14 clinical sites in Germany • 1 clinical site in Puerto Rico
Trial drugs	**Initial treatment period (Induction phase)** Lenalidomide 20 mg (10 mg if creatine clearance [CrCl] ≥30 mL/min but <60 mL/min) oral capsules once daily on Days 1 to 21 of every 28-day Cycle up to 12 cycles combined with rituximab 375 mg/m^2 IV every week in Cycle 1 and on Day 1 of every other 28-day Cycle from Cycles 3 through 11.
Permitted and prohibited medication	The following medications are prohibited during the study: • Systemic chronic corticosteroid at doses above 20 mg/day (prednisone/prednisolone or equivalent) during the treatment phase. A seven-day washout period before Cycle 1, Day 1 study drug dosing was required for these patients • All investigational therapies (drug or otherwise) and anticancer therapies, other than lenalidomide or rituximab were prohibited during the entire Treatment Period of the study
Primary outcomes (including scoring methods and timings of assessments)	**Interim analysis** Evaluate the efficacy of the initial treatment period of the study, where patients receive 12 cycles of R^2. Efficacy determination based on ORR by best response is the primary endpoint, using a modification of the 1999 IWGRC as determined for the IEE population.

Trial Name	MAGNIFY
Other outcomes used in the economic model/specified in the scope	**Interim analysis** Secondary outcomes: • Evaluate the safety of 12 cycles of initial therapy with R[2] • Evaluate other parameters of efficacy, such as CR rate, DOR, DOCR and TTR in the IITT population Exploratory outcomes: • Further evaluate efficacy of initial treatment, including PFS
Pre-planned subgroups	A number of subgroups were analysed for primary endpoint ORR and secondary endpoint CR. These are described in Appendix M

Key: CR, complete response; DOCR, duration of complete response; DOR, duration of response; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; IEE, induction efficacy population; IITT, induction intention-to-treat population; MCL, mantle cell lymphoma; MZL, marginal zone lymphoma; MALT, mucosa associated lymphatic tissue; ORR, overall response rate; PD, progressive disease; SD, stable disease; TTR, time to response.
Source: Celgene, 2017[74]; Celgene, 2019.[75]

Baseline characteristics

In the IEE, IITT and induction safety populations, there were 310, 370 and 359 patients, respectively. As discussed previously, the IEE population was considered the primary population of interest; therefore, the patient characteristics of this population are discussed in this section. Patient disposition for the IEE population is presented in Appendix D4.

Table 7 presents the baseline demographic and disease characteristics for the IEE population. In the IEE population, the majority of patients ([redacted]) were Stage IV at enrolment. There was approximately a 1:1 male to female ratio ([redacted]). Of the 310 patients, there were [redacted] FL patients and [redacted] MZL patients (of which [redacted] had nodal MZL). Of note, baseline disease characteristics were broadly similar between FL and MZL patients.

Overall, the baseline characteristics of patients in MAGNIFY represent a population with a poorer prognosis compared to the overall AUGMENT population, specifically for the following factors: median age ([redacted] vs. 62.5 years), Ann Arbor Stage III–IV ([redacted] vs. 73%), ECOG 0 ([redacted] vs. 68%), POD24 status ([redacted] vs. 33%). Also note that the median number of previous treatments for patients in MAGNIFY was greater than those in AUGMENT (2 vs. 1, respectively). Furthermore, [redacted] of the enrolled patients in MAGNIFY were R-refractory, the occurrence of which was similar

between FL and MZL patients ([redacted] respectively); however, for AUGMENT, all patients were non-R-refractory.

Demographic baseline characteristics for the IEE population (including distributions between FL and MZL patients) were similar to those reported for the IITT population and the induction safety population (See Appendix M2).

Table 7: Baseline demographic and disease characteristics, MAGNIFY – IEE population

	FL (N=247)	MZL (N=63)	Total (N=310)
Male, n (%)			
Median age, years (range)			
Age distribution, n (%)			
<65			
≥65			
Race, white, n (%)			
Histology (investigator review), n (%)			
FL			
Grade 1			
Grade 2			
Grade 3a			
MZL			
MALT (non-gastric)			
Gastric MALT			
Nodal			
Splenic			
Ann Arbor stage at enrolment, n (%)			
I			
II			
III			
IV			
ECOG score at enrolment, n (%)			
0			
1			

	FL (N=247)	MZL (N=63)	Total (N=310)
2			
POD24[a], n (%)			
Yes			
No			
Prior rituximab-containing therapy, n (%)			
Yes			
Prior rituximab-containing combination chemotherapy			
Prior rituximab monotherapy			
No			
R-refractory, n (%)			
Yes			
No			
Chemoresistant, n (%)			
Yes			
No			
Chemotherapy eligible, n (%)			
Yes			
No			
High tumour burden, n (%)			
Yes			
No			
Baseline bulky disease status[d], n (%)			
Yes			
No			

	FL (N=247)	MZL (N=63)	Total (N=310)
Missing			

Key: ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; IEE, induction efficacy evaluable; MZL, marginal zone lymphoma; MALT, mucosa associated lymphatic tissue; R, rituximab.
Notes: [a], POD24 is defined as relapse within two years of initial chemoimmunotherapy; [b], for 29 of these patients, this was their only prior systemic regimen; [c], for 17 of these patients, this was their only prior systemic regimen; [d], bulky disease is defined as a nodal or extranodal (except spleen) mass >7 cm in greater diameter or involvement of at least three nodal or extra-nodal sites (each with a diameter >3 cm).
Source: Celgene, 2018.[76]

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

Table 8 presents the hypothesis and associated statistical analysis methods adopted in AUGMENT and MAGNIFY.

AUGMENT

There were three analysis populations in the AUGMENT study, including the ITT, mITT and safety population (see Section B.2.3 for population definitions).

Primary efficacy analyses were conducted in the ITT population using Food and Drug Administration (FDA) censoring rules. The primary efficacy analyses were assessed by the IRC using a modified version of the 2007 IWGRC. Sensitivity efficacy analyses conducted in the ITT population (using EMA censoring rules) and mITT population were evaluated as supportive evidence and to assess the robustness of efficacy findings. Safety assessments for the study were conducted on the safety population.

MAGNIFY

There were four analysis populations in the MAGNIFY study, including IITT, IEE, induction safety and SPM safety population (see Section B.2.3 for population definitions).

All efficacy analyses were determined from investigator assessments using the modified 1999 IWGRC. Primary and secondary efficacy analyses for the interim analysis were conducted in both the IEE population and the IITT population (see Section B.2.3). Exploratory efficacy analyses were conducted in the induction safety population using both FDA and EMA censoring rules. Additionally, the data for MAGNIFY has been censored at 1-year (at the end of the initial treatment period), to eliminate the influence of post-randomization maintenance in this dataset.

Table 8: Summary of statistical analyses

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
AUGMENT	The primary objective of the study was to compare the efficacy of R^2 to R mono. Efficacy determination was based on PFS as the primary endpoint. The AUGMENT study was considered positive if the R^2 group was significantly superior to the rituximab group for the primary endpoint.	The analysis of the primary endpoint was planned when approximately 193 IRC-assessed PFS events were reached. The cut-off date for database lock was prespecified before database lock. KM estimates of PFS were provided, and the KM product limit method was used to estimate the survivorship function for PFS. Event rates at specific time points were estimated from KM curves. Medians together with two-sided 95% CIs were provided. The resulting PFS estimates were presented graphically.	Based on the rate of accrual anticipated in this study and 5% annual dropout rate, it was estimated that approximately 350 patients would be randomized in a 1:1 ratio to the two treatment arms and that PFS would be reached at 43 months. The basis for the power and sample size determination was a test of the equality of the overall time-to-event (i.e. PFS) curves between experimental and control treatment groups using a stratified log-rank test.	**EMA censoring rules** Event: • Death before first PD assessment while on study • Death between adequate assessment visits • All progressions and deaths, regardless of whether they occurred after next antilymphoma therapy or after ≥2 missed scheduled assessments Censored: • Patients with no baseline assessment were censored at randomization • Patients who did not progress or die and those that discontinued for any reason other than death or progression will be censored on the date of their last adequate assessment with

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
				evidence of no progression • Patients who died or progressed after more than one missed visit will be censored at the date of their last adequate assessment that revealed no progression
MAGNIFY	The study's objective is to demonstrate the efficacy of R^2 extended therapy followed by the optional lenalidomide single-agent extended versus the rituximab single-agent extended. For the final analysis, efficacy will be based on PFS as the primary endpoint. For the interim analysis, efficacy was based on ORR as the primary endpoint. The MAGNIFY study will be considered positive if the R^2 group is significantly superior to the rituximab group for the primary endpoint	The primary efficacy endpoint for the interim analysis presented here is ORR of the Initial Treatment Period (induction phase), as observed for the IEE population. Analysis of PFS is not presented as the primary endpoint for this interim analysis because the data have not reached maturity; however, PFS from first dose date of initial treatment are presented in this submission as an exploratory endpoint. It should be noted that the data for MAGNIFY has been censored at 1-year (at the end of the initial treatment period), to eliminate the influence of post-	Based on the rate of accrual anticipated in this study, and an annual 5% dropout rate, it is planned to enrol a total of 500 patients into the Initial Treatment Period. It is projected that approximately 314 patients will be randomized at a ratio of 1:1 into the two maintenance treatment arms based on an estimated response rate of the initial therapy. It is expected that the 191 PFS events that occur after the first dose date of extended therapy will be available in about 57 months from the beginning of the extended therapy, and the total study duration to reach the primary efficacy endpoint from the beginning of the Initial	**EMA censoring rules** Event: • Death • Death or progression between adequate assessment visits Censored: • Patients with no baseline tumour assessment will be censored at the informed consent date • Patients who did not progress or die will be censored on the date of their last adequate assessment with evidence of no progression

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		randomization maintenance therapy on this dataset	Treatment Period will be about 69 months	

Key: CI, confidence interval; EMA, European Medicines Agency; IEE, induction efficacy population; IRC, Independent Review Committee; ITT, intention-to-treat; KM, Kaplan–Meier; ORR, overall response rate; PD, progressive disease; PFS, progression-free survival; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo.
Source: Celgene, 2018[66]; Celgene, 2017.[74]

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

Table 9 provides a summary of the quality assessment with full details explained in Appendix D5. The AUGMENT and MAGNIFY studies were conducted under the ethical principles of Good Clinical Practice (GCP), according to the International Council for Harmonisation (ICH) Harmonized Tripartite Guideline, and with the ethical principles outlined in the Declaration of Helsinki. The accuracy and reliability of the clinical study data were assured by qualified investigators and an appropriate study centre, review of protocol procedures with the investigator and associated personnel before the study commenced. In addition, an independent external Data Monitoring Committee (DMC) assessed ongoing safety throughout the study and conducted the planned interim futility analysis.

AUGMENT

Randomization was successfully carried out in the AUGMENT study furnishing baseline demographics and disease characteristics that were generally well-balanced between treatment arms and reflective of a patient population with previously treated FL or MZL. The most common reason for withdrawal from the study was disease progression, and the numbers and reasons were similar between treatment groups. Patient withdrawals were accounted for within the efficacy assessments; for reasons other than disease, progression was accounted for with standard censoring methods. Patients, carers and investigators remained blinded throughout the study, and all outcome assessments were conducted in accordance with trial-validated methodology and based on the ITT principle. Subsequently, this double-blind randomization method ensured low levels of bias in the AUGMENT study.

MAGNIFY

The most common reason for withdrawal from MAGNIFY was disease progression, which was accounted for within the efficacy assessments; for reasons other than disease, progression was accounted for with standard censoring methods. Although this was designed as an open-label study, the efficacy endpoints are not subjectively assessed endpoints; therefore, a lack of blinding was not thought to have a

considerable effect on the outcome of the study. Furthermore, the results of interest for this submission are taken from the initial treatment period only and are therefore not affected by the open-label design.

Table 9: Quality assessment results for AUGMENT and MAGNIFY

Trial number (acronym)	AUGMENT	MAGNIFY (induction phase)
Was randomization carried out appropriately?	Yes	N/A
Was the concealment of treatment allocation adequate?	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	N/A
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	No
Were there any unexpected imbalances in drop-outs between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes
Source: Celgene, 2018[66]; Celgene, 2017.[74]		

B.2.6. Clinical effectiveness results of the relevant trials

AUGMENT

The data presented in this section are based on the 22 June 2018 data cut-off for the primary analysis. Efficacy analyses were conducted in the ITT population and based on data from IRC review, using the modified 2007 IWGRC. EMA censoring rules were applied to the analyses.

Primary endpoint: Progression-free survival

In total, there were 185 IRC-assessed events, which remained unchanged following the application of EMA censoring. At a median follow-up of 28.3 months, the R^2 arm

demonstrated superiority over the R mono arm with a [redacted] ($p<0.0001$). The median PFS was greater for the R^2 arm (39.4 months) compared with the R mono arm (14.1 months). Furthermore, the PFS rate was greater for the R^2 arm at both 1 year ([redacted] and 2 years ([redacted] compared with the R mono arm.

For FL patients, PFS improvements were consistent with those of the ITT population ([redacted]); however, for MZL patients, PFS was not significantly different between the R^2 and R mono arms (24.9 vs. 25.2 months), with a [redacted]). It is important to note that PFS outcomes in the MZL subgroup are difficult to interpret because of the small sample size (31 patients in the R^2 arm and 32 patients in the R mono arm) and imbalance in baseline prognostic factors (as discussed in Section B.2.3). Specifically, Ann Arbor Stage IV, elevated LDH, and "unfit for chemotherapy" were identified as significant prognostic factors ($p<0.05$). Multivariate analyses adjusting for the imbalance in these 3 significant prognostic factors in the MZL subgroup showed an adjusted PFS HR of [redacted]) in favour of the R^2 arm; this HR was similar to the PFS HR in the overall population ([redacted]
A summary of PFS, based on unadjusted analysis, is presented in Table 10.

In a post hoc analysis, among patients randomized to the R^2 arm, the median PFS was longer than that observed following the previous antilymphoma treatment (39.4 vs. 32.4 months), whilst the median PFS among patients in the R mono arm was shorter than that from the previous antilymphoma treatment (14.1 vs. 30.6 months). The observation in the control arm was consistent with established expectations that PFS will decrease with each successive treatment based on those currently available;[25] however, the longer PFS for patients treated with R^2 compared with that achieved in response to their previous treatment reverses that trend. This provides evidence for a clinically significant contribution from the novel mechanism of action of this combination that differs to currently available options.

Table 10: Progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in AUGMENT: ITT population

	FL		MZL		Overall	
	R^2 (N=147)	R mono (N=148)	R^2 (N=31)	R mono (N=32)	R^2 (N=178)	R mono (N=180)
Number of patients, n (%)						
With event						
Censored						
Median PFS (95% CI) (months)[a]						
PFS rate at 6 months (95% CI)						
PFS rate at 1 year (95% CI)						
PFS rate at 2 years (95% CI)						
p-value						
Hazard ratio (95% CI)						

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IRC, Independent Review Committee; ITT, intent-to-treat; IWGRC, International Working Group Response Criteria; MZL, marginal zone lymphoma; NE, not estimable; PFS, progression-free survival; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo.
Notes: a, median estimate is from Kaplan–Meier analysis; b, p-value from log-rank test stratified by three factors: previous rituximab treatment (yes, no), time since last antilymphoma therapy (≤2, >2 years), and disease histology (FL, MZL); c, from Cox proportional hazard model adjusting for the three stratification factors; d, p-value from log-rank test; e, from Cox proportional hazard model.
Source: Celgene, 2018.[77].

PFS Kaplan–Meier (KM) curves separated early, starting at the first on-study tumour assessment visit and persisted throughout the follow-up period beyond the 1-year study treatment duration. Figure 5 presents a KM curve of PFS for the ITT population.

Figure 5: Kaplan–Meier curve of progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in AUGMENT: ITT population

Notes: [a], p-value from stratified log-rank test. Stratification factors include the following: previous rituximab treatment (yes, no), time since last antilymphoma therapy (≤2, >2 years), and disease histology (FL, MZL). Hazard ratio and its CI were estimated from Cox proportional hazard model adjusting for the stratification factors.
Key: CI, confidence interval; EMA, European Medicines Agency; IRC, Independent Review Committee; ITT, intention-to-treat; IWGRC, International Working Group Response Criteria; KM, Kaplan–Meier; Len, lenalidomide; NE, not estimable; Pbo, placebo; Rit, rituximab.
Source: Celgene, 2018.[77]

Figure 6 and Figure 7, respectively, present the KM curves of PFS for FL and MZL patients in the ITT population.

Figure 6: Kaplan–Meier curve of progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in FL patients in AUGMENT: ITT population

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IRC, Independent Review Committee; IWGRC, International Working Group Criteria; ITT, intention-to-treat; KM, Kaplan–Meier; Len, lenalidomide; NE, not estimable; Pbo, placebo; Rit, rituximab.
Source: Celgene, 2018.[77]

Figure 7: Kaplan–Meier Curve of progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in MZL patients in AUGMENT: ITT population

Key: CI, confidence interval; EMA, European Medicines Agency; IRC, Independent Review Committee; ITT, intention-to-treat; IWGRC, International Working Group Criteria; KM, Kaplan–Meier; Len, lenalidomide; MZL, marginal zone lymphoma; NE, not estimable; Pbo, placebo; Rit, rituximab.
Source: Celgene, 2018.[77]

Secondary endpoints

The secondary efficacy endpoints for AUGMENT included OS, response rates (ORR, CR rate, DOR, DOCR and DCRR), EFS and TTNLT. Table 11 presents a summary of the secondary endpoints, with the exception of DOR, DOCR and DCRR, which are presented in Appendix O1.

Table 11: Summary of secondary endpoints in AUGMENT: ITT population

Endpoint	Overall	
	R^2 (N=178)	**R mono (N=180)**
Median OS, months (95% CI)[a]	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.61 (0.33, 1.13)[b]	
Best response, n (%)		
ORR (CR+PR)	138 (77.5)	96 (53.3)
95% CI[d]	70.7, 83.4	45.8, 60.8

p-value	<0.0001[e]	
CR rate	60 (33.7)	33 (18.3)
95% CI[d]	26.8, 41.2	13.0, 24.8
p-value	0.001[e]	
PR	78 (43.8)	63 (35.0)
SD	20 (11.2)	55 (30.6)
PD/ death	7 (3.9)	23 (12.8)
No evidence of disease	3 (1.7)	4 (2.2)
Unknown/ND/Missing	10 (5.6)	2 (1.1)
Median TTNLT, months (95% CI)[a]	NE (NE, NE)	32.2 (23.2, NE)
TTNLT rate at 2 years, % (95% CI)	73.6 (65.6, 80.1)	57.3 (49.3, 64.5)
Hazard ratio (95% CI)	0.54 (0.38, 0.78)[b]	
p-value	0.0007[g]	
Median EFS, months (95% CI)[a]	27.6 (22.1, NE)	13.9 (11.4, 16.7)
Hazard ratio (95% CI)	0.51 (0.38 to 0.67)[b]	
p-value	<0.0001[g]	

Key: CI, confidence interval; CR, complete response; EFS, event-free survival; FL, follicular lymphoma; IRC, Independent Review Committee; ITT, intent-to-treat; MZL, marginal zone lymphoma; ND, not done; NE, not estimable; ORR, overall response rate; OS, overall survival; PD, progressive disease; PR, partial response; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo; SD, stable disease; TTNLT, time to next antilymphoma treatment.
Notes: [a], median estimate is from Kaplan–Meier analysis; b, from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last antilymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c], from Cox proportional hazard model; [d], exact confidence interval for binomial distribution; [e], from CMH test adjusting for the three stratification factors; [f], from Fisher-Exact test; [g], from log-rank test adjusting for the three stratification factors; [h], from log-rank test; [i], exact confidence interval for binomial distribution.
Source: Leonard, et al. 2019; Celgene, 2018.[58, 77]

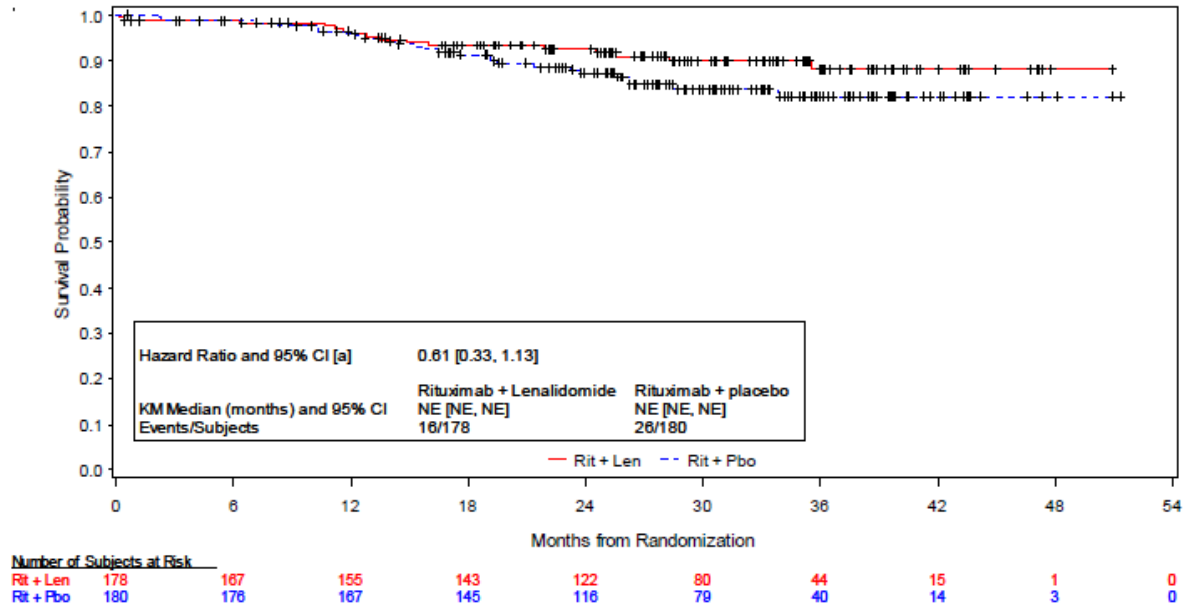
Overall survival

Overall survival data is relatively immature, with few events at time of analysis. In total, there were 16 deaths in the R^2 arm and 26 deaths in the R mono arm (median follow-up 28.3 months). Overall, there was a relative reduction of 39% in the risk of death (HR [95% CI]: 0.61 [0.33, 1.13]) for patients treated with R^2 compared with those treated with R mono. For both treatment arms, median OS was not estimable.

Both KM curves overlapped up to 1-year post-randomization (Figure 8); and the OS rate at 1-year was similar between the two treatment arms (95.8% vs. 96.0%, respectively). At 2 years, the OS rate was greater in the R^2 arm (92.6%) compared with the R mono arm (85.8%).

A KM curve of OS for the ITT population is presented in Figure 8.

Figure 8: Kaplan–Meier curve of overall survival in AUGMENT: ITT population



Key: CI, confidence interval; FL, follicular lymphoma; ITT, intention-to-treat; KM, Kaplan–Meier; Len, lenalidomide; MZL, marginal zone lymphoma; NE, not estimable; Pbo, placebo; Rit, rituximab.
Notes: [a], Hazard ratio and CI from Cox model adjusting for the stratification factors: previous rituximab treatment (yes, no), time since last antilymphoma therapy (≤2, >2 years), and disease histology (FL, MZL).
Source: Celgene, 2018.[77]

Response rates

In the ITT population, ORR was significantly greater for the R^2 arm than the R mono arm (78% vs. 53%; p<0.0001). The CR rate was also greater for the R^2 arm compared with the R mono arm (34% vs. 18%; p=0.001). Furthermore, of the patients who achieved a response (CR+PR) and a CR, the median time to response was similar between the R^2 and R mono treatment arms ([redacted]; and [redacted], respectively).

Time to next antilymphoma treatment

TTNLT offers a clinical endpoint with additional value beyond PFS when considering the applicability of clinical study data to actual clinical practice in this disease setting. PFS can be hard to interpret due to discrepant approaches to determining disease progression. Progression is proactively evaluated on-study through scheduled investigations, potentially prior to symptom development, while progression is more

likely to be discovered reactively in real-world practice after a patient presents with symptoms. Subsequent treatments, however, are generally initiated only following symptomatic progression, making TTNLT an endpoint more readily comparable between clinical studies and real-world practice. The importance of the TTNLT endpoint in this indolent disease was highlighted by the committee for TA513.[52] This concept is discussed further in Section B.2.13.

In the ITT population, 49 patients (28%) in the R^2 arm and 80 patients (44%) in the R mono arm received subsequent antilymphoma treatment.

For the R^2 arm, median TTNLT was not estimable; however, for the R mono arm, the median TTNLT was 32.2 months (HR [95% CI] 0.54: 0.38, 0.78; p=0.0007). Furthermore, at 2 years, 74% of patients in the R^2 arm and 57% of patients in the R mono arm had not received subsequent antilymphoma treatment.

An exploratory analysis of response rate to next antilymphoma treatment (RTNLT) is reported below.

Event-free survival

For EFS, an event was defined as documented progression, relapse, initiation of new antilymphoma treatment or death. In the ITT population, patients in the R^2 arm were 49% less likely to experience an event compared with those in the R mono arm (HR [95%]: 0.51 [0.38, 0.67]; p<0.0001). Furthermore, at 1 year, the EFS rate was greater in the R^2 arm ([redacted]) compared with the R mono arm ([redacted]).

Exploratory endpoints

A summary of the exploratory endpoints, including RTNLT, time to next chemotherapy treatment (TTNCT), are presented in Table 12. CR rate in FL patients (based on the 1999 WGRC), PFS on next antilymphoma treatment (PFS2) and patients with HT are presented in Appendix O1.

A greater proportion of patients receiving R^2 responded to subsequent treatment compared with patients receiving R mono (ORR: 57% vs. 36%; CR: 31% vs. 16%), lending weight to the hypothesis that the combination may re-sensitise patients to subsequent therapy through immune function enhancement.[3] Indeed, TTNCT was significantly improved for R^2 compared with R mono (p=0.0017); however, median

TTNCT was not estimable for either treatment arm. Furthermore, the TTNCT rate at 2 years was higher for the R^2 arm ([redacted]) than for the R mono arm ([redacted]).

Table 12: Summary of exploratory endpoints in AUGMENT: ITT population

Endpoint	Overall	
	R^2 (N=178)	**R mono (N=180)**
Median TTNCT, months (95% CI)[a]	NE (NE, NE)	NE (NE, NE)
TTNCT rate at 2 years, % (95% CI)	84.9 (77.2, 90.1)	66.5 (57.8, 73.8)
Hazard ratio (95% CI)	0.50 (0.32, 0.78)[b]	
p-value	0.0017[c]	
RTNLT		
ORR, n (% [95% CI][d])	28 (57.1 [42.2, 71.2])	29 (36.3 [25.8, 47.8])
p-value	0.0282[e]	
CR, n (% [95% CI][d])	15 (30.6 [18.3, 45.4])	13 (16.3 [8.9, 26.2])
p-value	0.0775[e]	

Key: CI, confidence interval; ITT, intention-to-treat; NE, not estimable; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo; RTNLT, response rate to next antilymphoma treatment; TTNCT, time to next antilymphoma chemotherapy treatment.
Notes: [a], median estimate is from Kaplan–Meier analysis; [b], from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes, no), time since last antilymphoma therapy (≤2, >2 year), and disease histology (FL, MZL); [c], p-value from log-rank test adjusted by the three stratification factors; [d], exact confidence interval for binomial distribution; [e], p-value obtained from Fisher-Exact test; [f], p-value obtained from CMH test adjusting for the three stratification factors; [g], 95% CI is based on the Clopper-Pearson exact method.
Source: Leonard, et al. 2019[58]; Celgene, 2018.[77]

Health-related quality of life

HRQL was assessed using the EORTC QLQ-C30 and EuroQol Five Dimension Three Level (EQ-5D-3L) questionnaire. The global health status/quality of life (GHS/QoL) domain of the QLQ-C30 was chosen as the primary patient reported outcome of interest. The remaining domains of the QLQ-C30 and the EQ-5D-3L were assessed as exploratory outcomes of interest.[78]

The following key endpoints were evaluated in the primary analyses to assess the impact of R^2 versus R mono on each HRQL domain of interest:

- Within- and between-group difference in mean change from baseline score at each post-baseline assessment visit (i.e. a cross-sectional analysis)

- Within- and between-group difference in the least-squares (LS) mean change at a pre-specified post-baseline assessment timepoint (i.e. a longitudinal analysis)
- The proportion of patients experiencing clinically meaningful deterioration at each post-baseline assessment visit
- Time to first clinically meaningful deterioration

Appendix P presents the results of the HRQL analysis. Overall, the results of the HRQL data analyses indicate there was no clinically meaningful difference in GHS/QoL or most of the exploratory domains between treatment groups. This analysis suggests HRQL (which was at the level equivalent to the general population at baseline) was generally uncompromised by the addition of lenalidomide to rituximab in patients with previously-treated FL or MZL. As such, R^2 is associated with more durable remissions and delayed progression (discussed in Section B.2.12), rather than substantially altering moment to moment HRQL.

MAGNIFY

The data discussed in this section have been taken from the interim analysis (data cut-off 10 August 2018). Primary efficacy analyses were conducted in the IEE population. Efficacy analyses conducted in the IITT population (secondary endpoints) were not considered of primary importance and are presented in Appendix O. Exploratory efficacy analyses, such as PFS, were conducted in the induction safety population.

Primary endpoint (interim analysis): Overall response rate

Response rates were based on best response and assessed using the modified 1999 IWGRC. In the IEE population, the ORR was 73%. For FL and MZL patients, the ORR was 75% and 65%, respectively. The CR rate (secondary endpoint) in the IEE population was 45% and was similar between FL and MZL patients (46% vs. 38%, respectively).

A summary of response rates by best response for the IEE population in the induction phase is presented in Table 13.

Table 13: Response rate by best response per 1999 IWGRC in the induction phase in MAGNIFY: IEE population

Best response in induction phase	FL (N=247)	MZL (N=63)	Overall (N=310)
Number of patients (%)			
ORR (CR+CRu+PR), n (% [95% CI[a]])	184 (74.5 [68.6, 79.8])	41 (65.1 [52.0, 76.7])	225 (72.6 [67.3, 77.5])
CR	[redacted]	[redacted]	[redacted]
CRu	[redacted]	[redacted]	[redacted]
CR rate (CR+CRu), n (% [95% CI][a])	114 (46.2 [39.8, 52.6])	24 (38.1 [26.1, 51.2])	138 (44.5 [38.9, 50.2])
PR	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]
Death w/o tumour assessment	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; FL, follicular lymphoma; IEE, induction efficacy evaluable; MZL, marginal zone lymphoma; ORR, overall response rate; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: [a], 95% CI based on the Clopper–Pearson exact method.
Source: Celgene, 2018.[59]

Secondary endpoints (interim analysis)

A summary of the secondary efficacy analyses conducted in the IITT population for MAGNIFY, including CR rate, DOR, DOCR and TTR are presented in Appendix O2.

Exploratory endpoint (interim analysis): Progression-free survival

PFS analyses were conducted in the induction safety population, and were based on investigator assessment, using the modified 1999 IWGRC. As with AUGMENT, EMA censoring rules were applied to the analyses.

In total, there were 103 investigator-assessed events. The PFS rate at 1 year was [redacted] A summary of PFS is presented in Table 14 and the KM curve is presented in Figure 9.

Table 14: Progression-free survival by investigator assessment per 1999 IWGRC with censoring rules based on EMA guidance in MAGNIFY: Induction safety population

	Total (N=359)
With event, n (%)	[redacted]
Censored, n (%)	[redacted]
PFS rate at 1 year (95% CI)	[redacted]

Key: CI, confidence interval; EMA, European Medicines Agency; IWGRC, International Working Group Response Criteria; PFS, progression-free survival.
Notes: [a], Statistics obtained from Kaplan–Meier method. 95% CI is based on Greenwood formula.
Source: Celgene, 2018.[79]

Figure 9: Kaplan–Meier curve of progression-free survival by histology based on EMA guidance in MAGNIFY: Induction safety population

Key: EMA, European Medicines Agency; FL, follicular lymphoma; MZL, marginal zone lymphoma.
Source: Celgene, 2018.[80]

B.2.7. Subgroup analysis

The subgroup data presented in this section focuses on the primary endpoint of each study (PFS for AUGMENT and response rate [ORR and CR rate] for MAGNIFY).

AUGMENT

Figure 10 presents a summary of subgroup analyses for PFS by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance.

Figure 10: Forest plot for subgroup analyses of IRC-assessed progression-free survival per 2007 IWGRC with EMA censoring in AUGMENT: ITT Population

Notes: Hazard ratio and its CI were estimated from unstratified Cox model except overall ITT Population, for which HR and its CI were estimated using Cox model adjusted by the three stratification factors.
Key: CI, confidence interval; EMA, European Medicines Agency; EU, European Union; FL, follicular lymphoma; GELF, Groupe d'Etude des Lymphomes Folliculaires; HR, hazard ratio; IRC, Independent Review Committee; ITT, intent-to-treat; IWGRC, International Working Group Response Criteria; MZL, marginal zone lymphoma; US, United States.
Source: Celgene, 2018.[77]

Across all predefined subgroups, PFS favoured the R^2 arm compared to the R mono arm. Furthermore, the difference between treatment arms was statistically significant

between each subgroup (with the exception of the Ann Arbor stage 1-2 and MZL subgroups).

Note especially that the factors suspected to contribute to worse prognosis (discussed in Section B.1.3) did not impact the relative effectiveness of R^2. The HRs for patients at second-line (one previous antilymphoma therapy) and those at third-line or above (>1 prior antilymphoma therapy) were consistent (HR: [redacted] vs. HR: [redacted]), with both groups favouring the R^2 arm.

In addition, a post-hoc subgroup analysis (based of FDA guidance) assessing PFS in FL patients found similar results between patients with and without POD24 (i.e. progression within 24 months of initial therapy). For patients with POD24, the HR was 0.41 (95% CI: 0.24, 0.68) and for patients with no POD24, the HR was 0.43 (95% CI: 0.28, 0.65), both groups favouring the R^2 arm.[81] The same analysis also showed the POD24 subgroup had similar response data compared to that of the overall ITT population.

MAGNIFY

For MAGNIFY, the subgroup analyses were only conducted in FL patients. ORR and CR rate was analysed for the relevant subgroups in the IEE population and the results are presented in Figure 11 and Figure 12, respectively.

In the IEE Population, responses were consistently observed within various subgroups, with ORRs ranging from [redacted] and CR rates ranging from [redacted].

Figure 11: Forest plot for subgroup analyses of overall response rate per 1999 IWGRC in MAGNIFY: IEE population

Key: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; IEE, induction efficacy evaluable; IEE, induction efficacy evaluable; IWGRC, International Working Group Criteria; ORR, overall response rate.
Source: Celgene, 2018.[82]

Figure 12: Forest plot for subgroup analyses of complete response rate (CR/CRu) per 1999 IWGRC in MAGNIFY: IEE population

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; IEE, induction efficacy evaluable.
Source: Celgene, 2018.[83]

As with AUGMENT, the factors suspected to contribute to worse prognosis (discussed in Section B.1.3) did not impact the relative effectiveness of R^2.

The ORR and CR rate for patients were favourable for both non-R-refractory and R-refractory patients (ORR: 78% vs. 63%; CR rate: 47% vs. 40%), demonstrating the consistent effectiveness of R^2 for both of these populations.

B.2.8. Meta-analysis

A meta-analysis to pool the relevant R^2 studies, AUGMENT and MAGNIFY, has not been conducted, with the rationale that pooling the studies would potentially add more complexity without additional benefit due to the following factors:

- There were differences in baseline characteristics between the two studies that were identified at the advisory board as being important (previous rituximab treatment, age, refractory to last prior regimen, line of therapy, disease stage and ECOG status), suggesting that patients in MAGNIFY have poorer prognosis than those in AUGMENT. Comparing the R^2 treated IITT population of MAGNIFY with the R^2-treated population of AUGMENT:
  - MAGNIFY included both R-refractory and non-R-refractory patients whereas AUGMENT included only non-R-refractory patients
  - Patients were older (55.4% vs. 46.1% were aged ≥65 years)
  - Patients were more heavily pre-treated with rituximab (95.8% vs. 85.4% had prior rituximab)
  - A greater percentage of patients were refractory to the last prior regimen (38.7% vs. 16.9%)
  - Patients were more heavily pre-treated (43.5% vs. 57.3% of patients had only one prior therapy)
- There is limited follow-up of MAGNIFY patients once the data has been censored at the end of the initial treatment period (maximum follow-up approximately 10 months), compared with the AUGMENT patients (maximum follow-up of approximately 45 months)

When comparing the subgroup of the MAGNIFY IITT population that only includes R^2-treated patients who were non-R-refractory with the AUGMENT ITT population who were treated with R^2, similar differences in patient populations were observed (except for the proportion of patients refractory to the last prior regimen, which was greater for AUGMENT).

B.2.9. Indirect and mixed treatment comparisons

No head-to-head data are available for R^2 versus any of the comparators of interest to this submission; only R mono was compared with R^2 within the AUGMENT RCT (Sections B.2.2-B.2.7). As such, for all relevant comparators, indirect treatment comparisons (ITCs) must be attempted using:

- Published evidence identified from the SLR (Section B.2.1).
- Evidence from a UK real-world evidence (RWE) registry, the HMRN database.

Note that the HMRN data were used to help address the limitations associated with the ITCs using published evidence. Key limitations include: the paucity of published data on comparator treatments for relapsed/refractory FL and MZL; and the incomparability of populations in terms of prior rituximab treatment – a key treatment effect modifier.

ITCs with data from published evidence

Data in this section are taken from the ITC report.[84]

Data sources

As described in Section B.2.1, two RCTs were identified in the SLR for comparators that are relevant for this submission and consideration in the ITC. The Van Oers RCT comparing R-CHOP with CHOP,[63] and the GADOLIN RCT comparing O-Benda with Benda mono.[64]

Because the final evidence base provides no common comparator linking R^2 and the comparators of interest, traditional ITC methods using anchored comparisons[85] cannot be applied. The ITC used aggregate data for comparators as available from the publications for the included studies, whilst individual patient data (IPD) was used for R^2 from the AUGMENT and MAGNIFY studies. Due to the exclusion of R-refractory populations from the AUGMENT study, AUGMENT was used solely for comparisons in the non-R-refractory population and MAGNIFY for comparisons in the R-refractory population. For the comparator studies, data were extracted for patient and study characteristics, objective response (OR), OS and PFS medians/rates, and AEs. Of note, TTNLT data was not consistently reported, and therefore was not included in this ITC. Where OS and/or PFS KM curves were available, these were digitized and used to generate pseudo-IPD. This enabled a range of survival models to be fitted to the data. The Guyot method was used to construct pseudo-individual patient event and censoring times from the digitized KM curves.[86]

Outcomes for analysis

The outcomes of interest in the ITC were as follows:

- Binary endpoints

- Efficacy: ORR, CR rate

The data required for analysis were the number of patients (n) and number of events (r) in each treatment arm. If the study reported the proportion of patients with events (p) instead of r, r will be calculated as r=p x n, rounded to the nearest integer.

- Time to event endpoints
  - OS (IPD/KM)
  - PFS (IPD/KM)

For the MAGNIFY study, data from the R^2 induction phase (pre-randomized maintenance) were used. Outcomes were calculated in all patients that started R^2 induction therapy, and not just in patients who were randomized into the maintenance phase. The maintenance phase in the MAGNIFY study is not concordant with anticipated clinical practice in the post-R^2 setting. As such, patients who had not progressed (for PFS only), died or been censored for other reasons by the start of the maintenance phase (12 months) had their OS and PFS censored at that point, to ensure the outputs are not influenced by this maintenance phase. The same definition of OS and PFS as in the clinical SAP for the MAGNIFY induction phase was used for consistency.

All analyses were performed in either the non-R-refractory or R-refractory patient populations according to the comparator of relevance. Analysis of efficacy endpoints were performed for the pooled FL and MZL population (because of their similar prognosis in the relapsed/refractory population – see Section B.1.3) for R-CHOP, and FL only for O-Benda, as this comparator does not have a license for MZL.

An overview of the studies, treatments, populations and outcomes available from the R^2 and two comparator studies of interest are presented in Appendix D2.

Statistical methods

Given that traditional ITC methods using anchored comparisons could not be applied, unanchored indirect comparisons (UAIC) of individual arms from different studies were performed.[87] A typical method for conducting such 'cross-study' comparisons and adjusting for differences in patient characteristics across studies is the matching-adjusted indirect comparison (MAIC) method, which accounts for

known imbalances in any effect modifiers (EM) or prognostic variables (PV) between the studies.

List of potential effect modifiers and/or prognostic variables

An effect modifier is a covariate that alters the effect of treatment on outcomes, so that the treatment is more or less effective in subgroups defined by different levels of the effect modifier. A prognostic variable is defined as a covariate that affects outcome (regardless of treatment). A covariate can be an effect modifier, a prognostic variable, both, or neither. The following is a list of potential effect modifiers/prognostic variables that would ideally be adjusted for in a MAIC, as identified and validated by external clinical experts.[1] UK clinical experts identified the highest priority covariates and they are marked with asterisks. Covariates shown in *italic text* had no data reported in the comparator studies and so could not be adjusted for in the MAICs.

- *Previous exposure to rituximab
- FLIPI (Follicular Lymphoma International Prognostic Index) components:
  - *Age (mean, or median if mean no reported, or % >60 years if neither reported)
  - Ann Arbor Stage (III-IV)
  - *Nodal sites (>4)*
  - High LDH
- *Refractory to last therapy
- *Prior lines of therapy 1 vs. 2 vs. >2 (or mean/median if categories not reported)
- *FLIPI risk group (low vs. intermediate vs. high)
- FLIPI2+ components:
  - *Serum beta-2 microglobulin high*
  - Bone marrow involvement
  - *Diameter of largest node >6 cm*
  - *Haemoglobin <12 dL/L*
- *FLIPI2 risk group (low vs. intermediate vs. high)*
- *Time from last treatment*
- *POD24*
- ECOG performance status (0–1 vs. 2+)

- Presence of B-symptoms

ECOG performance status (PS) was dropped from the MAICs because there were very few ECOG PS 2+ patients in AUGMENT/MAGNIFY, and the comparator studies also either had a small number of ECOG PS 2+ patients (hence were balanced) or did not report these data. The rationale for selecting these covariates as potential modifiers/prognostic variables is as follows:

- Previous exposure to rituximab appears to be a prognostic factor, with a greater response expected in rituximab naïve patients[88]
- FLIPI and FLIPI2 are validated prognostic indexes[89, 90]
- The link between early relapse (POD24) and survival in FL is described in Casulo et al. 2015[91]
- ECOG performance status has been used previously as an indicator of cancer response/progression[92]
- Previous lines of therapy and time from last treatment are likely to reflect the extent and severity of disease

If the adjustment resulted in an expected sample size and/or adjusted number of patients that was too small for analysis, then the list of variables used for adjustment was reduced before analysis. This was done to maintain the maximum number of the most clinically important variables in the adjustment. Several combinations of variables were explored. However, note that excluding known imbalanced covariates from matching may result in populations with differing levels of effect modifiers/prognostic variables on each treatment, which can bias the analysis results.

Populations used for matching

The R^2 patient characteristics were matched to the comparator patient characteristics within the population of interest according to histology and R-refractory status. The Van Oers, 2006 study[63] contained FL patients only, for which baseline characteristics and results were available. The Sehn, 2016[64] study included

a mixed histology population, for which the individual patient characteristics and results of interest for FL patients relevant to the reported analysis were presented.

Multiple studies

For the R^2 studies with IPD available, AUGMENT was used for the comparisons in the non-R-refractory population and MAGNIFY was used for comparisons in the R-refractory population.

Software

Analyses were performed in R version 3.5.1 or later.

Method

For an unanchored MAIC, it is assumed there is a treatment k_I (in this case, R^2) that has been studied in a population s_I for which there is IPD available. There is a comparator of interest k_A that has been studied in a population s_A, for which there is only aggregate data. The aim of the method is to re-weight the observed IPD results for k_I in population s_I to make it more similar to population s_A, thus enabling a comparison of k_I and k_A in a more comparable population.

The weights were calculated using the method of Signorovitch et al., 2010, 2012[93, 94]:

- Re-centre the IPD patient covariates XsI by subtracting the aggregate data mean covariate value $\bar{X}$sA to create XsI'
- The weights are then the values $\hat{\alpha}$ that minimize the following equation:

$$\sum_{j=1}^{n_I} \exp\left(\alpha^T \boldsymbol{X}'_{jsI}\right)$$

Analysis was then performed on the reweighted data using standard models for binomial or survival data.

Standard errors for the Cox model and parametric models were calculated using robust sandwich estimators[95] to account for the fact that the weights are estimated rather than known. Note that robust sandwich estimators were not available for the generalized gamma distribution and so usual standard errors are presented for this

distribution. The generalized gamma distribution is mostly used to better inform the choice of parametric model rather than to estimate the treatment effect.

An additional sensitivity analysis was performed for the Cox model using bootstrapping rather than robust sandwich estimators, as this has been shown to result in approximately correct standard errors, confidence intervals and coverage rates for the Cox model.[96] Bootstrapping was also performed for all parametric models including a treatment covariate except for generalized gamma (due to convergence issues). Ten thousand bootstrap simulations were run per model.

A summary of the two comparator studies, potential modifiers/prognostic variables covariate sets analysed, and populations used for both matching and analysis is presented in Table 15 with further detail in Appendix D.2.

Table 15: Comparator studies, EM/PV covariate set analysed, and populations used for both matching and analysis

Population of interest	Outcome	N	EM/PV set analysed	R^2 population used for matching	Comparator population used for matching	R^2 population used for analysis	Comparator population used for analysis
R-CHOP (Van Oers, 2006)							
FL+MZL	OS, PFS	234 FL, 0 MZL	All except previous rituximab	All	All (=FL)	All	All
FL+MZL	ORR, CR rate	234 FL, 0 MZL	All except previous rituximab	All	All	All	All
O-Benda (Sehn, 2016)							
FL	OS & PFS (censored at end of induction), ORR	164 FL	All	FL	FL	FL	FL
FL	CR rate	164 FL	All	FL	All (CR rate not reported for FL)	FL	All (CR rate not reported for FL)

Key: CR, complete response; EM, effect modifiers; FL, follicular lymphoma; MZL, marginal zone lymphoma; OS, overall survival; ORR, overall response rate; PFS, progression-free survival; PV, prognostic variables; R^2, rituximab plus lenalidomide; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Results

Data extracted from included studies for relevant outcomes

A summary of data extracted for the key efficacy outcomes from two comparator studies included in the ITC are presented in Appendix D2.

Non-R-refractory population

The results in this section are presented for OS and PFS only; a summary of the results for ORR and CR rate are presented in Appendix D2.

Summaries of potential effect modifiers/prognostic variables and matching results for all covariates for the R^2 (AUGMENT) versus R-CHOP (Van Oers) comparison are provided in Appendix D2. A summary of number of patients and events in the MAIC analyses for OS and PFS for the R^2 versus R-CHOP comparison is also provided in Appendix D2.

The KM curves for OS and PFS for the R^2 versus R-CHOP comparison are presented in Figure 13 and Figure 14, respectively. Table 16 provides an overall summary of the results from the MAIC analyses for R^2 versus R-CHOP. For this MAIC, it was not possible to match using all available potential modifiers/prognostic variables and results are presented using the maximum set of covariates (all except previous rituximab exposure). The inability to match to previous rituximab exposure was problematic as clinical experts deemed this covariate important. Although results were not statistically significant, R^2 led to improved OS ([redacted] and improved PFS ([redacted]), when compared with R-CHOP.

Figure 13: Kaplan–Meier plot of overall survival comparing R^2 (unadjusted and adjusted) with R-CHOP in the non-R-refractory population, for combined FL and MZL patients

Key: FL, follicular lymphoma; MZL, marginal zone lymphoma; R, rituximab; R^2, rituximab plus lenalidomide; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Figure 14: Kaplan–Meir plot of progression-free survival comparing R^2 (unadjusted and adjusted) with R-CHOP in the non-R-refractory population, for combined FL and MZL patients

Key: FL, follicular lymphoma; MZL, marginal zone lymphoma; R, rituximab; R^2, lenalidomide plus rituximab; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Table 16: Summary of MAIC results for R^2 versus R-CHOP (from Van Oers, 2006) for FL and MZL histologies

Outcome	Adjustment covariates	R^2 N, adjusted	R-CHOP N	R^2 median (OS/PFS) or rate, adjusted	R-CHOP median (OS/PFS) or rate	HR or OR (95% CI)	p-value
OS KM	All available except % previous R exposure	78.8	234	NR	NR	[redacted]	[redacted]
PFS KM	All available except % previous R exposure	78.8	234	30.4m	33m	[redacted]	[redacted]

Key: CI, confidence interval; FL, follicular lymphoma; HR, hazard ratio; KM, Kaplan–Meier; MAIC, matching-adjusted indirect comparison; MZL, marginal zone lymphoma; NR, not reported; OR, odds ratio; OS, overall survival; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

R-refractory population

The MAIC results in this section are presented for OS and PFS only, MAIC results for ORR and CR rate are presented in Appendix D2.

Summaries of potential effect modifiers/prognostic variables and matching results for all covariates for the R^2 versus O-Benda comparison are provided in Appendix D2. Matching results are also provided for studies with, ORR and CR rate in Appendix D2.

Table 17 provides a summary of the number of patients and events in the MAIC analyses for OS and PFS for the R^2 versus O-Benda comparison.

Table 17: Number of patients and events in the MAIC analyses for the R^2 versus O-Benda comparison

Treatment	N	Events	Median or %
OS, MAGNIFY censored at end of induction			
R^2	76.3	4.9	NR
O-Benda	164	38	NR
PFS, MAGNIFY censored at end of induction			
R^2	76.3	17.7	NR
O-Benda	164	93	26

Key: O-Benda, obinutuzumab plus bendamustine; OS, overall survival; MAIC, matching-adjusted indirect comparison; NR, not reported; PFS, progression-free survival; R^2, rituximab plus lenalidomide.

Figure 15 and Figure 16 present the KM curves for OS and PFS for the R^2 versus O-Benda comparison when MAGNIFY was censored at the end of the induction phase.

Figure 15: Kaplan–Meier curve for the MAIC analysis of OS, when MAGNIFY was censored at the end of the induction phase, comparing R^2 with O-Benda in the R-refractory population, for FL patients

Key: FL, follicular lymphoma; MAIC, matching-adjusted indirect comparison; O-Benda, obinutuzumab plus bendamustine; OS, overall survival; R, rituximab; R^2, rituximab plus lenalidomide.

Figure 16: Kaplan–Meier curve for the MAIC analysis of PFS, when MAGNIFY was censored at the end of the induction phase, comparing R^2 with O-Benda in the R-refractory population, for FL patients

Key: FL, follicular lymphoma; MAIC, matching-adjusted indirect comparison; O-Benda, obinutuzumab plus bendamustine; R, rituximab; R^2, rituximab plus lenalidomide.

A summary of the results from the MAIC analyses for R^2 (MAGNIFY) versus O-Benda (GADOLIN) is provided in Table 18. For this MAIC, it was possible to match using all available potential modifiers/prognostic variables. When OS & PFS were censored at the end of the MAGNIFY induction phase, there was no significant difference in OS ( ) or PFS ( ), between R^2 and O-Benda.

Table 18: Summary of MAIC results for R^2 versus O-Benda (from Sehn, 2016) for FL only

Outcome	Adjustment covariates	R^2 N, adjusted	O-Benda, N	R^2 median (OS/PFS) or rate, adjusted	Comparator median (OS/PFS) or rate	HR or OR (95% CI)	p-value
OS KM, censored at end of induction	All available	76.3	164	NR	NR	[redacted]	[redacted]
PFS KM, censored at end of induction		76.3	164	NR	26m	[redacted])	[redacted]

Key: CI, confidence interval; FL, Follicular lymphoma; HR, hazard ratio; KM, Kaplan–Meier; MAIC, matching-adjusted indirect comparison; NR, not reported; OR, odds ratio; OS, overall survival; O-Benda, obinutuzumab plus bendamustine; PFS, progression-free survival; R^2, rituximab plus lenalidomide.

Uncertainties from the ITC

It was not possible to adjust for all the potential effect modifiers/prognostic variables in the analysis. This was either due to no data being available for those variables being reported in the comparator studies, or insufficient patient numbers to inform matching. It was therefore necessary to assume that such variables are either not actual effect modifiers/prognostic variables, or that they are balanced across the treatment arms in the analysis.

For the MAIC between R^2 and R-CHOP in the non-rituximab refractory population, potential modifiers/prognostic variables that were known to be imbalanced between treatments had to be removed from the matching to prevent the effective sample size becoming too small for analysis. This means the key assumption of the MAIC (that all modifiers/prognostic variables are accounted for) may not hold. The affected covariate was previous rituximab exposure, which impacted the R^2 versus R-CHOP comparison (85.4% of R^2 patients had received prior rituximab compared to 0% of R-CHOP patients). Given that higher response rates have been demonstrated in rituximab naïve patients compared with patients previously treated with rituximab (75.0% vs. 57.1%), this imbalance biases against R^2.[88]

The R^2 IPD has been matched to the distribution of patient/study characteristics in the comparator studies. The comparator studies may differ between analyses, but the results are only applicable to the population of the specific comparator study in each analysis. For the relative treatment effect to be applicable in an alternative population, a further assumption must be made that there are either no EMs, or that the distribution of EMs is the same in the comparator study's population and the alternative population.

The R^2 patient characteristics were matched to the comparator patient characteristics within the population of interest according to histology and R-refractory status. The Van Oers, 2006 study contained FL patients only, for which baseline characteristics and results were available. The Sehn, 2016 study included a mixed histology population, for which the individual patient characteristics and results of interest for FL patients relevant to the reported analysis were presented.

The definitions of outcomes differed slightly between studies, which may have affected the results. For ORR, CR rate and PFS, different studies used different response criteria, and some used investigator assessment whilst others used an IRC.

In considering the limitations of the analyses and published evidence additional data were sought in order to further explore comparisons of relevance.

ITCs with data from HMRN

Introduction

As discussed previously, due to the limitations of the ITCs using published evidence, ITCs comparing R^2 to R-CHOP and R-CVP for non-R-refractory patients and O-Benda for R-refractory patients using UK-specific RWE, from HMRN, were explored. Due to small patient numbers for non-R-refractory patients receiving R-CHOP and R-CVP in the HMRN database, clinical expectation that R-CHOP and R-CVP would have similar efficacy in a relapsed/refractory setting and empirical data demonstrating this to be the case, efficacy analyses compared R^2 to the pooled R-CHOP/R-CVP (Appendix D3 includes detail of the KM plots and Cox PH models that support the assumption of equal efficacy).

There were 63 patients identified as receiving either R-CVP or R-CHOP as 2L+ treatment. Comparisons were made for the key time to event outcomes collected within the AUGMENT clinical study (OS, TTNLT and PFS).

There was no data for R-refractory patients receiving O-Benda in the HMRN database (due to this regimen only being recently available) and so this data source was not used for this population.

HMRN database

The HMRN is a population-based cohort comprising a total population of 3.8 million people covering the former adjacent UK Cancer Networks of Yorkshire and the Humber & Yorkshire Coast. The HMRN was established in 2004 to provide robust, generalizable data to inform clinical practice and research and collects detailed information about all haematological malignancies in the region.

The HMRN identified [redacted] patients who had received ≥1 prior line of chemotherapy for treatment of FL and were identified as being non-R-refractory or R-refractory after each treatment line.

For the subgroup of patients who were non-R-refractory, [redacted] patients received R-CVP and [redacted] patients received R-CHOP as a second or later line therapy (2L+), although most patients [redacted] received these treatments in second-line (2L). Patients could be included in both treatment subgroups if they had received both treatments in different lines of therapy, for example, R-CHOP in 2L and R-CVP in 3L. As described previously, data were pooled given the small patient numbers and similar efficacy between treatments. This assumption of equivalent efficacy is further supported by clinical expert opinion.

As described in section B.2.13, TTNLT is an important outcome to examine for comparisons of clinical study and real-world data in this disease setting. The definition of TTNLT as used for the HMRN analysis is time to documentation of new anti-lymphoma treatment from 'baseline'. The definition of PFS as used for the HMRN analysis is time from 'baseline' to disease progression (including transformation to diffuse large B-cell lymphoma) or death due to any cause and the definition of OS was time from start of treatment to date of death or if still alive censored at 18th December 2018.

Statistical methods

As for the ITCs that used published data to inform the comparator evidence, MAICs were performed using the methodology as described in Signorovitch et al., 2010, 2012[93, 94] and referenced in the NICE Decision Support Unit (DSU) and Technical Support Document (TSD) 18,[87] in an attempt to adjust for the differences observed in demographic and disease characteristics between the data sources. The outcomes of interest were OS, TTNLT and PFS.

Matching variables

The baseline characteristics that were commonly collected by the HMRN and the AUGMENT study are presented in Appendix D3. The list of potential modifiers/prognostic variables discussed previously in the context of the ITC with published data, was used to identify the matching variables. A key treatment effect

modifier/prognostic factor that was not collected by the HMRN was the FLIPI risk category. However, three of the four FLIPI components were collected (only LDH was not collected). Matching was therefore performed for the following variables:

- Age ≥60 years (FLIPI component)
- Ann Arbor Stage III-IV (FLIPI component)
- Nodal sites >4 (FLIPI component)
- Prior rituximab treatment
- Prior lines of therapy (1 vs. 2 vs. >2)
- POD24 status

Only 34 of the 63 R-CVP/R-CHOP-treated patients were fully staged, thereby providing information on nodal sites and Ann Arbour stage. For matching, it was therefore assumed that the missing data was equally distributed across the categories of these two factors.

Results

Following the matching procedure, the weighted baseline characteristics for AUGMENT patients were compared with the R-CVP/R-CHOP HMRN population. The MAIC has led to re-weighted AUGMENT covariates that are the same as in the R-CVP/R-CHOP HMRN population (matching results are presented in Appendix D3).

Table 19 provides a summary of the number of patients and events in the MAIC analyses for OS, TTNLT and PFS.

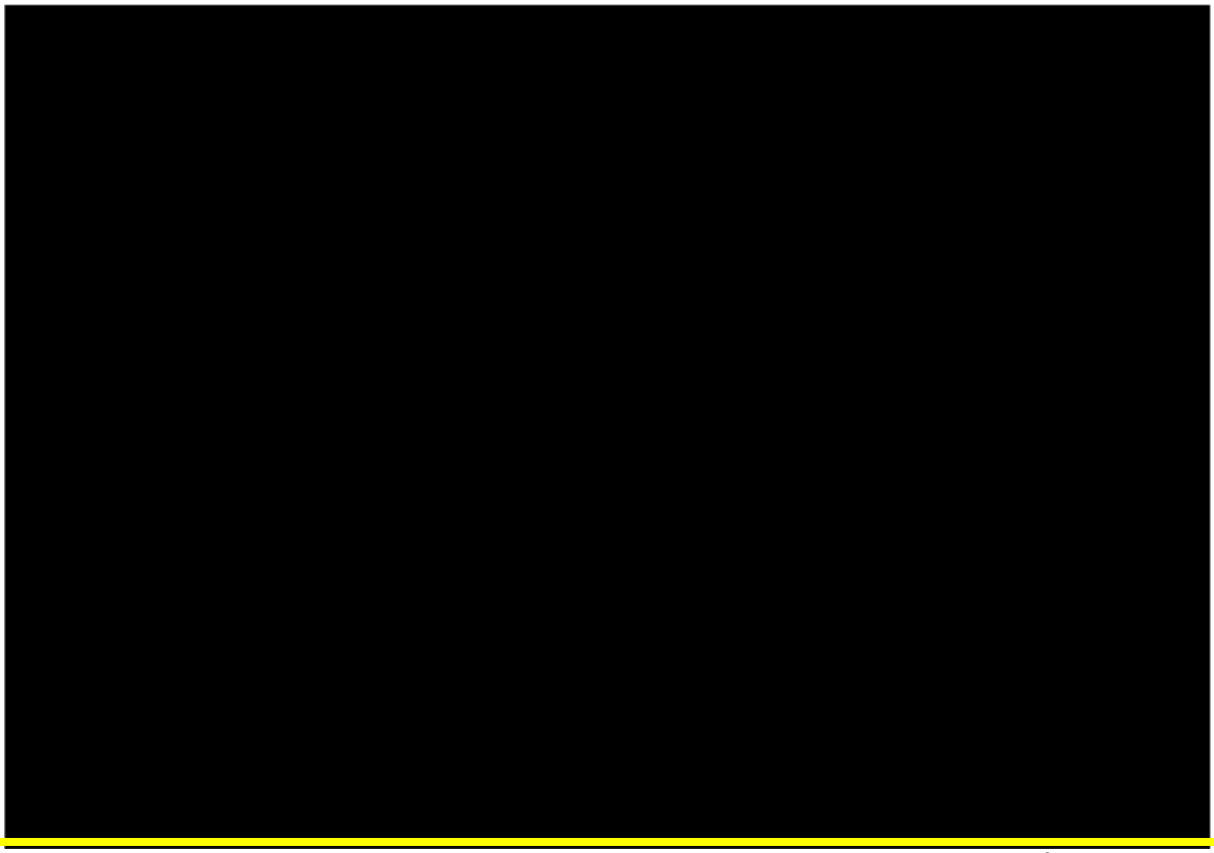
Table 19: Number of patients and events in the MAIC analyses for the R^2 (AUGMENT) versus R-CVP/R-CHOP (HMRN) comparison

Treatment	N	Events	Median survival
OS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
PFS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
TTNLT			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]

Key: HMRN, Haematological Malignancy Research Network; MAIC, matching-adjusted indirect comparison; NA, not applicable; OS, overall survival; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; TTNLT, time-to-next antilymphoma therapy.

The KM curves comparing R^2 (weighted and unweighted) to R-CVP/R-CHOP for OS, PFS and TTNLT are presented in Figure 17, Figure 18 and Figure 19, respectively. For all three endpoints, the weighting does not considerably alter the R^2 KM curves. The curves suggest R^2 has a significant survival benefit compared to R-CVP/R-CHOP and a benefit for TTNLT; however, with a more modest PFS benefit. This latter observation may be a function of the PFS assessments for the AUGMENT and HMRN populations being conducted differently (proactively on-study and reactively in the real-world setting) as described below and in section B.2.13. The more like-for-like TTNLT comparison provides supporting evidence for the OS benefit observed with R^2.

Figure 17: Kaplan–Meier curve for the MAIC analysis of OS comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population



Key: MAIC, matching-adjusted indirect comparison; OS, overall survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Figure 18: Kaplan–Meier curve for the MAIC analysis of PFS comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population

Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Figure 19: Kaplan–Meier curve for the MAIC analysis of TTNLT comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population

Key: MAIC, matching-adjusted indirect comparison; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; TTNLT, time-to-next antilymphoma treatment.

HRs from the Cox Proportional-Hazard models comparing R^2 and R-CVP/R-CHOP are presented in Table 20. Log-cumulative hazard plots for these three outcomes (see Section B.3.3) suggest that the proportional hazards assumption may not hold. However, the HRs and corresponding CIs support the interpretation of the KM curves: R^2 has survival benefit compared to R-CVP/R-CHOP and a benefit for TTNLT, with modest PFS improvement.

Table 20: Results from Cox Proportional Hazard models comparing R^2 and R-CVP/R-CHOP

Outcome	R^2, adjusted N	R-CVP/R-CHOP N	HR (95% CI)[a]
OS	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]
TTNLT	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival; PFS, progression-free survival; R^2, rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone TTNLT, time to next antilymphoma treatment.
Notes: [a], bootstrapped CI.

Uncertainties relating to the use of HMRN data

As for the MAIC analyses with published data, it was not possible to adjust for all of the potential effect modifiers/prognostic variables in the analysis. Data was not available to match on FLIPI risk group.

The comparability of data sources for these analyses is uncertain due to limitations when comparing RCT data to RWE. Some of these limitations are as follows:

- In AUGMENT, monitoring of patients is frequent and includes regularly scheduled imaging investigation as per the protocol, whereas, HMRN patients are monitored less frequently and without routine imaging investigations. This has particular relevance with respect to PFS; mandated imaging per study protocol will identify progression events including those that are asymptomatic and do not require immediate treatment (i.e. a proactive approach to evaluation). By contrast, in the clinical setting, imaging is not routine and progression events are most likely determined only after symptomatic presentation of patients requiring treatment (i.e. reactive evaluation). It is therefore possible that progression events are recorded later in the real-world setting than in AUGMENT, which is likely to overestimate PFS for the HMRN patients
- Time to Next Anti-Lymphoma Treatment provides a more comparable endpoint across the data sets; treatment will generally be initiated in this disease setting only in symptomatic patients, a trigger common to both study patients and real-world patients. However, follow-up of patients on-study is likely to have occurred

more frequently than in the real-world setting, providing more frequent opportunities for reporting of symptoms prompting treatment and potentially underestimating TTNLT in the study population

- The definition of PFS differs across the two data sources; these analyses compare PFS by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance for AUGMENT and time from baseline to disease progression (including transformation to diffuse large B-cell lymphoma) or death due to any cause for the HMRN analyses

Furthermore, although follow-up of HMRN patients was longer than in AUGMENT, sample sizes per treatment received from the HMRN database were relatively small. Pooling of the patients who had received R-CVP or R-CHOP was conducted in order to reduce any bias from a small sample size.

B.2.10. Adverse reactions

AUGMENT

Data from this section are taken from the 22 June 2018 database cut-off; safety analyses were conducted in the safety population.

Treatment exposure

Overall, the median lenalidomide/placebo treatment duration was [redacted] for the R^2 arm and [redacted] months for the R mono arm. The median rituximab treatment duration was also similar between the R^2 and R mono arms ([redacted] vs. [redacted], respectively). Of note, the median treatment duration was consistent between FL and MZL patients.

A greater proportion of patients in the R^2 arm completed all 12 cycles of planned study treatment (lenalidomide or placebo) compared with patients in the R mono arm (71% vs. 62%, respectively). Rituximab dosing was similar between the R^2 arm and the R mono arm, with the majority of patients completing the planned five cycles of rituximab treatment (89% vs. 90%, respectively).

The majority of patients in both the R^2 arm (86%) and R mono arm (87%) received a starting lenalidomide or placebo dose of 20 mg. The average daily dose of lenalidomide/placebo was consistent between treatment arms with a median dose of

20 mg per day. The median relative dose intensity of lenalidomide for R^2-treated patients was [redacted] (vs. 99% for placebo), and the proportion of patients who had a relative dose intensity of ≥90% and <110% was lower in the R^2 arm (55%) than the R mono arm (83%). For rituximab, the median relative dose intensity was similar between the R^2 and R mono arms [redacted] vs. 99%, respectively). Again, the proportion of patients who had a relative dose intensity of ≥90% and <110% was lower in the R^2 arm (77%) than the R mono arm (91%).

A summary of treatment exposure by study drug for AUGMENT is presented in Appendix F1.

Adverse events

Summary of treatment-emergent adverse events

Table 21 presents a summary of the TEAEs during AUGMENT. TEAEs were reported in 174 patients (99%) in the R^2 arm and 173 patients (96%) in the R mono arm. More patients in the R^2 arm (69%) experienced a Grade 3 or 4 TEAE compared with those in the R mono arm (32%), and only two patients in each treatment arm reported a Grade 5 TEAE. Additionally, a greater proportion of patients reported serious adverse events in the R^2 arm (26%) compared with those in the R mono arm (14%).

A greater proportion of patients experienced TEAEs leading to dose reductions of lenalidomide or placebo in the R^2 arm than in the R mono arm (26% vs. 3%, respectively). TEAEs leading to dose interruptions of lenalidomide or placebo were also more frequent in the R^2 arm compared with the R mono arm (64% vs. 26%, respectively). Similarly, a greater proportion of patients in the R^2 arm (34%) had at least one TEAE leading to dose interruption of rituximab compared with those in the R mono arm (21%). Furthermore, the proportion of patients who had at least one TEAE which caused them to discontinue therapy was slightly higher in the R^2 arm compared with the R mono arm for both lenalidomide or placebo therapy (9% vs. 5%), and for rituximab therapy (3% vs. 1%).

Table 21: Summary of treatment-emergent adverse events, AUGMENT – Safety population

	Total	
	R^2 (N=176)	**R mono (N=180)**
Number of patients (%)		
Any TEAE	174 (98.9)	173 (96.1)
Len related	159 (90.3)	118 (65.6)
R related	132 (75.0)	105 (58.3)
Grade 3–4 TEAE	121 (68.8)	58 (32.2)
Len related	101 (57.4)	38 (21.1)
R related	57 (32.4)	19 (10.6)
Grade 5 TEAE	2 (1.1)	2 (1.1)
Any SAE	45 (25.6)	25 (13.9)
Len related	23 (13.1)	8 (4.4)
R related	13 (7.4)	3 (1.7)
Any TEAE leading to dose reduction of Len/Pbo	46 (26.1)	6 (3.3)
Any TEAE leading to dose interruption of Len/Pbo	112 (63.6)	47 (26.1)
Any TEAE leading to dose interruption of R	60 (34.1)	37 (20.6)
Any TEAE leading to discontinuation of Len/Pbo	15 (8.5)	9 (5.0)
Any TEAE leading to discontinuation of R	6 (3.4)	2 (1.1)

Key: Pbo, placebo; Len, lenalidomide; R, rituximab; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[77]

Most common treatment-emergent adverse events

In the safety population, TEAEs that occurred more frequently (≥10% difference) in the R^2 arm than the R mono arm included the following: neutropenia (58% vs. 22%), diarrhoea (31% vs. 23%), constipation (26% vs. 14%), cough (23% vs. 17%), upper respiratory tract infection (18% vs. 13%) and leukopenia (20% vs. 9%).

The difference in the number of Grade 3 or 4 TEAEs between treatment arms (shown in Table 21) was largely driven by Grade 3 or 4 events of neutropenia and leukopenia. Neutropenia occurred in 88 patients (50%) in the R^2 arm compared with

23 patients (13%) in the R mono arm, and leukopenia occurred in 12 patients (7%) in the R^2 arm compared with three patients (2%) in the R mono arm.

The most common TEAEs, occurring in more than 10% of patients, are presented in Appendix F1.

MAGNIFY

Treatment exposure

At the 10 August 2018 database cut-off, the overall median treatment duration for the induction safety population was [redacted] months. For lenalidomide, the median treatment duration was [redacted] months; and for rituximab the median treatment duration was [redacted] months. In total, [redacted] patients ([redacted]%) completed all 12 cycles of induction treatment.[97] It should be noted that the low completion percentage is related to the maturity of the study and not withdrawals and discontinuations alone. As such, [redacted] patients ([redacted]%) are still continuing on both lenalidomide and rituximab in the induction phase.

The majority of patients [redacted]%) received a starting lenalidomide dose of 20 mg. The median average daily dose of lenalidomide was 18.3 mg/day and the median relative dose intensity was [redacted] mg/week. For rituximab, the median average daily dose was [redacted] mg/day and the median relative dose intensity was [redacted] mg/week. The proportion of patients who had a relative dose intensity of ≥90% and <110% was [redacted]% for lenalidomide and [redacted]% for rituximab.

A summary of treatment exposure during the initial treatment period is presented in Appendix F2.

Adverse events

All AEs were assessed starting after the patient signed the informed consent and until 28 days after they discontinued taking the study drug. AEs that lead to patients discontinuing the study were followed until the problem was resolved or stabilized.

Summary of treatment-emergent adverse events

Table 22 presents a summary of TEAEs in the induction safety population during MAGNIFY. Overall, TEAEs were reported in [redacted] patients ([redacted]%). A total of [redacted]

patients ([redacted]%) reported a Grade 3 or 4 TEAE and only seven patients ([redacted] reported a Grade 5 TEAE. Serious TEAEs were reported in [redacted] patients ([redacted]

In total, [redacted] patients ([redacted]%) reported a TEAE leading to a dose reduction of lenalidomide. A greater proportion of patients experienced a TEAE leading to a dose interruption of lenalidomide compared with those leading to a dose interruption of rituximab ([redacted]% vs. [redacted]%, respectively). Similarly, more patients reported a TEAE leading to early discontinuation of lenalidomide compared with early discontinuation of rituximab ([redacted]% vs. [redacted]%, respectively).

Table 22: Overview of treatment-emergent adverse events, MAGNIFY – induction safety population

	Total (N=359)
Number of patients (%)	
Any TEAE	[redacted]
Len related	[redacted]
R related	[redacted]
Grade 3–4 TEAE	223 (62.1)
Len related	[redacted]
R related	[redacted]
Grade 5 TEAE	[redacted]
Any serious TEAE	[redacted]
Len related	[redacted]
R related	[redacted]
Any TEAE leading to dose reduction of Len	[redacted]
Any TEAE leading to dose interruption of Len	[redacted]
Any TEAE leading to dose interruption of R	[redacted]
Any TEAE leading to early discontinuation of Len	[redacted]
Any TEAE leading to early discontinuation of R	[redacted]

Key: FL, follicular lymphoma; Len, lenalidomide; Len, lenalidomide; MZL, marginal zone lymphoma; R, rituximab; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[98]

Most common treatment-emergent adverse events

In the induction safety population, the most commonly reported TEAEs (≥10% of patients) included: [redacted] ([redacted]%), [redacted] ([redacted]%), [redacted] ([redacted]%), [redacted] ([redacted]%), [redacted] ([redacted]%), [redacted] ([redacted]%) and [redacted] ([redacted]).

The most common TEAEs, occurring in more than 10% of patients, are presented in Appendix F2.

Adverse events profile for relevant comparators

A summary of common and very common AEs observed with R^2, and all the relevant comparator treatments or regimen components, as reported in the SmPC for each is presented in Appendix F3.

In general, R^2 has a safety profile that is consistent with the established individual profiles of lenalidomide and rituximab and differs from chemotherapy-based regimens.[58] In a head-to-head comparison with R-chemo in a first-line RCT in patients with FL, R^2 was associated with a lower incidence of any grade TEAEs compared with R-chemo.[5] Another RCT examining R^2 versus R-chemo in untreated FL patients found no unexpected toxicities associated with R^2.[99] The study reported that associated AEs were manageable and that there were seemingly fewer compared with R-chemo.

Consistent with this, Table 34 in Appendix F3 suggests that R^2 exhibits a different individual AE profile than its comparators. In particular, R^2 is not as commonly associated with the standard toxicities associated with chemotherapy treatments (e.g., vomiting, peripheral neuropathy, alopecia).

B.2.11. Ongoing studies

There are no ongoing studies that are anticipated to provide additional evidence for R^2 within the next 12 months.

B.2.12. Innovation

R^2 represents a significant innovation in the management of patients with previously treated FL and MZL, recently becoming the first chemotherapy-free combination immunotherapy regimen licensed in this setting by the US Food and Drug Administration. The regimen is currently pending approval in the EU. All existing options available in the UK contain standard chemotherapy components (except R mono, which is used in only ~1% of UK patients). R^2 provides an opportunity for clinicians to offer patients with relapsed/refractory disease an alternative to retreatment with the same chemotherapy-based approach used in their first-line treatment, therefore representing a step-change in the management of FL and MZL.[3] Furthermore, R^2 is anticipated to be the first treatment to be licensed specifically for treatment of MZL in Europe with the potential to address the substantial unmet need that exists in this population.[1]

Notably, R^2 is characterised by immune enhancement rather than by the immune suppression observed with standard chemotherapy in NHL.[100, 101] Combining lenalidomide with rituximab provides a means to restore and enhance patients' own immune function through complementary and synergistic mechanisms of action.[3, 4]. Lenalidomide has been shown to rapidly increase NK and T-cells, as well as re-activate dysfunctional NK- and T-cells. It also enhances immune synapse formation, leading to killing of FL and MZL cells, while rituximab leads to NK cell–mediated ADCC of FL and MZL cells. As such, combining lenalidomide with rituximab has demonstrated an ability to enhance tumour cytotoxicity.[3, 4] R^2 is anticipated to provide durable benefit as an alternative to current standards of care in relapsed/refractory FL and MZL. Furthermore, the direct comparison with R mono in the AUGMENT study demonstrated that R^2 significantly improved the primary endpoint, PFS (39.4 months [95% CI: 22.9–NR] vs. 14.1 months [95% CI: 11.4–16.7]; HR: 0.45 [95% CI: 0.33–0.61]; p<0.0001), and the best overall response rate (77.5% vs. 53.3%; p<0.0001) in addition to other secondary endpoints. Importantly,

patients treated with R^2 demonstrated a longer PFS than they had experienced on their prior regimen (39.4 months v 32.4 months), reversing the trend for deteriorating PFS with successive treatments, and providing evidence for the clinical value of a regimen with a mechanism of action distinct from existing standard of care interventions.

As R^2 is a chemotherapy-free combination regimen, it has an AE profile different to those associated with existing chemotherapy-based options.[100, 101] In an in vitro model of myeloid differentiation, lenalidomide demonstrated a reversible arrest in neutrophil maturation, with no loss of cell viability that was distinct from the irreversible neutrophil cytotoxicity of the chemotherapeutic agent, bendamustine.[3] This is consistent with, and helps to explain, the lower rates of Grade 3 or 4 neutropenia observed with R^2 versus R-chemo as reported for the first-line RELEVANCE study.[5] Furthermore, the safety profiles of the components of R^2, rituximab and lenalidomide, are already established and familiar to clinicians, thereby helping to minimize the safety fears associated with use of a novel treatment.

B.2.13. Interpretation of clinical effectiveness and safety evidence

Overall, the clinical evidence available provides an appropriate base to inform the assessment of clinical effectiveness and cost-effectiveness of R^2 for the treatment of previously treated FL and MZL.

Existing treatment options leave a substantial unmet need in the management of relapsed/refractory FL and MZL. Patients with FL who are non-R-refractory are generally offered R-CHOP and R-CVP, whilst those who are R-refractory have more limited options but most likely receive O-Benda, via the CDF. Patients with relapsed/refractory MZL have a broadly similar management to that of patients with relapsed/refractory FL, albeit without access to O-Benda. Existing chemotherapy-based treatment options do not fully address the immune dysfunction associated with FL and are associated with well-documented toxicities (e.g., vomiting, peripheral neuropathy, alopecia).[5] There is a need for novel treatment approaches, such as R^2, that target the immune microenvironment and offer alternatives to repeated exposure to chemotherapy-based interventions, with benefit derived from changes in both mechanism of action and toxicity profile.[55]

The efficacy of R^2 has been demonstrated across the spectrum of relapsed/refractory FL and MZL. In AUGMENT (non-R-refractory patients only), PFS was significantly improved for R^2 versus R mono (39.4 vs. 14.1 months; HR 0.45 [95% CI: 0.33, 0.61]; p<0.0001), an outcome consistently observed across a wide range of pre-defined subgroups. TTNLT, an endpoint with greater clinical significance to patients than PFS,[52] and one more readily applicable to comparison with real-world data, was significantly improved for R^2 versus R mono (HR: 0.54 [95% CI 0.38, 0.78] p=0.0007). Furthermore, OS showed a relative reduction of 39% in the risk of death (HR 0.61 [95% CI 0.33, 1.13]) for patients treated with R^2 versus R mono.

ORR was significantly improved with R^2 versus R mono (77.5% vs. 53.3%; p<0.0001). Additionally, patients in the R^2 arm were more likely to respond to subsequent anti-lymphoma treatment than those in the R mono arm (ORR: 57% vs. 36%; CR: 31% vs. 16%), lending weight to the hypothesis that R^2 may re-sensitise patients to subsequent treatment.[3]

Results from interim analysis of the MAGNIFY study are supportive of the AUGMENT data. The ORR for the overall IEE population was [redacted], and the PFS rate at 1 year for patients in the induction safety population was [redacted]. In particular, MAGNIFY demonstrated favourable outcomes for R^2 in both R-refractory and non-R-refractory patients. MAGNIFY also provides favourable outcomes for both the FL and MZL populations, supporting the AUGMENT multivariate analyses outcome.

Overall, across both the AUGMENT and MAGNIFY studies, R^2 demonstrated a predictable, manageable and acceptable safety profile, consistent with those of the individual components of R^2 (lenalidomide and rituximab) with no new safety signals observed for R^2. Neutropenia, a known adverse event associated with lenalidomide, was well managed for patients receiving R^2 through dose modifications, with few patients requiring treatment discontinuation (8.5% discontinuing lenalidomide, and 3.4% discontinuing rituximab). Further supportive safety data are available from a head-to-head study comparing R-chemo and R^2 in untreated FL patients, in which R^2 was associated with a lower incidence of TEAEs, manageable AEs, and no unexpected toxicities.[5] When AE profiles are compared between R^2 and components

of existing treatment regimens for FL and MZL, R^2 exhibits a different AE profile including lower risk of Grade 3-4 neutropenia and febrile neutropenia.[5]

Indeed, in the AUGMENT study, R^2 was found to have no detrimental impact on HRQL in patients with previously-treated FL or MZL despite the temporary impact of symptoms (e.g. fatigue, constipation, appetite loss, and diarrhoea) during the treatment period. In the context of the significant extension of PFS provided by R^2 compared with rituximab alone, the AE profile for R^2 appears to be outweighed by its clinical benefit.

In the absence of head-to-head comparisons of R^2 with relevant comparators, indirect comparisons were conducted using RWE from the UK HMRN database (for R-CHOP and R-CVP in the non-R-refractory population) and published evidence from a SLR (O-Benda in the R-refractory population). Published literature was also available for R-CHOP, though comparisons to this had limitations.

In the non-R-refractory population, relevant comparators were R-CHOP and R-CVP. Due to limitations of using published literature for the ITC (primarily that the100% rituximab-naïve population in Van Oers is not reflective of UK practice, and the lack of data for R-CVP in the relapsed/refractory setting), the economic base-case analysis was conducted using data from UK HMRN. To overcome the small patient numbers available from HMRN and given clinical opinion that R-CHOP and R-CVP may have equivalent efficacy in this population, the data for these patients were pooled.

Results derived from the Cox proportional hazard models using the HMRN data reported benefits in OS ([redacted]) and TTNLT ([redacted]) for R^2 compared to R-CVP/R-CHOP. The benefit for PFS is more modest ([redacted].

The availability of TTNLT data from HMRN is important, as it provides reassurance regarding the OS benefit described for R^2 compared with R-CHOP/R-CVP given the more modest improvement in PFS. The evaluation of PFS often differs between clinical studies and real-world data. In a clinical study, progression events are determined through imaging investigations performed on a protocol-mandated schedule (in AUGMENT, these investigations were conducted every 12 weeks).

Using this approach, particularly in indolent diseases like FL and MZL, progression may be documented substantially ahead of a patient having sufficient symptoms to warrant initiation of subsequent treatment. In clinical practice, imaging investigations are not conducted routinely and most likely occur only when a patient presents with symptoms. Therefore, disease progression is likely identified at a later stage than in the setting of a clinical study, often at the point at which subsequent treatment is imminently required. Accordingly, comparing PFS from clinical studies to that derived from real-world data may well be misleading and bias against the clinical study data. TTNLT, by contrast, represents an outcome triggered in a similar fashion in both settings, thus providing a more like-for-like comparison than PFS. The improvement in TTNLT for R^2 compared to HMRN-derived data for R-CHOP/R-CVP is more convincing than the improvement in PFS and more closely aligned with the observation for OS. This consideration is in line with a recent NICE appraisals in FL, where TTNLT has been recognised as a clinical endpoint that is more relevant to patients in clinical practice than PFS (TA513); furthermore, during the appraisal of O-Benda for R-refractory FL (TA472), the Evidence Review Group (ERG) specifically requested TTNLT data at clarification stage as they felt this was a relevant outcome, despite not being listed in the NICE final scope.[52] [48]

Consistent with the HMRN data, when using published literature, R^2 leads to improved OS ([redacted]), and PFS ([redacted]) when compared with R-CHOP. Given clinical opinion that R-CHOP and R-CVP may have equivalent efficacy in this population, it is assumed that similar results would be expected for a comparison between R^2 and R-CVP. The comparison of TTNLT between R^2 and R-CHOP or R-CVP could not be conducted because the publication identified via the SLR for R-CHOP did not provide TTNLT data.

In the R-refractory population, the only relevant comparator was O-Benda, for which, one published RCT was identified from the SLR that was used to compare with R^2 data from MAGNIFY using MAIC methodology. Results of the analyses found that when OS and PFS were censored at the end of the MAGNIFY induction phase (to remove the potential of confounding by the maintenance phase of the study), there were no significant differences between R^2 and O-Benda for OS ([redacted]:

[redacted]) or PFS ([redacted]). However, CRR was superior for R^2 over O-Benda (38.6% vs 16.7%; OR: 3.15 [95% CI: 1.69, 5.86]; p=0.0003), although there were no significant differences between R^2 and O-Benda for ORR (59.7% vs 67.7%; OR: 0.71 [95% CI: 0.38, 1.32]; p=0.2763).

Strengths and limitations of the clinical evidence base

Strengths

The pivotal study underpinning this submission is the Phase III, multicentre, double-blind, randomised AUGMENT study versus R mono in patients with non-R-refractory FL and MZL. The submission is further supported by the induction phase of the ongoing multicentre, open-label, randomised, Phase IIIb MAGNIFY study with both R-refractory and non-R-refractory patients. The AUGMENT and MAGNIFY studies were conducted in line with GCP guidelines, with steps taken to minimise the risk of bias, including establishment of independent external Data Monitoring Committees to provide oversight of safety and efficacy considerations, as well as study conduct.

In view of the absence of head-to-head data for R^2 versus relevant comparators (R-CHOP, R-CVP and O-Benda), multiple approaches were explored to conduct indirect comparisons with all published clinical study data available via an SLR, as well as with real-world data from the UK HMRN database.

The outcomes used in AUGMENT and MAGNIFY are relevant to clinical practice. The primary endpoint of the AUGMENT study was PFS, and was supported by clinically relevant secondary endpoints, OS and TTNLT. As discussed earlier, TTNLT has previously been recognised as a more clinically relevant endpoint for patients compared with PFS and provides a like-for-like endpoint to use in comparison with real-world data; its inclusion increases confidence in the results of these analyses. Quality of life was also assessed by means of the cancer-specific EORTC QLQ-C30 and the generic EQ-5D-3L, and no detrimental impact was observed was observed as a result of treatment with R^2 compared to R mono, strengthening the risk/benefit evaluation for R^2.

The patients in the AUGMENT study are similar to the non-R-refractory population in UK clinical practice, as evidenced by the comparison of baseline characteristics of

the patients with relapsed/refractory FL between the study and the HMRN database in Table 28 in Appendix D3. Clinical advisors confirmed that HMRN is a valid source for such data and that these data are considered representative of the UK as a whole, both in terms of patient population and clinical practice.[1]

The MAGNIFY study provides a substantial number of R-refractory patients in whom favourable outcomes have been observed, in addition to providing supportive data for the outcomes reported in the AUGMENT study.

Limitations

A key limitation of the evidence base is the lack of head-to-head to data with a relevant comparator. AUGMENT is an RCT comparing R^2 to R mono, which is a treatment rarely used in UK clinical practice.[1] The limitations of comparator evidence were further confounded by the inability to conduct formal indirect comparisons using RCT evidence because of a lack of a common comparator, which would be typically used to form an evidence network.

Furthermore, there is a distinct lack of published evidence for relevant comparators in relapsed/refractory FL and MZL. In a SLR, only two relevant comparator studies were identified – one RCT for R-CHOP, and one for O-Benda. No published literature was available for R-CVP. Therefore, for purposes of constructing ITC analyses using published evidence, it has been necessary to assume equivalent efficacy for R-CHOP and R-CVP, an assumption acknowledged as reasonable by UK clinical experts. In an effort to alleviate the issues around limited published comparator data, further data were obtained from the UK HMRN database of FL patients, for the comparators, R-CHOP and R-CVP.

Another limitation was the inability to adjust for all potential modifiers/prognostic variables in the non-R-refractory population, using either published literature or HMRN data.

Limitations associated with the HMRN analyses were small patient numbers and the resultant requirement to pool data for R-CHOP and R-CVP, a lack of data in MZL patients, limited data in R-refractory patients, and limitations in the comparability of the data sets. A key limitation in the latter regard was the difference in the scheduling

of imaging investigations to determine progression events (i.e., protocol-mandated in AUGMENT and ad-hoc in HMRN; see Section B.2.9 for more details).

End of life considerations

The life expectancy of relapsed/refractory FL with rituximab-based treatment exceeds 24 months,[25] and as such, R^2 is not relevant for end-of-life considerations. The number of patients anticipated to receive second-line chemotherapy or beyond in England each year and therefore be eligible for R^2 is [redacted] patients ([redacted] with FL and [redacted] with MZL; note: these numbers do not add to 448 due to rounding).[38]

B.3. Cost effectiveness

B.3.1. Published cost-effectiveness studies

A systematic review of the published literature was conducted to identify all relevant economic evaluations/modelling studies for the treatment of adult patients with relapsed/refractory FL or MZL.

The search was conducted on 8 February 2019 using the following electronic databases:

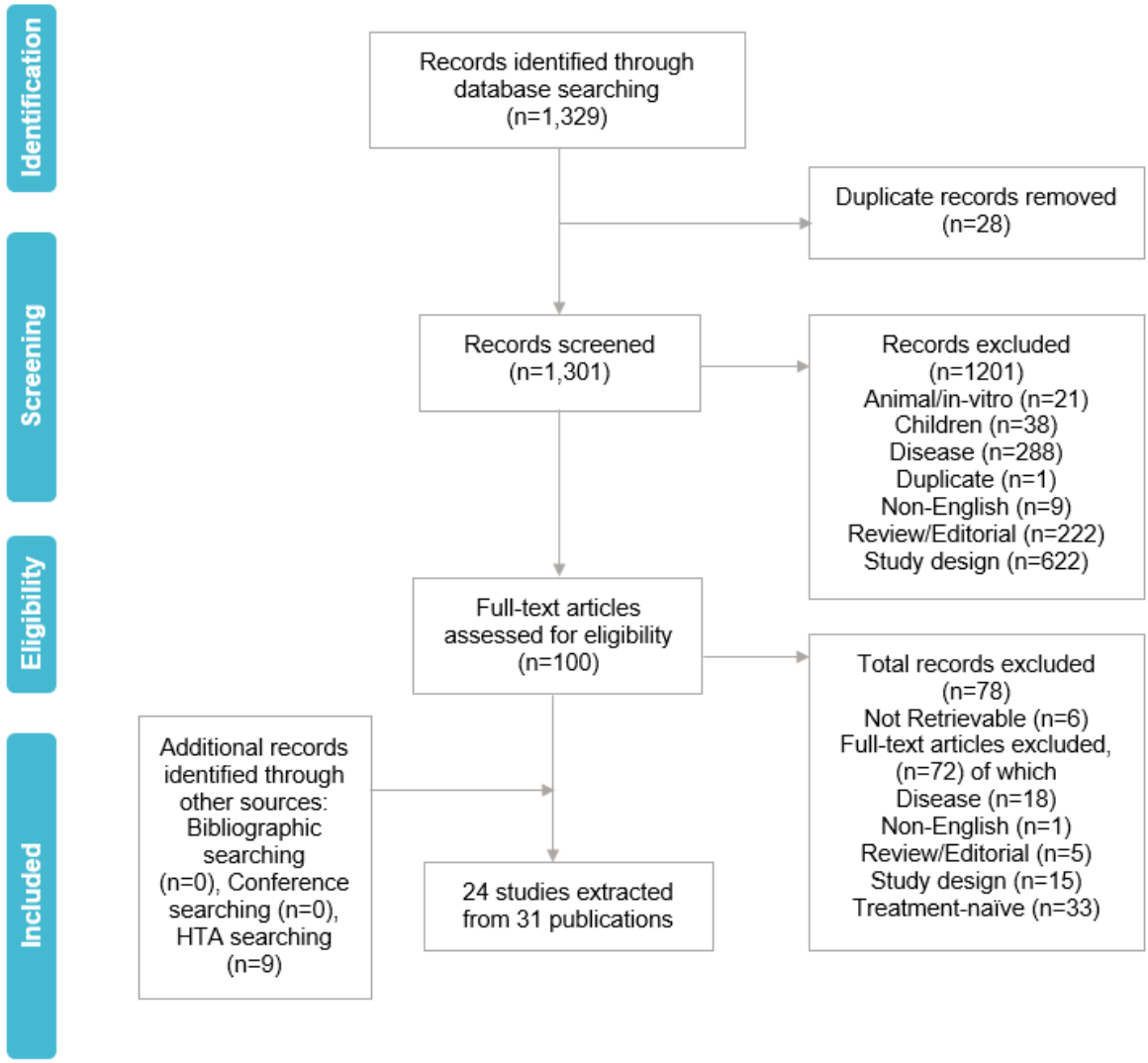
- MEDLINE® In-Process (using Pubmed.com)
- Embase® and MEDLINE® (using Embase.com)
- EconLit® (using Ebsco.com)
- The Cochrane Library, including the following:
  - NHS Economic Evaluation Database (NHS EED)
  - Health Technology Assessment Database (HTAD)

Note: Electronic searches for NHS EED and HTAD were performed via the University of York Centre for Reviews and Dissemination (CRD) platform for records archived to 2015.

Additionally, conference proceedings from the last 2 years and data available on health technology assessment (HTA) websites were searched to identify recently completed or ongoing studies of interest.

The details for the studies are presented in Figure 20 using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.[102]

Figure 20: PRISMA flow diagram for cost-effectiveness studies



Key: HTA, health technology assessment; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

A summary of the published cost-effectiveness studies is presented in Table 23 and Table 24 presents the summary of previous NICE submissions.

Table 23: Summary list of published cost-effectiveness studies

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Guzauskas et al.[103]	2018	• Perspective: US payer perspective • Time horizon: Life time • Cycle length: 1 month	The model is made up of 3 health states: • PFS • PD • Death	• Median: 62	• G + Benda: 5.54 • Benda: 4.30	• G + Benda: $121,000 • Benda: $62,900	• G + Benda vs Benda: $47,000
Haukaas et al.[104]	2018	• Perspective: Norwegian healthcare payer • Time horizon: 20 years • Cycle length: 1 month	The model is made up of 3 health states: • PFS • PD • Death	• 62	• G + Benda: 4.67 • Benda: 3.65	• G + Benda: €98,849 • Benda: €51,570	• G + Benda vs Benda: €46,438
Wang et al.[38]	2018	• Perspective: NHS perspective • Time horizon: Lifetime • Cycle length: NR	A discrete event simulation model was constructed	• NR	• Second-line treatment[a] without SCT: 8.34 • Second-line treatment[a] with SCT: 12.15	• Second-line treatment[a] without SCT: £36,000 • Second-line treatment[a] with SCT: £60,261	• NR

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
AWMSG [Idelalisib][105]	2017	• Perspective: NHS/Personal Social Services perspective • Time horizon: 15-year • Cycle length: 1 week	The model is made up of 5 health states: • Pre-progression on treatment • Pre-progression off treatment • Post-progression • Palliative care • Death	• Median (range): 64 (33-87)	• Idelalisib: 3.00 • Standard chemo: 2.37 • BSC: 1.97	• Idelalisib: CIC • Standard chemo: £36,447 • BSC: £28,306	• Idelalisib vs Standard chemo: CIC • Idelalisib vs BSC: CIC
Guzauskas et al.[106]	2017	• Perspective: US payer perspective • Time horizon: Lifetime • Cycle length: NR	The model is made up of 3 health states: • PFS • PD • Death	• NR	<u>Progressed patients</u> • G-chemo: 1.78 • R-chemo: 2.01	<u>Cost for progression</u> • G + chemo: $75,650 • R + chemo: $85,585	• NR
SMC [Obinutuzumab][107]	2017	• Perspective: NR • Time horizon: NR • Cycle length: NR	The model is made up of 3 health states: • PFS • PD • Death	• NR	• G-Benda vs R-chemo: 1.53	• G-Benda vs R-chemo: £42,775	• G-Benda vs R-chemo: £27,988

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
SMC [Idelalisib][107]	2015	• Perspective: NHS and Personal Social Services • Time horizon: Lifetime (10 years) • Cycle length: NR	The model is made up of 3 health states: • Pre-progression • Post progression • Death, with an indirect health state of palliative care prior to death	• NR	• Idelalisib vs further chemo and/or rituximab retreatment: 0.35	• Idelalisib vs further chemo and/or rituximab retreatment: £22,217	• Idelalisib vs further chemo and/or rituximab retreatment: £62,653
[b]Blommestein et al.[108]	2014	• Perspective: Healthcare perspective • Time horizon: Lifetime (20 years) • Cycle length: 1 month	The model is made up of 3 health states: • PFS • After progression survival • Death	EORTC trial • Median (range): 55 (26-80) Real-world Mean (SD) • R: 61 (13) • Obs: 59 (12)	Scenario 1 • R: 7.84 • Obs: 6.46 Scenario 2 • R: 7.81 • Obs: 6.44 Scenario 3 • R: 8.65 • Obs: 6.54	Scenario 1 • R: €56,608 • Obs: €39,182 Scenario 2 • R: €100,424 • Obs: €67,756 Scenario 3 • R: €88,582 • Obs: €64,846	Effectiveness source: EORTC20981 trial and real-world data Scenario 1 • R vs Obs: €12,655 Scenario 2 • R vs Obs: €23,821 Scenario 3 • R vs Obs: €10,591

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
				Propensity matched Real-world Mean (SD) • R: 61 (13) • Obs: 59 (10)			
		• NR	• NR	Median (range) • R: 62 (30-92) • Obs: 60 (34-82)	• NR	• NR	Effectiveness source: population-based registry • R vs Obs: € 5,156
		• NR	• NR	Median • R: 61 • Obs: 61	• NR	• NR	Cost-effectiveness ratio Effectiveness source: two Dutch population-based registries (PHAROS and HemoBase) • R: Between € 3,614 and € 5,156

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Pink et al.[109]	2012	• Perspective: UK NHS costing perspective • Time horizon: 30 years • Cycle length: 1 month	The model is made up of 3 health states: • PFS • Progressed follicular lymphoma • Death	• NR	• R-CHOP Simulation: 4.23 Original:4.22 • CHOP Simulation: 3.70 Original: 3.33	• R-CHOP Simulation: £22,728 Original: £21,608 • CHOP Simulation: £17,355 Original: £14,722	R-CHOP vs CHOP • Simulation: £9,076 • Original: £7,721
Vandekerckhove et al.[110]	2012	• Perspective: UK NHS perspective • Time horizon: Lifetime • Cycle length: 1 month	The model is made up of 3 health states: • PFS • PD • Death	• NR	<u>Discounted</u> • Bortezomib + R: 9.76 • R: 8.41 <u>Undiscounted</u> • Bortezomib + R: 14.16 • R: 11.71	• Bortezomib + R: £56,362 • R: £37,701	• Bortezomib + R vs R: £13,774
Soini et al.[111]	2011	• Perspective: Health care provider perspective • Time horizon: Lifetime	The model is made up of 3 health states: • PFS • PD • Death	• NR	• R-CHOP-R: 5.21 • R-CHOP: 4.72 • CHOP: 3.90	• R-CHOP-R: €68,331 • R-CHOP: €59,521 • CHOP: €49,562	• R-CHOP-R vs. R-CHOP: €18,147 • R-CHOP-R vs. CHOP: €14,360

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		• Cycle length: 1 month					• R-CHOP vs. CHOP: €12,123
Deconinck et al.[112]	2010	• Perspective: French National Health Service perspective • Time horizon: Lifetime (30 years) • Cycle length: 1 month	The model is made up of 3 health states: • PFS • PD • Death	• Median (range): 54 (27-80)	• R: 4.72 • Obs: 3.68	• R: €71,314 • Obs: €62,251	• R vs Obs: €8,729
Papadakis et al.[113]	2010	• Perspective: UK National Healthcare System • Time horizon: Lifetime (25 year) • Cycle length: NR	The model is made up of 4 health states: • PF1 • PF2 • PD • Death	• NR	• NR	• Cost of 2nd line R: cost saving £198 • Cost of supportive care incurred at disease progression: cost saving £906	NR

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Capote et al.[114]	2008	• Perspective: Spanish National Health System • Time horizon: 10 years • Cycle length: 1 month	The model is made up of 3 health states: • PFS • PD • Death	• NR	• R: 4.11 • Obs: 3.26	<u>Cost per patient</u> • R: €22,458.20 • Obs: €14,432.14	• R vs Obs: €9,358.49
Chiattone et al.[115]	2008	• Perspective: Brazilian Private Healthcare System • Time horizon: Lifetime • Cycle length: NR	• NR	• NR	• R: 4.73 • Obs: 3.20	• R vs Obs: R$60,576	• R vs Obs: R$39,576

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Hayslip et al.[116]	2008	• Perspective: US healthcare system • Time horizon: 5 years • Cycle length: 6 months	The model is made up of 6 health states: • Disease-free survivor • Undergoing salvage treatment • Subsequent remissions • Refractory disease • Transplantation • Death	• Range: 65-70	• NR	• NR	Adjuvant R vs Obs • Discounted: $19,522 • Unadjusted: $16,586
Hornberger et al.[117]	2008	• Perspective: US societal • Time horizon: Lifetime (30 years) • Cycle length: 21 days	The model is made up of 3 health states: • PF • Time after progression • Death	• NR	• R-CVP: -0.19 • CVP: -0.22	• R-CVP: $34,466 • CVP: $36,610	• NR

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Kasteng et al.[118]	2008	• Perspective: healthcare provider • Time horizon: 30 years • Cycle length: 1 month	The model is made up of 3 health states: • PF • PD • Death	• NR	• R: 4.29 • Obs: 3.38	• R: €39,617 Obs: €28,156	R vs Obs: €12,584
SMC [Rituximab][119]	2006	• Perspective: NR • Time horizon: 30 years • Cycle length: 1 month	The model is made up of 3 health states: • PF • PD • Death	• NR	• R: 0.89	• R Lifetime cost: £21,600 per patient Net cost: £6,886	Cost-effectiveness ratio • R: £7,721

Study	Year	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
SMC [Ibritumomab tiuxetan][120]	2005	• Perspective: NR • Time horizon: 15 year • Cycle length: 1 month	• NR	• NR	• Ibritumomab tiuxetan vs conventional care[d]: 0.38	• Ibritumomab tiuxetan vs conventional care[d]: £8,535 per patient	• Ibritumomab tiuxetan vs conventional care[d]: £22,445

Key: AWMSG, All Wales Medicines Strategy Group; Benda, bendamustine; BSC, best supportive care; Chemo, chemotherapy; CHOP, cyclophosphamide, doxorubicin, vincristine and prednisone; CIC, commercial in confidence; CVP, cyclophosphamide, vincristine and prednisone; DHAP, dexamethasone, cytarabine and cisplatin; EORTC, European Organisation for Research and Treatment of Cancer; ESHAP, etoposide, methylprednisolone, cytarabine, cisplatin; ERG, evidence review group; G, Obinutuzumab; G-CSF, granulocyte colony-stimulating factor; ICER, incremental cost-effectiveness ratio; K-M, Kaplan–Meier, NHS, National Health Service; NR, not reported; Obs, observation; PD, progressed disease, PF1, progression-free survival at 1st line maintenance; PF2, progression-free survival at 2nd line; PFS, progression-free survival; PF, progression-free QALY, quality-adjusted life year; R, rituximab; SCT, stem cell transplant; SD, standard deviation; SMC, Scottish Medicines Consortium.
Note: [a]Second-line treatment consists of chemotherapy CHOP (-R), CVP (-R), chlorambucil (-R), bendamustine (-R), DHAP (-R), ESHAP (-R); radiotherapy; supportive care: G-CSF, transfusion.
[b]This study is linked to two other studies which have same objective and perspective, but patient size is different as source of effectiveness varies. Therefore, the data from link studies is reported in separate row.
[c]ERG modifications: amended discounting logic; increased cost of drug administration; revised calculation of relapsed treatment costs; inclusion of £5,000 per patient terminal care costs; replace projected overall and progression-free survival estimates with K-M estimates at 1500 days.
[d]Conventional care arm: composed of a range of other treatments including chemotherapy, radiotherapy and stem cell transplant
Scenario 1: Effectiveness based on trial efficacy; costs based on treatment protocol EORTC20981
Scenario 2: Effectiveness based on trial efficacy; costs based on matched real-world patients
Scenario 3: Effectiveness based on real-world evidence; costs based on matched real-world patients

Table 24: Summary list of published NICE submissions

Study	Year	Treatment setting	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
NICE [TA513][52]	2018	Treatment-naïve FL patients	• Perspective: NHS and Personal Social Services perspective • Time horizon: Lifetime (40 years) • Cycle length: 1 month	The model is made up of 3 health states: • PFS (on/off treatment) • PD (Early PD and late PD) • Death	• NR	<u>Disease progression < 2 years</u> • G-chemo + G: 0.28 • R-chemo + R: 0.42 <u>Progression > 2 years</u> • G-chemo + G: 2.49 • R-chemo + R: 2.65	<u>Progressive disease</u> Supportive care and subsequent treatment costs • G-chemo: £10,201 • R-chemo: £11,873	• NR
						<u>ERG assessment</u> <u>Discounted QALY gain</u> <u>Progression < 2 years:</u> • G-chemo + G: 0.24 • R-chemo + R: 0.33 <u>Progression > 2 years:</u> • G-chemo + G: 2.17 • R-chemo + R: 2.40	<u>ERG assessment</u> <u>Progressive disease</u> Supportive care and subsequent treatment costs • G-chemo: £ 9,762 • R-chemo: £11,455	<u>ERG assessment</u> • NR

Study	Year	Treatment setting	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
NICE [TA472][48]	2017	Refractory FL patients	• Perspective: NHS and Personal Social Services perspective • Time horizon: Lifetime (25 year) • Cycle length: 1 month	The model is made up of 3 health states: • PFS (on/off treatment) • PD • Death	• 62.06	• G-Benda + G: 4.23 • Benda: 2.92	• G-Benda + G: Data redacted • Benda: £23,889	• G- Benda + G vs Benda: Data redacted
NICE [TA 226][51]	2011	Treatment-naïve FL patients	• Perspective: NHS and Personal Social Services perspective • Time horizon: Lifetime (25 year) • Cycle length: 1 month	The model is made up of 4 health states: • PFS at 1st line maintenance (PF1) • PFS at 2nd line (PF2) • PD • Death	• NR	PF2 with R-chemo-R • R: 1.75 • Obs: 1.78 PF2 with chemo-Obs: • R: 0.10 • Obs: 0.17 Total progressed survival • R: 1.11 • Obs: 1.21	PF2 (with R-chemo-R and with chemo-Obs) • R: £38,571 • Obs: £38,246 PD (with R-chemo-R and with chemo-Obs) • R: £10,779 • Obs: £11,682	• NR

Study	Year	Treatment setting	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
NICE [TA137][121]	2008	Relapsed or refractory FL patients	• Perspective: NHS in England and Wales (for estimating costs) Perspective of the patient with values from the general public (for health outcomes) • Time horizon: Lifetime (30 years) • Cycle length: 1 month	The model is made up of 5 health states: • PF – in the induction setting • PF – in a maintenance setting • PF – but not in the induction or maintenance settings • PD • Death	Median (range) • R: 53 (29-76) • Obs: 55 (27-80)	2-arm model QALY: Discounted (Undiscounted) • R: 4.23 (4.72) • Obs: 3.33 (3.68)	2-arm model Cost: Discounted (Undiscounted) • R: £21,608 (£24,082) • Obs: £14,722 (£16,855)	2-arm model • R vs Obs: £7,721
						4-arm model QALY: Discounted (Undiscounted) • R-CHOP followed by R: 4.09 (4.56) • R-CHOP followed by Obs: 3.63 (4.01) • CHOP followed by R: 3.72 (4.13) • CHOP followed by Obs: 3.09 (3.40)	4-arm model Cost: Discounted (Undiscounted) • R-CHOP followed by R: £28,585 (£30,821) • R-CHOP followed by Obs: £23,054 (£25,189) • CHOP followed by	4-arm model • CHOP followed by R vs CHOP followed by Obs: £9,076 • R-CHOP followed by R vs CHOP followed by R: £16,749 • R-CHOP followed by R vs R-CHOP followed by Obs: £11,904 • R-CHOP followed by R

Study	Year	Treatment setting	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
							R: £22,389 (£25,005) • CHOP followed by Obs: £16,658 (£18,728)	vs CHOP followed by Obs: £11,910 • CHOP followed by R vs R-CHOP followed by Obs: Dominant • R-CHOP followed by Obs vs CHOP followed by Obs: £11,916

						ERG assessment (4-arm model) ERG modifications[a] but using original outcome projections • R-CHOP followed by R vs CHOP followed by R: 0.37 • R-CHOP followed by R vs R-CHOP followed by Obs: 0.47 • R-CHOP followed by R vs CHOP followed by Obs: 1.00 • CHOP followed by R vs R-CHOP followed by Obs: 0.10 • CHOP followed by R vs CHOP followed by Obs: 0.63 • R-CHOP followed by Obs vs CHOP followed by Obs: 0.54 ERG modifications including K-M outcome estimates • R-CHOP followed by R vs CHOP followed by R: 0.20 • R-CHOP followed by R vs R-CHOP followed by Obs: 0.17	ERG assessment (4-arm model) ERG modifications[a] but using original outcome projections: • R-CHOP followed by R vs CHOP followed by R: £8,849 • R-CHOP followed by R vs R-CHOP followed by Obs: £7,686 • R-CHOP followed by R vs CHOP followed by Obs: £12,149 • CHOP followed by R vs R-CHOP followed by Obs: -£1,163	ERG assessment (4-arm model) ERG modifications[a] but using original outcome projections: • R-CHOP followed by R vs CHOP followed by R: £23,882 • R-CHOP followed by R vs R-CHOP followed by Obs: £16,509 • R-CHOP followed by R vs CHOP followed by Obs: £12,108 • CHOP followed by R vs R-CHOP followed by Obs: -£12,232 • CHOP followed by R vs CHOP

						• R-CHOP followed by R vs CHOP followed by Obs: 0.47 • CHOP followed by R vs R-CHOP followed by Obs: 0.03 • CHOP followed by R vs CHOP followed by Obs: 0.27 • R-CHOP followed by Obs vs CHOP followed by Obs: 0.30	• CHOP followed by R vs CHOP followed by Obs: £3,300 • R-CHOP followed by Obs vs CHOP followed by Obs: £4,463 <u>ERG modifications including K-M outcome estimates</u> • R-CHOP followed by R vs CHOP followed by R: £8,660 • R-CHOP followed by R vs R-CHOP followed by Obs: £7,289 • R-CHOP followed by R vs CHOP followed by Obs: £12,157	followed by Obs: £5,214 • R-CHOP followed by Obs vs CHOP followed by Obs: £8,298 <u>ERG modifications including K-M outcome estimates:</u> • R-CHOP followed by R vs CHOP followed by R: £42,982 • R-CHOP followed by R vs R-CHOP followed by Obs: £42,192 • R-CHOP followed by R vs CHOP followed by Obs: £25,978 • CHOP followed by R vs R-CHOP followed by Obs: £47,734

Study	Year	Treatment setting	Model settings	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
							• CHOP followed by R vs R-CHOP followed by Obs: -£1,371 • CHOP followed by R vs CHOP followed by Obs: £3,497 • R-CHOP followed by Obs vs CHOP followed by Obs: £4,867	• CHOP followed by R vs CHOP followed by Obs: £13,122 • R-CHOP followed by Obs vs. CHOP followed by Obs: £16,488

Key: Benda, bendamustine; Chemo, chemotherapy; CHOP, cyclophosphamide, doxorubicin, vincristine and prednisone; ERG, evidence review group; FL, follicular lymphoma; G, Obinutuzumab; ICER, incremental cost-effectiveness ratio; K-M, Kaplan–Meier, NHS, National Health Service; NICE, National Institute for Health and Care Excellence; NR, not reported; Obs, observation; PD, progressed disease, PFS, progression-free survival, PF, progression-free; QALY, quality-adjusted life year; R, rituximab; TA, technology appraisal.
Note: [a]ERG modifications: amended discounting logic; increased cost of drug administration; revised calculation of relapsed treatment costs; inclusion of £5,000 per patient terminal care costs; replace projected overall and progression-free survival estimates with K-M estimates at 1500 days.

No economic studies that included R^2 were identified. Therefore, a de novo analysis was required for this analysis. The majority of economic evaluations found in the economic SLR included a model structure based around progression with three health states (or sub sets of): progression-free, progressive disease and death (see Table 23). In particular, all of the previous NICE submissions identified in Table 24 had model structures using progression ranging from three to five health states. In Table 23, four economic evaluations were identified that had a UK perspective and were of potential value to inform this submission (Wang et al.[122], Pink et al.[109], Vandekerckhove et al.[110], and Papadakis et al.[113]). More details of how these evaluations have informed the de novo analysis are discussed in Section B.3.2.

B.3.2. Economic analysis

Patient population

The patient population considered in the model is in line with the proposed licence, that is, adult patients with previously treated FL or MZL. Due to the similar prognosis of FL and MZL patients, and the difficulty in sourcing MZL-specific data, FL and MZL populations were pooled throughout the economic analysis (see Section B.1.3 and B.2.9).

The model is split into two subpopulations: non-R-refractory and R-refractory. R-refractory patients have different comparators to patients who are non-R-refractory, and the data sources used to inform the efficacy of R^2 are different between these populations (see Section B.3.3).

The patient cohort considered in the model varies per population. Given the limitations of the data sources for the comparators (discussed in Section B.2.9), each comparator is considered individually using the most appropriate data source. MAICs have been used to match the R^2 population to the comparator population; this is necessary for comparator data where IPD are not available (i.e. O-Benda). For those comparators that use HMRN data, this was the preferred approach by the Committee and the ERG in the Idelalisib submission (ID1379).[123] Consequently, baseline patient characteristics in the model vary depending on the comparator efficacy source used. As discussed in Section B.2.9 the efficacy data for R-CHOP and R-CVP from HMRN are associated with small patient numbers but are shown to

perform similarly, and their efficacy is therefore taken from their pooled HMRN data. Efficacy data for O-Benda are derived from the Phase III GADOLIN trial.

The patient starting age and gender are matched to the data source used for the comparator arms. Body surface area (BSA) data are taken from individual patients in the AUGMENT study. A summary of the patient characteristics per comparison is presented in Table 25.

Table 25: Model patient characteristics

Comparison: R^2 vs	Mean or median age (years)[a]	% Female[a]	Mean BSA (m^2)[b]
R-CHOP/R-CVP	[redacted]	[redacted]	1.85
O-Benda	63.0	44.5%	

Key: Benda, bendamustine; BSA, body surface area; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; HMRN, Haematological Malignancy Research Network; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a]Source for R-CHOP and R-CVP is HMRN data.[124] Source for O-Benda is from GADOLIN trial.[64]
[b]Source is from AUGMENT.

Model structure

As discussed in Section B.3.1, the majority of economic evaluations found in the SLR included a model structure based around progression with three health states (or subsets of): progression-free, progression and death (see Table 23 and Table 24).

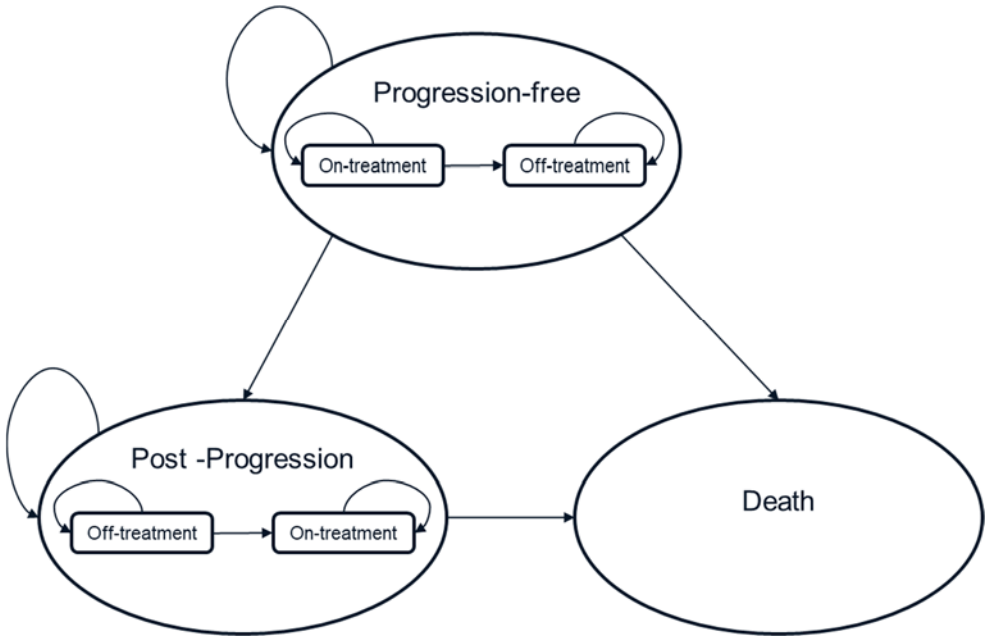
Wang et al. used a discrete event simulation model that was designed to capture the real-word treatment strategies for patients with previously untreated FL.[38] The model consisted of multiple treatment pathways using data from HMRN to inform the probabilities of each patient moving through each stage of the model. Although this model is appropriate for estimating costs of real-world treatment pathways for FL patients, various data inputs are required, as are assumptions on the 'standard' treatment pathway to be modelled. Patients with FL and MZL may be treated with many different treatments over the course of their disease; therefore, capturing each possible treatment pathway is complex. Given limitations in the available evidence base, and the remit of the single technology's appraisal (STA) being to focus on the

cost effectiveness of a specific intervention rather than attempting to identify optimal treatment pathways, this model structure was not pursued.

In line with the SLR findings, prior NICE submissions for FL and the primary endpoint of the AUGMENT study, the economic model used to evaluate the cost effectiveness of R^2 for previously treated FL and MZL patients has been structured around progression.[48, 52, 121]

A cohort-level model was developed with three health states: progression-free (PF), post-progression (PP) and death. The PF and PP health states include sub health states based on whether patients are on or off treatment as defined according to time on treatment (ToT) and TTNLT (Figure 21).

Figure 21: Cost-effectiveness model structure



Partitioned survival versus state transition models

Both partitioned survival model (PSM) and state transition model (STM) structures and the data sources available to inform them were considered in line with NICE DSU TSD 19.[125] PSMs allow the proportion of patients in each health state to be defined by the individual survival curves extrapolated from the study data or hazard ratios. This structure is most commonly used in oncology models and is an established method with straightforward implementation and explanation.[125] It does

not require the definition of explicit transitions between health states and automatically incorporates time dependencies in the event rates.

Alternatively, STMs incorporate an explicit link between clinical endpoints and allow sensitivity around the post-progression survival (PPS) inputs given the uncertainty within the study data. However, the data required to inform a STM are lacking. The relevant comparators for this submission are not included in the head-to-head study with R^2 and so the data available for informing PPS for these treatments are reduced to available published data or alternative sources.

The ERG for TA472 considered the company's approach to modelling using an STM unreliable due to the discrepancies between the predicted and observed data. They consequently amended the model to a PSM, acknowledging the limitations and immaturity of the data. However, treatment effect assumptions were incorporated so that extrapolated outcomes were considered more clinically plausible.[48]

In this analysis, given the limitations to both approaches, the PSM has been selected as the most appropriate model structure. Not only is this model structure more common within the oncology setting, but it makes use of the PFS and OS data directly, ensuring that estimated survival outcomes versus observed outcomes are matched. This structure additionally allows the use of 'sub health states' without the complication of modelling each individual and possible transition, for which data are not always available (i.e. on/off treatment within the PF and PP health states). Clinical validation was sought to ensure that the most clinically plausible curves were selected for both PFS and OS in the base case, in addition to treatment effect scenarios to explore long-term uncertainty. Validation of the model outcomes was also conducted and is described in Section B.3.10.

How patients move through the model

In each cycle, patients can either remain in their current health state or progress to a subsequent health state. All patients start in the PF health state; here, patients start 'on treatment', then either remain on treatment or come off treatment before progressing per cycle. From PF, patients can either progress and move to the PP health state or die. Within PP, patients can have a treatment-free interval before receiving subsequent therapy. As discussed in Section B.2.6 patients who progress

via radiographic scans may not necessarily receive subsequent treatment immediately if not symptomatic, resulting in TTNLT being longer than progression. Therefore, in terms of costs of subsequent treatments, patients' quality of life and impact on OS, the time in which patients receive their next treatment is a more relevant endpoint than the time at which progression is documented. From PP, patients can either remain in the PP state or die in each cycle. The proportions of patients within each health state (and sub health state) are calculated as follows:

- PF (on-treatment) = ToT data
- PF (off-treatment) = PFS – ToT
- PP (off-treatment) = TTNLT – PFS
- PP (on-treatment) = OS – TTNLT
- Death = 1 – OS

Modelling utility

Utilities for each health state are based on the observed EQ-5D-3L data from AUGMENT, with published literature used in scenario analysis.

Modelling drug cost

All treatments are modelled in line with the current summary of product characteristics (SmPC) indications using ToT data to inform the proportion of patients receiving each dose per cycle or mean number of doses.

Modelling subsequent therapies

Subsequent treatments are costed as one-off costs when patients enter the PP (on-treatment) health state. Subsequent treatment costs are based on subsequent treatment usage in AUGMENT and HMRN post initial therapy.

Modelling resource use

Resource use costs were defined according to the length of time patients were within the PF or PP health state. Resource information was based on the European Society for Medical Oncology (ESMO) guidelines[14] and resource estimates from previous FL submissions.[48, 52] AE costs were calculated as one-off costs applied at the first cycle of the model and based upon AUGMENT, MAGNIFY or literature sources. End-of-life costs were also applied to each patient upon death.

Table 26 summarizes the key features of the economic analysis in comparison with previous appraisals in the same disease area within the relapsed/refractory setting: TA137 and TA472.[48, 121] As per the NICE reference case, all health effects were measured in quality-adjusted life years (QALYs), a 3.5% discount rate was used for utilities and costs, and the perspective is of the NHS and Personal Social Services (PSS).

A cycle length of 28 days is used in the model. This captures the majority of the treatment cycle lengths that are included in the model and was considered sufficiently short to accurately capture key clinical outcomes and dosing regimens, given a 40-year time horizon. Half cycle-correction is also applied as this accounts for any events that happen during the model cycle.

Table 26: Features of the economic analysis

Factor	Previous appraisals		Current appraisal	
	TA137[121]	**TA472[48]**	**Chosen values**	**Justification**
Time horizon	30 years. Lifetime horizon.	25 years. Lifetime horizon, less than 1% patients alive at 25 years.	40 years	General population survival was modelled based on ONS life tables[126] for the age-matched patient populations using the mean age (63-65 from Table 25). All patients have died by the end of the time horizon.
Treatment waning effect	The treatment effect was limited to 5 years in the model, whereby the monthly risk of disease progression or death was equivalent to the observation group.	The revised post-consultation base case assumed that treatment effect would last 5.5 years	5 years has been chosen as the base case with other time points varied in scenario analysis.	This is consistent with the previous appraisals accepted by NICE.

Factor	Previous appraisals		Current appraisal	
	TA137[121]	TA472[48]	Chosen values	Justification
Source of utilities	PF and PD utilities sourced from the Oxford outcomes study, 2005 (same study as Wild et al., 2006[127])	PF and PD utility values were taken from Wild et al., 2006[127], sourced from a systematic literature review	Utilities derived from the AUGMENT EQ-5D-3L data. Published literature utilities used in scenario analysis.	Utilities collected from a relevant patient population and used to inform specific health states to the model.
Source of costs	Routine management costs were based on outpatient visits to avoid double counting of costs captured in other sections. An outpatient visit every 3 months for PF and monthly for PD.	Assumptions and ESMO guidelines informed the frequencies for disease management costs. PF was split by 0-6 months, 6-30 months and 30+ month periods.	Similar frequencies were used as per TA472. The timings of the PF frequencies changed depending on the treatment selected (see Section B.3.5 for further explanation)	These were chosen to be consistent with the most recent NICE submission in a similar patient population and are informed by the ESMO guidelines.

Key: EQ-5D-3L, 3-level EQ-5D; ESMO, European Society for Medical Oncology; NICE, National Institute for Health and Care Excellence; ONS, Office for National Statistics; PD, progressed disease; PF, progression-free; PFS, progression-free survival; TA, technology appraisal.

Intervention technology and comparators

The R^2 dosing regimen within the model is lenalidomide 20 mg orally once daily on Days 1–21 of repeated 28-day cycles for up to 12 cycles of treatment. Rituximab is given as 375 mg/m^2 every week in Cycle 1 (Days 1, 8, 15 and 22) and Day 1 of every 28-day cycle for Cycles 2–5. This is in line with the recommended dose in the SmPC.[4] Patients with moderate renal impairment start on a dose of 10 mg of lenalidomide if CrCl is ≥30 ml/min but <60 ml/min. These criteria were met by [redacted] of patients in AUGMENT and [redacted] in MAGNIFY (R-refractory population), and these proportions are used to inform the starting dose in the model for the non-R-refractory and R-refractory populations, respectively.

The comparators considered in the model are listed below and the dosing schedules in the base case are in line with their market authorization where possible (see Section B.1.3 for justifications of chosen comparators).

Scenarios were investigated that amended the dosing to be in line with the efficacy source if this was different to the SmPC. These are detailed in Table 27.

Table 27: Model intervention and comparators with dosing schedules (induction)

Comparator	Treatment	Dosing (base case)			Source
		Dose	Cycle length	Duration	
R[2]	Lenalidomide	20 mg Days 1–21 or 10 mg if CrCl ≥30 and <60	28 days	Up to 12 cycles	SmPC[4] MAGNIFY[128] (scenario)
	Rituximab	375 mg/m^2 Day 1, 8, 15 and 22 in Cycle 1, Day 1 in Cycle 2–5 (Cycle 3, 5, 7, 9 and 11 in scenario)			
R-CHOP	Rituximab	375 mg/m^2 Day 1	21 days	Up to 8 cycles	SmPC[129]
	Cyclophosphamide	750 mg/m^2 Day 1			South East London Cancer Network[130]
	Doxorubicin	50 mg/m^2 Day 1			
	Vincristine	1.4 mg/m^2 Day 1 (max 2 mg)			
	Prednisolone	100 mg Day 1–5			
R-CVP	Rituximab	375 mg/m^2 Day 1	21 days	Up to 8 cycles	SmPC[129]
	Cyclophosphamide	750 mg/m^2 Day 1			South East London Cancer Network[131]
	Vincristine	1.4 mg/m^2 Day 1 (max 2 mg)			
	Prednisolone	100 mg Day 1–5			
O-Benda	Obinutuzumab	1,000 mg days 1, 8 and 15 cycle 1, day 1 cycle 2-6	28 days	Up to 6 cycles	SmPC[132]
	Bendamustine	90 mg/m^2 days 2 and 3 cycle 1 and days 1 and 2 cycles 2-6.			TA472[48]
Key: CrCl, creatine clearance; Q1W, every week; SmPC, summary of product characteristics; TA, technology appraisal.					

B.3.3. Clinical parameters and variables

Time-to-event data are used to inform the proportion of patients in each health state over the time horizon of the model. The data sources used to inform each time to event for each comparison were chosen based on the availability of endpoints and robustness of the comparison. As discussed in Section B.2.9 two sources of efficacy were available for R-CHOP in the relapsed/refractory setting; van Oers and HMRN.[63, 124] Both sources only had FL patients, and despite van Oers being a Phase III study and therefore usually the preferred source of evidence for a matched comparison, there were concerns over the age and population of the study. The van Oers study was conducted before the rituximab era; therefore, all patients in the study were rituximab naïve. Prior rituximab exposure is an important effect modifier, as patients that are rituximab naïve are expected to have better response rates than rituximab-sensitive patients.[88] Thus, failure to match on this covariate biases against R^2. Furthermore, as nearly all patients receive rituximab upfront in UK clinical practice, matching to this study would not be reflective of clinical practice.[133] In addition, although PFS and OS were reported for R-CHOP, no data were reported for TTNLT, a key clinical endpoint reflective of a patients requirement for further treatment. The availability of TTNLT data was vital for the economic analysis as it was deemed the most appropriate endpoint for capturing subsequent treatment costs, quality-of-life impact and OS impact, as discussed in Section B.3.2. In addition, no trial-based evidence was found for R-CVP. HMRN data were therefore considered more appropriate for the base case, as these data were reflective of the current UK population and included PFS, OS and TTNLT data, which could be used in the model.

The HMRN data demonstrated that both OS and PFS outcomes were similar between R-CHOP and R-CVP (see Section B.2.9 and Appendix D). Clinical opinion suggests that in the relapsed/refractory setting, it would not be unreasonable to assume the efficacy of R-CHOP and R-CVP to be similar. Therefore, HMRN data for R-CHOP and R-CVP were pooled and used for both comparisons within the economic model. In addition to the evidence of similar efficacy, this also allowed the HMRN cohort to be of a reasonable size versus individual comparisons (63 versus 33). A scenario analysis with the ITC results of R^2 compared to R-CHOP using the

van Oers study has been presented with the assumption that this also represents the efficacy of R-CVP. Details of how this scenario was derived are presented in Appendix T.

For the comparison with O-Benda, GADOLIN was considered the most robust source of evidence and contained all the outputs required for the economic model.[48, 64]

All time-to-event data have been extrapolated using parametric survival distributions: exponential, generalized gamma, Gompertz, log-normal, log-logistic and Weibull. The selection of the most appropriate distribution for the base case has been made in accordance with NICE TSD 14.[134] Firstly, log cumulative hazard plots have been produced to evaluate how the hazards change over time and whether the proportional hazards assumption holds between the two treatment arms. Secondly, visual inspection and comparison of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) were then used to compare which distribution fits the KM data best. Thirdly, clinical validation was used to ensure that the final extrapolated curve was clinically plausible. Finally, the overall selected distributions of all the endpoints were reviewed to ensure that no implausible curves crossing (e.g. TTNLT with OS) were possible. Given a patient's mean starting age in the model (63-65 years), 15 years was determined to be a reasonable time point during which curves crossing would be considered implausible. Curves were therefore inspected for crossing up to a cut-off of 15 years, and any curve selections that did show evidence of crossing before this time point were deemed implausible choices.

Treatment effect

The treatment waning effect is considered in the model and assumes that any treatment effect of R^2 only lasts up to 5 years. After this time point, the comparator hazard of progressing or dying is applied to the R^2 arm. A time point of 5 years was selected as the base case as this is consistent with previous NICE submissions in the same disease area. In TA472, the company assumed that the treatment effect of O-Benda would last 5.5 years, in line with the longest follow-up from the GADOLIN study.[48] The duration of treatment benefit for rituximab in TA137 was assumed to be 5 years, which the Committee felt was reasonable.[121] As there is no evidence to

suggest for how long the exact duration of treatment effect would be expected to last, other time points are tested in scenario analyses.

Overall survival

R^2 versus R-CHOP/R-CVP

For R^2 in comparison with R-CHOP/R-CVP, data from AUGMENT were matched to pooled data from HMRN for R-CHOP and R-CVP (see Section B.2.9). The OS KMs for R^2 (re-weighted) and R-CHOP/R-CVP are shown in Figure 22.

Figure 22: OS – R^2 (re-weighted) versus R-CHOP/R-CVP

Figure 23 presents the corresponding log-cumulative hazard versus time plot. The lines meet at approximately 500 days and then diverge thereafter, suggesting that the proportional hazards assumption is not appropriate. Additionally, as described in the NICE DSU TSD 14, it is unnecessary to rely upon the proportional hazards assumption when IPD are available.[134] As such, stratified statistical models have been used.

Figure 23: OS: log-cumulative hazard versus time plot – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Table 28 presents the AIC and BIC fit statistics for each distribution. For R^2, exponential is statistically the best fitting; for R-CHOP/R-CVP, these are exponential or log-logistic with the difference between the fit statistics being minimal. All curves appear to fit the KM data reasonably well for R^2 (see Figure 24). For R-CHOP/R-CVP, the curves slightly over- or under-estimate the observed data due to the

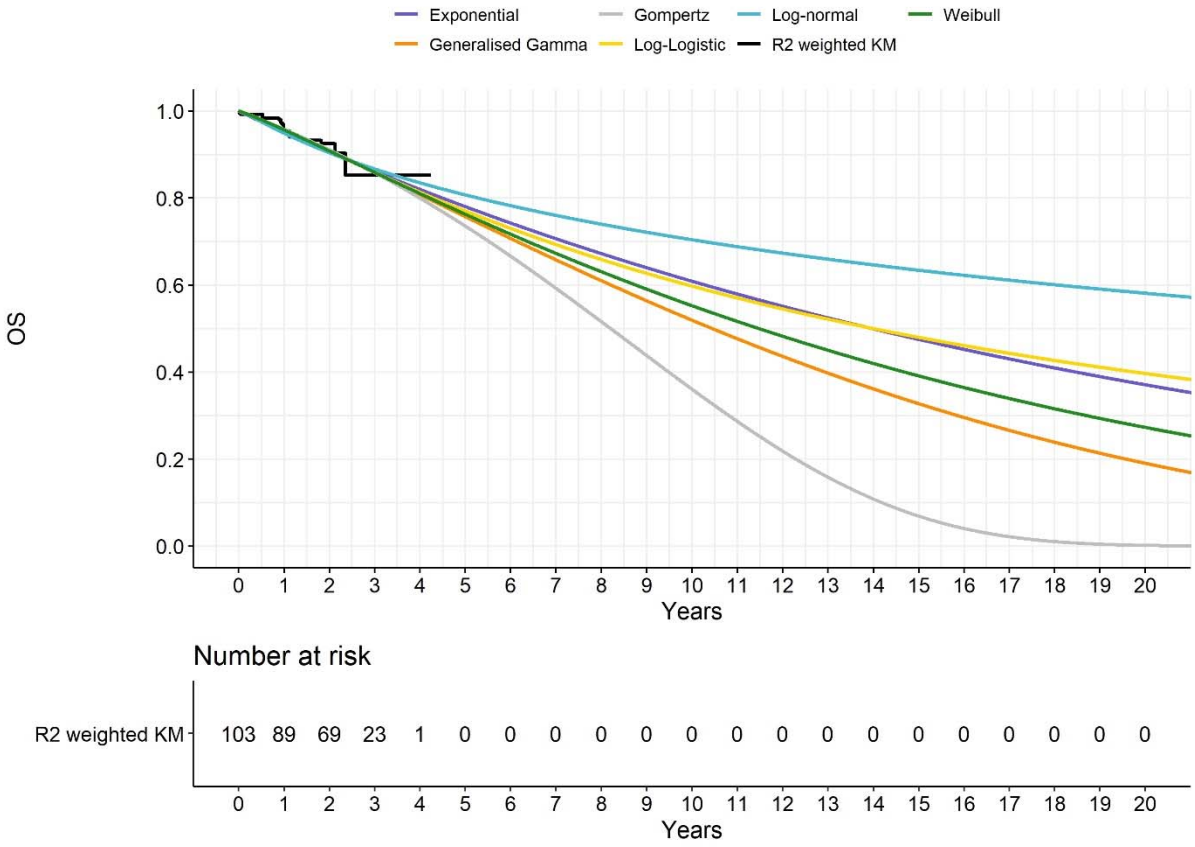
'stepped' nature of the KM data (see Figure 25). For R-CHOP/R-CVP, the curves estimate that OS at 20 years is between 16% and 32%. Given that the AIC/BIC for Weibull suggests it is a reasonable fit, and this distribution was considered appropriate for the relapsed/refractory setting in TA137 (the same population as this appraisal) this has been used for the base case. The same distribution has also been selected for R^2 to estimate OS.

Table 28: OS: AIC and BIC – R^2 versus R-CHOP/R-CVP

Distribution	R^2		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	**232.53**	**235.71**	598.43	**600.58**
Generalized gamma	236.30	245.84	600.05	606.48
Gompertz	234.25	240.62	598.05	602.34
Log-logistic	234.38	240.74	**598.03**	602.32
Log-normal	236.29	242.65	598.18	602.47
Weibull	**234.33**	**240.70**	**599.59**	**603.88**

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: Bold = statistically best fit; **Purple** = selected for base case.

Figure 24: OS parametric curves for R^2 in the non-rituximab refractory population



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 25: OS parametric curves for R-CHOP/R-CVP in the non-rituximab refractory population

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

The curves are adjusted to use general population mortality if the hazard of death is greater than the age- and gender-matched general population estimated from the ONS life tables[126], at any time point. The final OS curves used for the model base case are presented in Figure 26, demonstrating the curves at each stage of adjustment (unadjusted parametric curve → adjustment for treatment waning → general population mortality adjustment).

Figure 26: OS: Final curves – R^2 versus R-CHOP/R-CVP

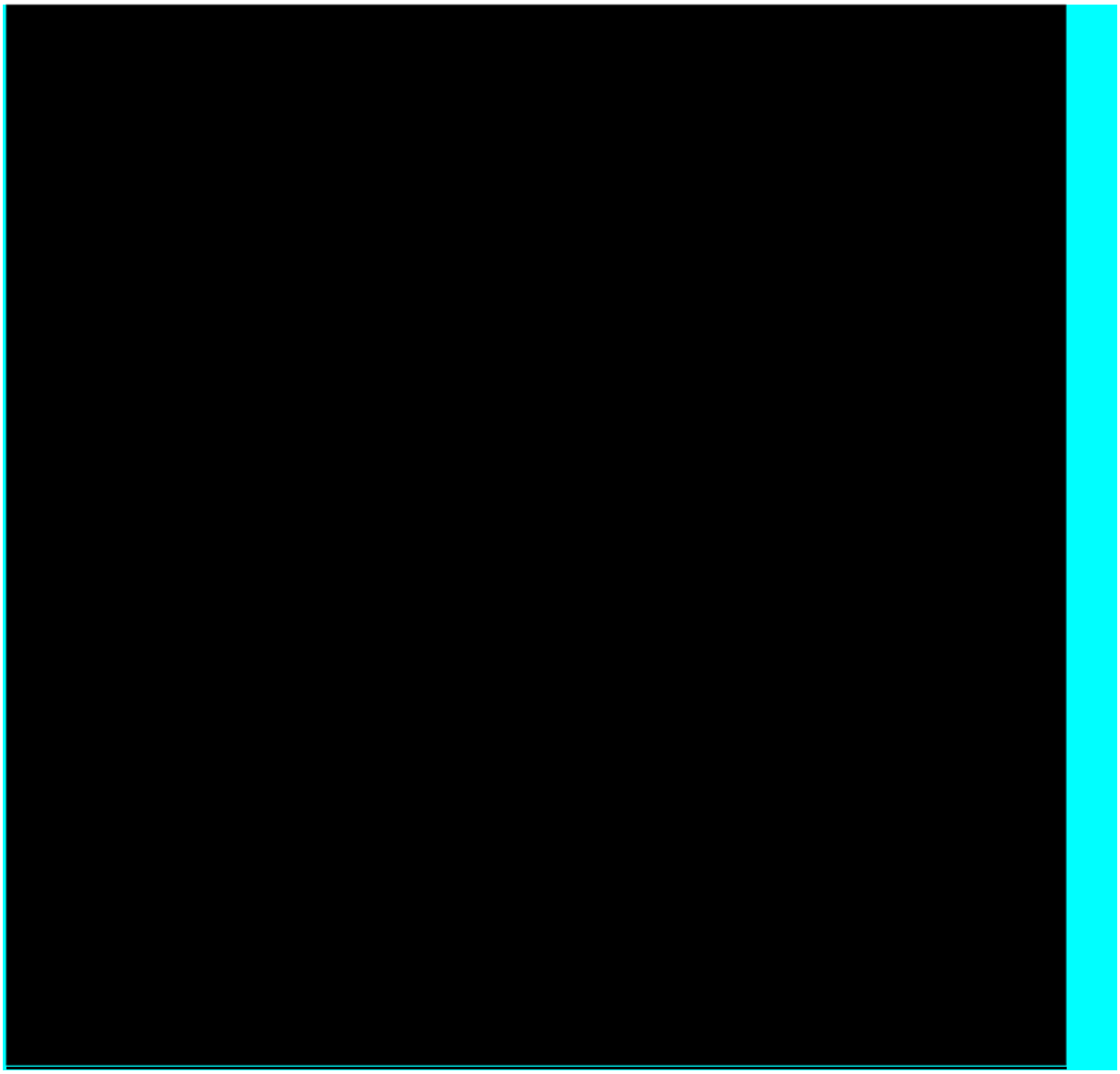
Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

R^2 versus obinutuzumab + bendamustine

For R^2 in comparison with O-Benda, OS data for the R-refractory population from the induction phase of MAGNIFY was matched to the O-Benda arm in the GADOLIN study (see Section B.2.9).

The corresponding KM curves overlap (see Figure 27). Therefore, equivalent OS has been assumed for the base case analysis. As the GADOLIN study provides an additional 4 years of follow-up for informing long-term extrapolation compared with MAGNIFY, parametric survival curves were fit to the O-Benda arm. These parametric survival curves were then assumed to apply to both the R^2 and O-Benda arms. This assumption was explored in the scenario analysis by fitting parametric curves to the matched R^2 arm to inform OS for the R^2 arm in the model (Appendix U).

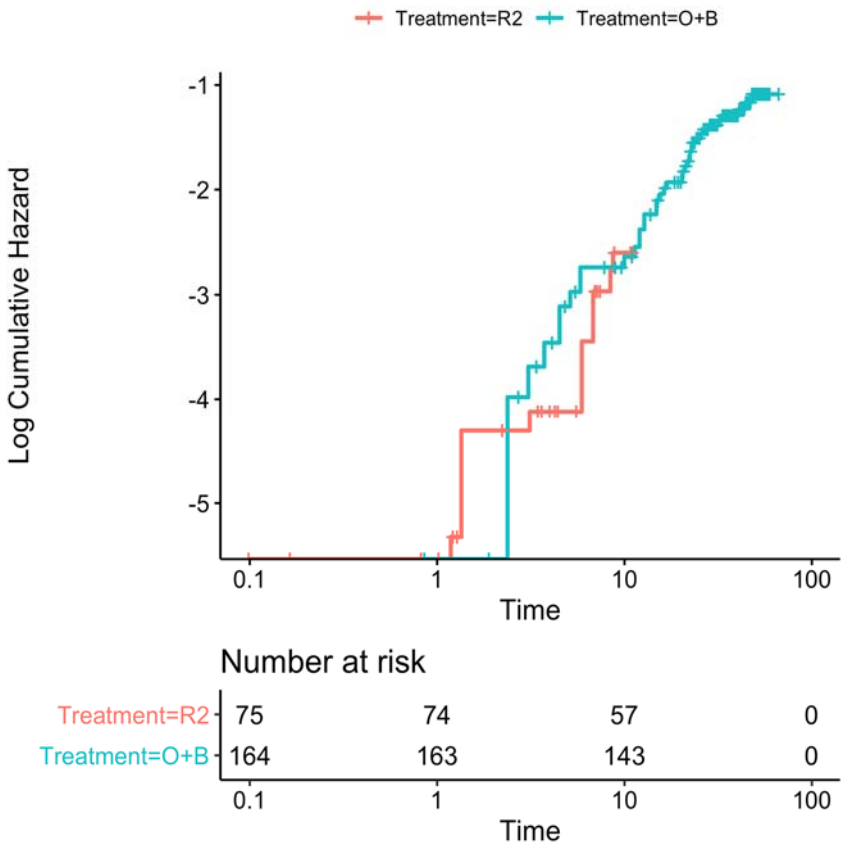
Figure 27: OS – matched R^2 versus O-Benda[84]



Key: Benda, bendamustine; O, obinutuzumab; OS, overall survival; R^2, lenalidomide plus rituximab.

Figure 28 presents the corresponding log-cumulative hazard versus time plot. The lines initially cross then come together at approximately 9 months, suggesting that proportional hazards assumption is not appropriate. As such, stratified statistical models have been used.

Figure 28: OS: log-cumulative hazard versus time plot – matched R^2 versus O-Benda



Key: Benda, bendamustine; O, obinutuzumab; OS, overall survival; R^2, lenalidomide plus rituximab.

Table 29 presents the AIC and BIC fit statistics for each distribution for O-Benda. Exponential or generalized gamma are statistically the best fitting. All curves appear to fit the KM data reasonably well, with a slight over-estimation for O-Benda between 1.8 to 3.3 years, but fan out after the study data, indicating different long-term estimates (see Figure 29). As R-refractory patients are expected to have worse prognosis than non-R-refractory patients, the survival estimated for O-Benda using the generalized gamma distribution seems optimistic. Therefore, based on BIC, visual fit and clinical plausibility, exponential has been chosen for the base case.

Table 29: OS: AIC and BIC – O-Benda

Distribution	O-Benda	
	AIC	BIC
Exponential	**455.64**	**458.74**
Generalized gamma	**453.40**	462.70
Gompertz	454.77	460.97
Log-logistic	456.02	462.22
Log-normal	453.47	459.67
Weibull	457.33	463.52

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; Benda, bendamustine; O, obinutuzumab; OS, overall survival.
Bold = statistically best fit; **Purple** = selected for base case.
Note: For the base case, R^2 is assumed to have equal OS to O-Benda.

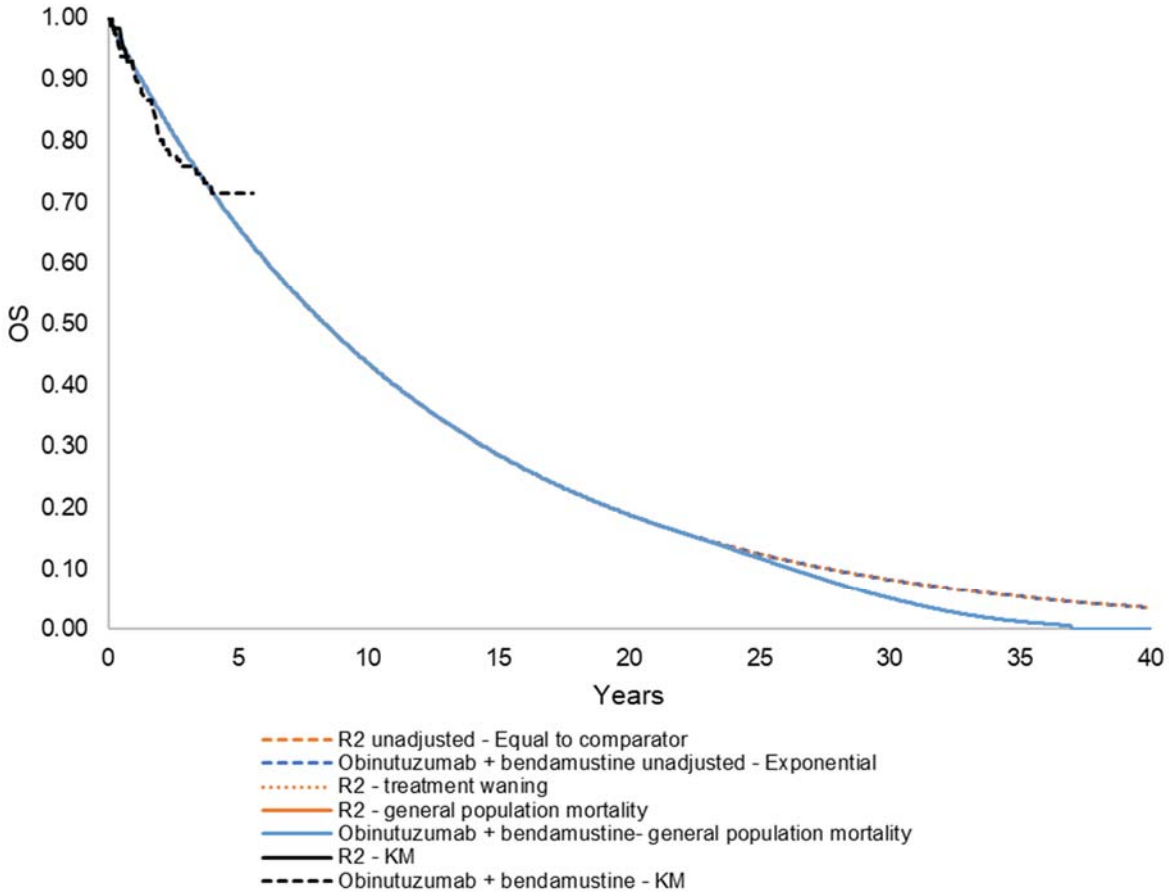
Figure 29: OS parametric curves for O-Benda in the rituximab refractory population

Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; OS, overall survival; R^2, lenalidomide plus rituximab.

As per the comparison of R2 vs. R-CHOP/R-CVP, general population mortality was applied as a competing risk. The final base-case OS curves, demonstrating curves at each stage of adjustment, are presented in Figure 30. As parametric survival curves

are assumed to apply to both the R^2 and O-Benda arms at base case, only the extrapolated curves for O-Benda are visible.

Figure 30: OS: Final curves – R^2 versus O-Benda



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; OS, overall survival; R^2, lenalidomide plus rituximab.

Progression-free survival

R^2 versus R-CHOP/R-CVP

For R^2 in comparison with R-CHOP/R-CVP, as with the OS data, PFS data from AUGMENT (IRC with EMA censoring rules) were used and matched to the patient population from the R-CHOP/R-CVP cohort from HMRN (see Section B.2.9).

When comparing the associated KM curves for R^2 and R-CHOP/R-CVP, these appear to be diverge from initiation and then converge and overlap at approximately 800 days (see Figure 31), suggesting the relative treatment effect of R^2 vs. R-

CHOP/R-CVP is non-constant. This is supported by the log-cumulative hazard plot which is non-parallel (Figure 32). Therefore, in the base case analysis, PFS for R^2 was modelled using the KM data until the maximum follow-up (46.7 months), beyond which the comparator hazard was applied to extrapolate (see below). This approach ensures the relative treatment effect of R^2 vs. R-CHOP/R-CVP based on the MAIC is accurately reflected. Moreover, the extrapolation method utilises the longer follow-up from the HMRN dataset (11.6 years; additional 7.7 years vs. AUGMENT) to inform the hazard, which is conservatively applied to both arms over lifetime.

Figure 31: PFS – R^2 (re-weighted) versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 32: PFS: log-cumulative hazard versus time plot – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Parametric curves have been fitted to each arm for scenario analysis so that the impact of the assumptions used in the base case analysis on cost-effectiveness can be tested. Due to the non-proportional hazards described above, stratified statistical models have been used.

Table 30 presents the AIC and BIC fit statistics for each distribution. For R^2, exponential or log-logistic are statistically the best fitting; for R-CHOP/R-CVP Weibull is the best statistical fit. Based on visual inspection, all curves appear to fit the KM data reasonably well for R^2; however, the majority of the curves for R-CHOP/R-CVP either slightly under- or over-estimate the actual observed data (see Figure 33 and Figure 34). The Weibull and generalized gamma curves appear to fit the R-CHOP/R-CVP data best in comparison to the other distributions.

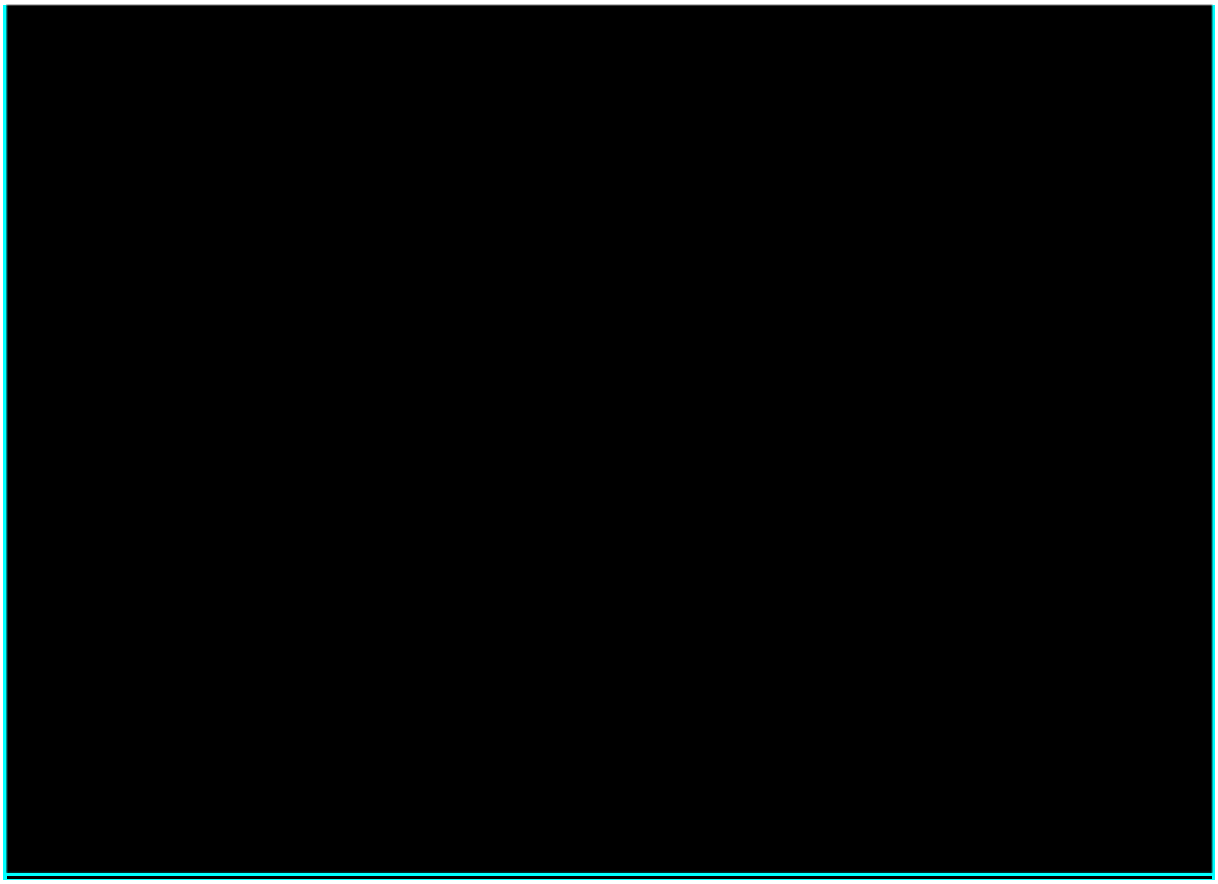
The Weibull PFS curve crossed the TTNLT curve at approximately 8 years for R-CHOP/R-CVP. In practice, it is unlikely that patients would start their next treatment prior to progressing, so this was considered implausible. Therefore, the second-best fitting curve based on AIC/BIC (generalized gamma) has been selected as the base case.

Table 30: PFS: AIC and BIC – R^2 versus R-CHOP/R-CVP

Distribution	R^2		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	609.50	**612.68**	696.16	698.30
Generalized gamma	610.69	620.24	**671.80**	**678.23**
Gompertz	610.80	617.17	684.73	689.01
Log-logistic	**607.88**	614.24	673.73	678.02
Log-normal	611.74	618.10	677.98	682.27
Weibull	609.02	615.38	**670.65**	**674.94**

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Bold = statistically best fit; **Purple** = selected for base case.
Note: For the base case, R^2 is assumed to follow KM data and then to match the comparator effectiveness.

Figure 33: PFS parametric curves for R^2 in the non-rituximab refractory population



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Note: For the base case, R^2 is assumed to assumed to follow the KM and then use the hazard of R-CHOP/R-CVP.

Figure 34: PFS parametric curves for R-CHOP/R-CVP in the non-rituximab refractory population

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

The curves are adjusted to ensure that the long-term PFS estimates are never predicted to be higher than TTNLT or OS. The final PFS curves, demonstrating the curves at each stage of adjustment, are presented in Figure 35.

Figure 35: PFS: Final curves – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

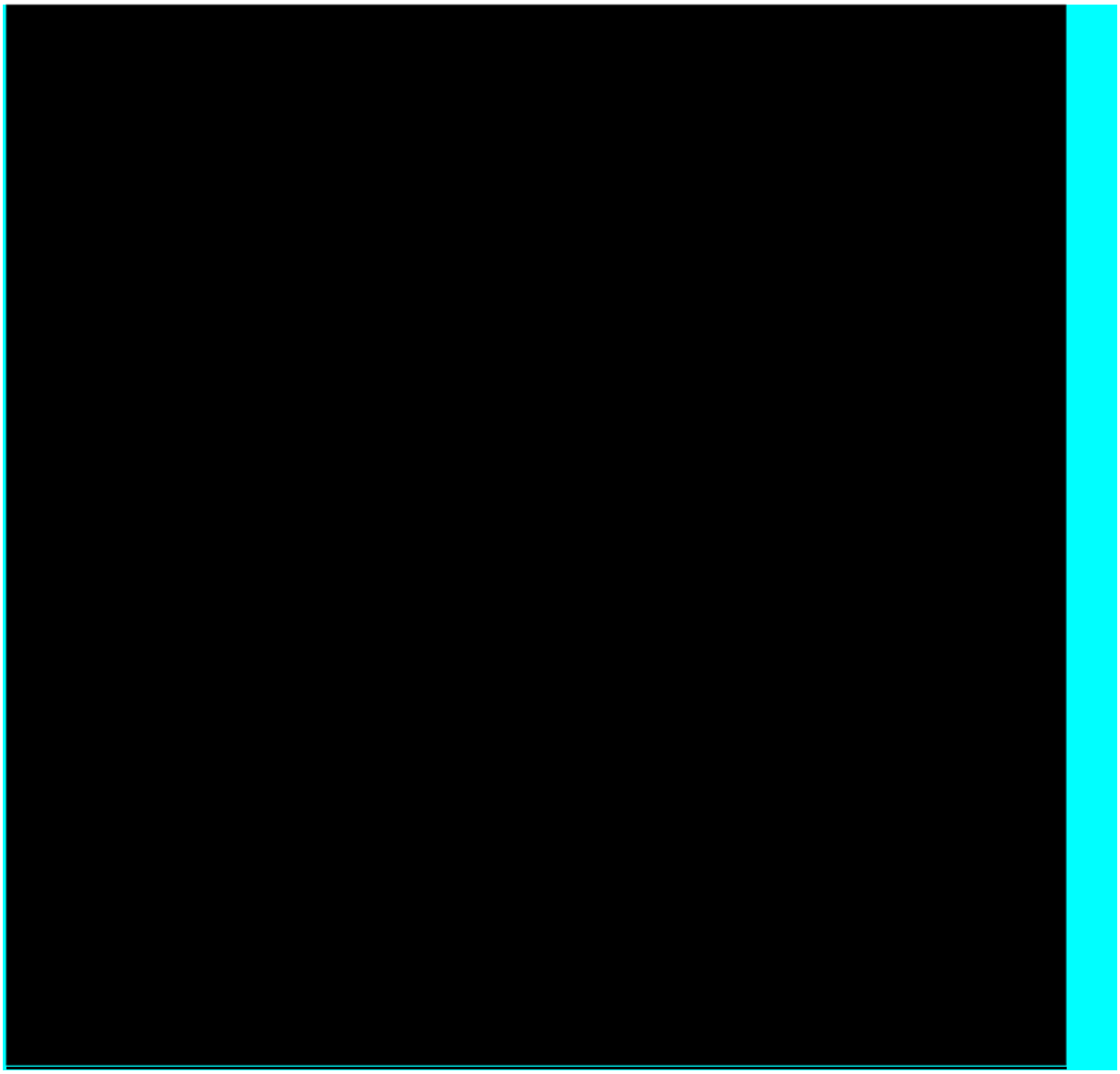
R^2 versus obinutuzumab + bendamustine

For R^2 in comparison with O-Benda, PFS data for the R-refractory population from the induction phase of MAGNIFY was matched to the O-Benda arm in the GADOLIN study (see Section B.2.9). Parametric survival curves were then fit to the matched outcomes.

When comparing the associated KM curves for R^2 and O-Benda, these appear to be diverge from initiation and then converge at 9.8 months (see Figure 36), suggesting the relative treatment effect of R^2 vs. O-Benda is non-constant. This is further supported by the log-cumulative hazard plot which is non-parallel (Figure 37). Therefore, in the base case analysis, PFS for R^2 was modelled using the KM data until the maximum follow-up (11.0 months), beyond which the comparator hazard was applied to extrapolate (see below). This approach ensures the relative treatment

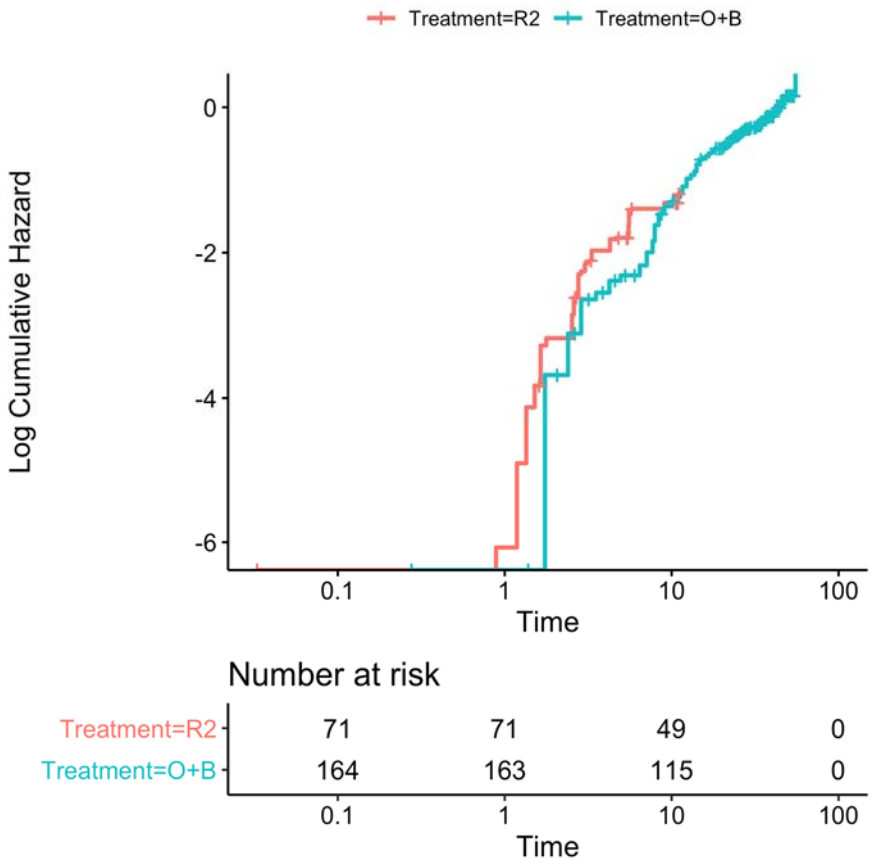
effect of R^2 vs. O-Benda based on the MAIC is accurately reflected. Moreover, the extrapolation method utilises the longer follow-up from the GADOLIN study (4.6 years; additional 3.7 years vs. MAGNIFY) to inform the hazard, which is conservatively applied to both arms over lifetime.

Figure 36: PFS – Matched R^2 versus O-Benda[84]



Key: Benda, bendamustine; O, obinutuzumab; O+B, obinutuzumab plus bendamustine; PFS, progression-free survival; R^2, lenalidomide plus rituximab.

Figure 37: PFS: log-cumulative hazard versus time plot – matched R^2 versus O-Benda



Key: Benda, bendamustine; O, obinutuzumab; O+B, obinutuzumab + bendamustine; PFS, progression-free survival; R^2, lenalidomide plus rituximab.

Parametric curves have also been fitted to each arm for scenario analysis so that the impact of the assumptions used in the base case analysis on cost-effectiveness can be tested. Due to the non-proportional hazards described above, stratified statistical models have been used.

Table 31 presents the AIC and BIC fit statistics for each distribution. For R^2, generalized gamma and exponential are statistically the best fitting, while for O-Benda, this is the log-normal. For R^2, the PFS KM is stepped at approximately 6 months and 9 months, causing the curves to slightly over-estimate the observed data between those times (see Figure 38). Additionally, after the study data, the curves

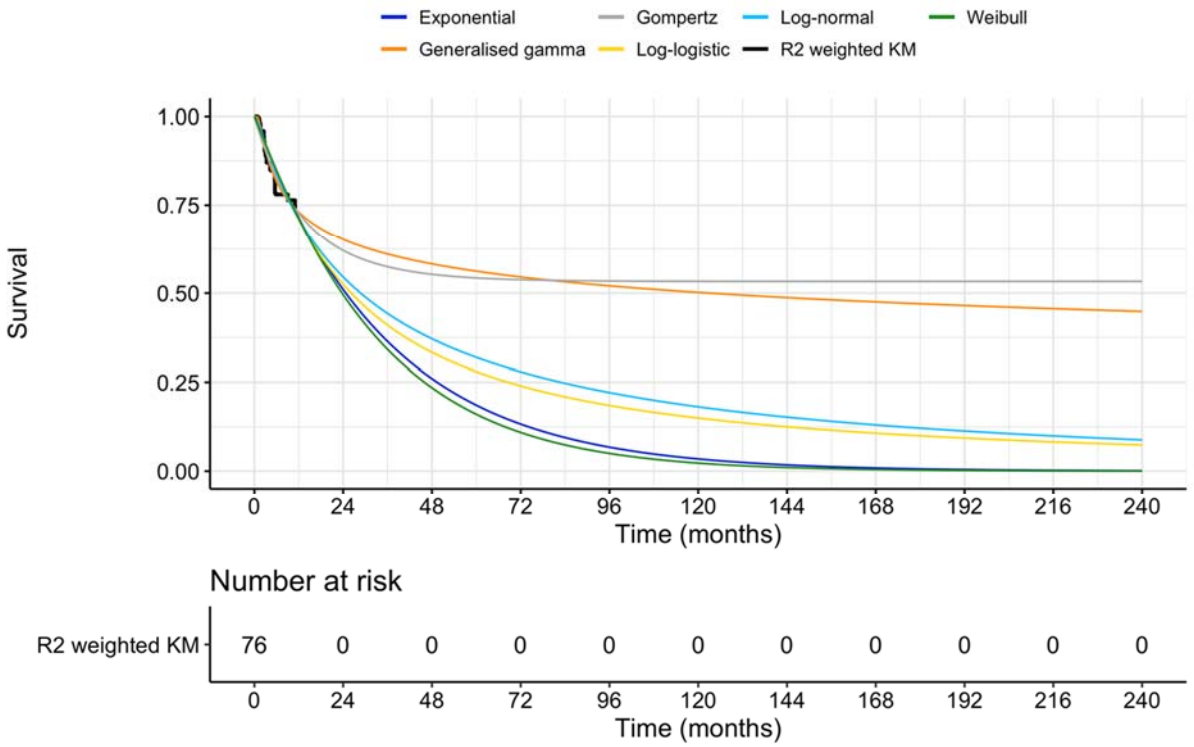
fan out, suggesting different long-term estimations. The curves for O-Benda slightly over-estimate the observed data between 10 months and 2.5 years, and then slightly under-estimate the observed data (Figure 39). Log-normal is statistically the best fitting curve for O-Benda and fits the data well; therefore, this has been selected for the base case.

Table 31: PFS: AIC and BIC – R^2 versus O-Benda

Distribution	R^2		O-Benda	
	AIC	BIC	AIC	BIC
Exponential	163.71	**166.28**	871.57	874.67
Generalized gamma	**160.45**	168.17	866.09	875.38
Gompertz	165.10	170.25	873.52	879.72
Log-logistic	164.94	170.09	868.24	874.44
Log-normal	163.07	168.22	**864.58**	**870.78**
Weibull	165.67	170.82	872.88	879.08

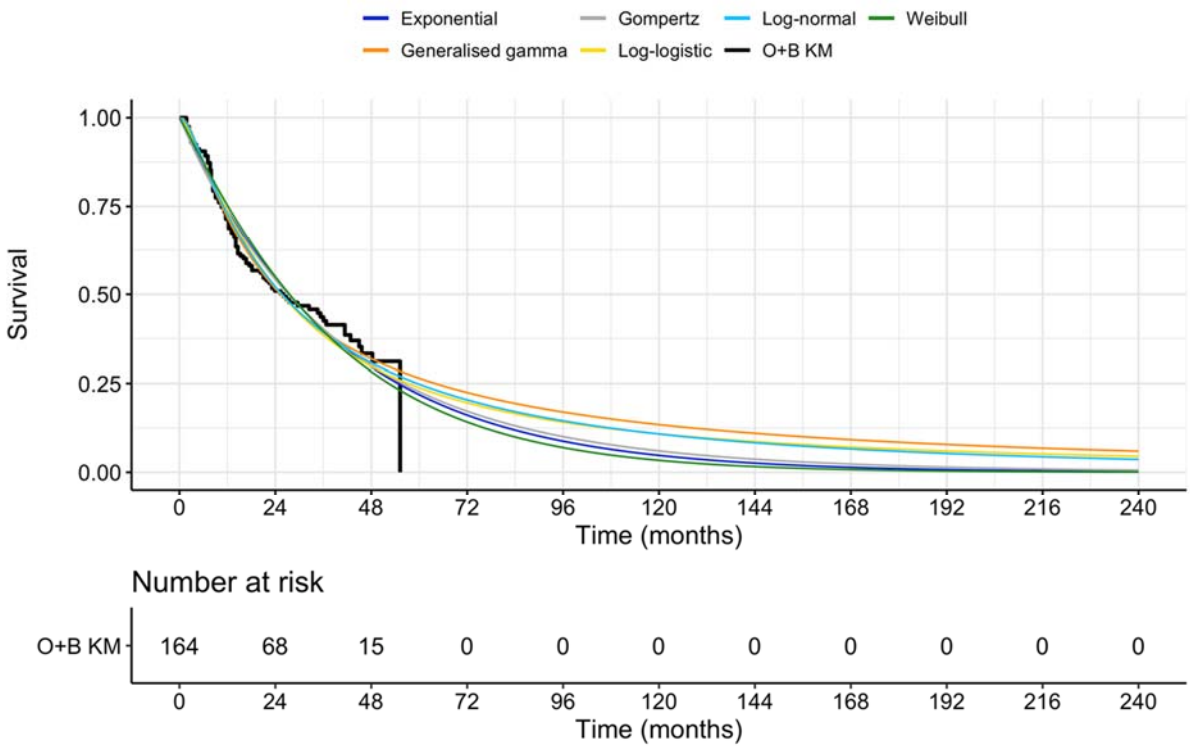
Key: AIC, Akaike information criterion; Benda, bendamustine; BIC, Bayesian information criterion; O, obinutuzumab; PFS, progression-free survival; R^2, lenalidomide plus rituximab.
Bold = statistically best fit; **Purple** = selected for base case.
Note: For the base case, R^2 is assumed to follow KM data and then to match the comparator effectiveness.

Figure 38: PFS parametric curves for R^2 in rituximab refractory population



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; PFS, progression-free survival; R^2, lenalidomide plus rituximab.
Note: For the base case, R^2 is assumed to follow the KM and then use the hazard of O-Benda.

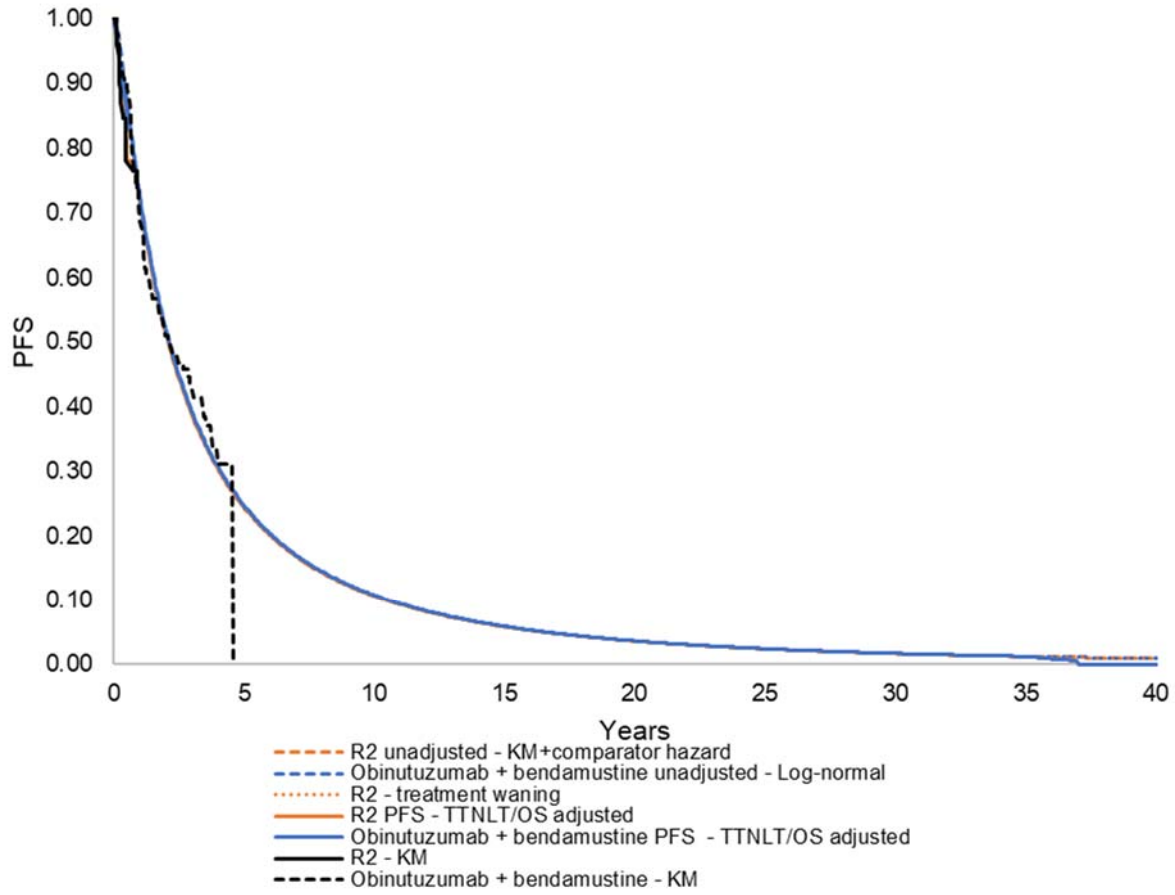
Figure 39: PFS parametric curves for O-Benda in rituximab refractory population



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; O+B, obinutuzumab + bendamustine; PFS, progression-free survival; R^2, lenalidomide plus rituximab.
Notes: KM for O-Benda was digitized from Cheson et al. (2018), which includes the drop at approximately 4.5 years[135]

The curves are adjusted to ensure that the long-term PFS estimates are never predicted to be higher than TTNLT or OS. The final PFS curves used for the model base case are presented in Figure 40, demonstrating the curves at each stage of adjustment.

Figure 40: PFS: Final curves – R^2 versus O-Benda



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; OS, overall survival; PFS, progression-free survival; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

Time to next anti-lymphoma treatment

R^2 versus R-CHOP/R-CVP

As for OS and PFS, TTNLT data for R^2 from AUGMENT were matched to the R-CHOP/R-CVP cohort from HMRN (see Section B.2.9).

Figure 41 presents the log-cumulative hazard versus time plot between R^2 and R-CHOP/R-CVP from AUGMENT and HMRN. The lines of the log cumulative hazards meet in some places, but it could be argued that the proportional hazards assumption is not unreasonable. As it is not definitive from the log-cumulative hazard plot, and to be consistent with the OS and PFS data, stratified statistical models have been used. Unstratified models have been used in scenario analysis.

Figure 41: TTNLT: log-cumulative hazard versus time plot – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

Table 32 presents the AIC and BIC fit statistics for each distribution. For R^2, exponential is statistically the best fitting, whereas for R-CHOP/R-CVP log-normal is the best statistically fitting. Based on visual inspection, all curves appear to fit the KM data reasonably well for R^2; however, some of the curves for R-CHOP/R-CVP over-estimate the observed data at approximately 2.5 years (see Figure 42 and Figure 43). If exponential is selected, this causes the PFS and TTNLT curves to cross at approximately 7 years in both the R^2 and R-CHOP/R-CVP pooled arms, based on the PFS curve selected for the base case analyses. Therefore, the exponential was considered inappropriate.

Given that log-normal is the best statistically fitting extrapolation this has been selected as the base case for R-CHOP/R-CVP. Furthermore, as all curves fit the

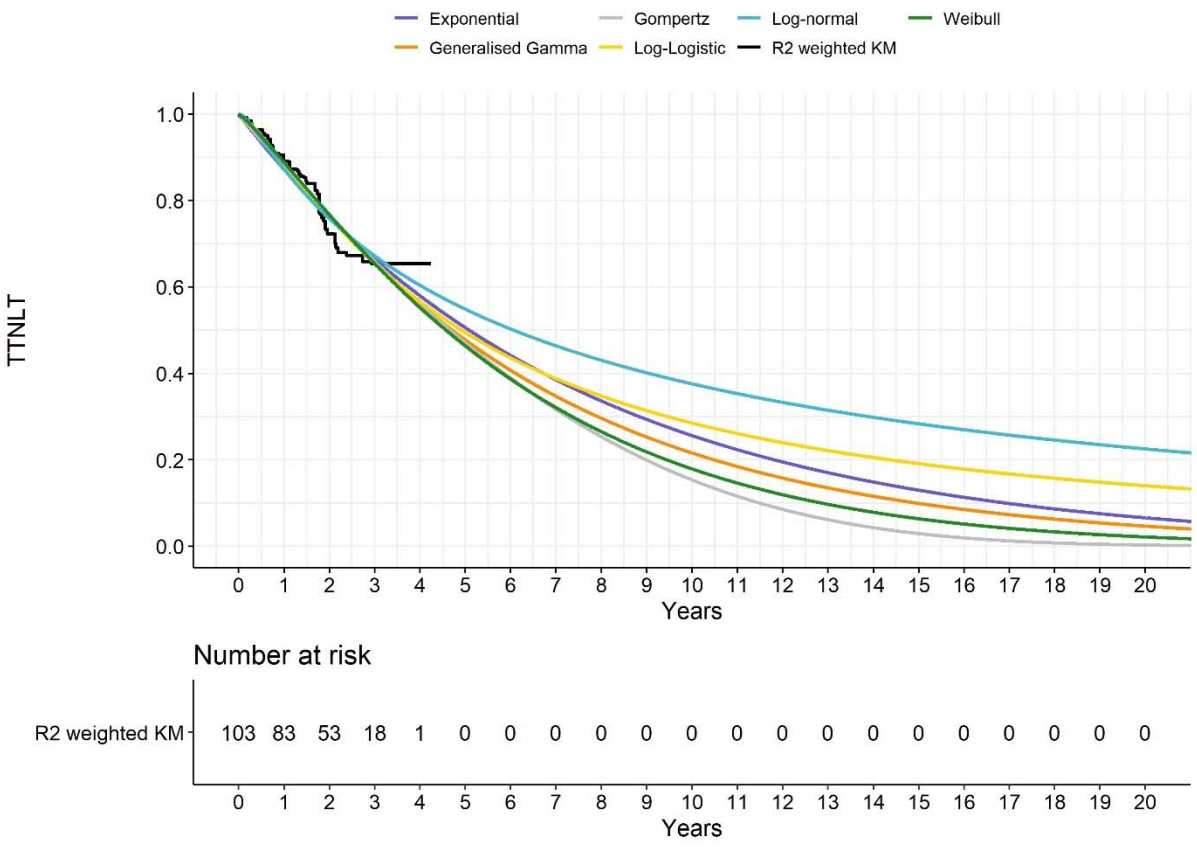
data reasonably well, and to be consistent with the R-CHOP/R-CVP, the log-normal distribution has been chosen for the base case for R^2.

Table 32: TTNLT: AIC and BIC – R^2 versus R-CHOP/R-CVP

Distribution	R^2		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	**509.98**	**513.16**	652.37	654.51
Generalized gamma	513.14	522.69	648.65	655.08
Gompertz	511.84	518.20	647.18	651.47
Log-logistic	510.80	517.17	647.77	652.06
Log-normal	**513.38**	**519.74**	**646.97**	**651.26**
Weibull	511.23	517.59	651.70	655.99

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.
Bold = statistically best fit; **Purple** = selected for base case.

Figure 42: TTNLT parametric curves for R^2 in the non-rituximab refractory population



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

Figure 43: TTNLT parametric curves for R- CHOP/R-CVP in the non-rituximab refractory population

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

The curves are adjusted to ensure that the long-term TTNLT estimates are never predicted to be higher than OS. The final TTNLT curves used for the model base case are presented in Figure 44, demonstrating the curves at each stage of adjustment.

Figure 44: TTNLT: Final curves – R^2 versus R-CHOP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.

R^2 versus obinutuzumab + bendamustine

Seeing as only the induction phase of MAGNIFY could be used for the analysis, TTNLT was not available (see Section B.2.4). Given that the matched analysis for OS and PFS between R^2 and O-Benda appear to overlap, it is assumed that the TTNLT would also be similar. Therefore, the TTNLT treatment from GADOLIN is used for both R^2 and O-Benda.[48]

TTNLT reported in the TA472 submission for O-Benda was digitized with pseudo patient-level data created using the guyot algorithm.[86] Parametric survival curves were then fit to the data to extrapolate beyond the study period. Table 33 presents the AIC and BIC fit statistics from the parametric distributions. Although Gompertz is shown to be the best fitting statistically, this distribution has an implausible plateau at 5 years, suggesting that no patients receive a next treatment after this time (see Figure 45). Log-normal is the second-best fitting according to AIC/BIC and visually

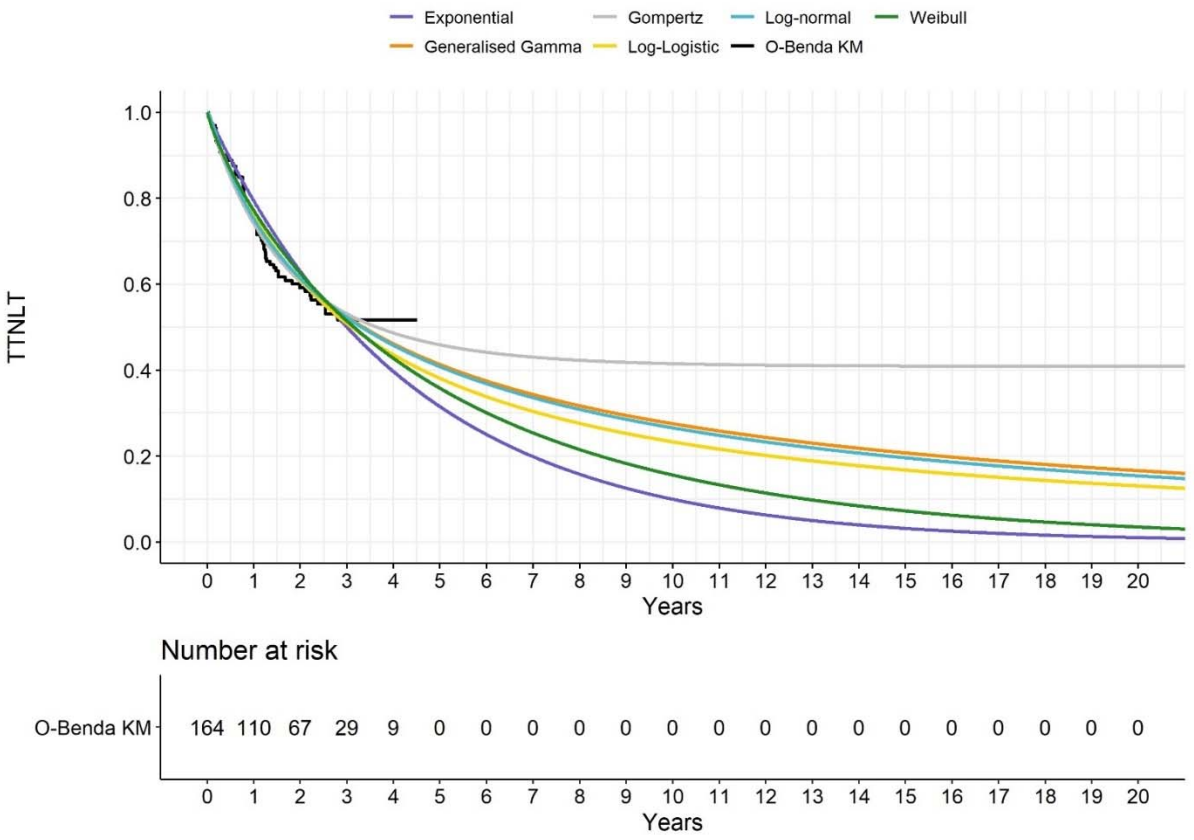
fits the data well. In addition, this curve does not violate any of the time point crossing rules described previously. Therefore, this curve was selected as the base case.

Table 33: TTNLT: AIC and BIC – GADOLIN (O-Benda)

Distribution	AIC	BIC
Exponential	675.31	678.41
Generalized gamma	671.88	681.18
Gompertz	**668.54**	**674.74**
Log-logistic	671.01	677.21
Log-normal	**669.92**	**676.12**
Weibull	675.04	681.24

Key: AIC, Akaike information criterion; Benda, bendamustine; BIC, Bayesian information criterion; O, obinutuzumab; PFS, progression-free survival; TTNLT, time to next antilymphoma treatment.
Bold = statistically best fit; **Purple** = selected for base case.

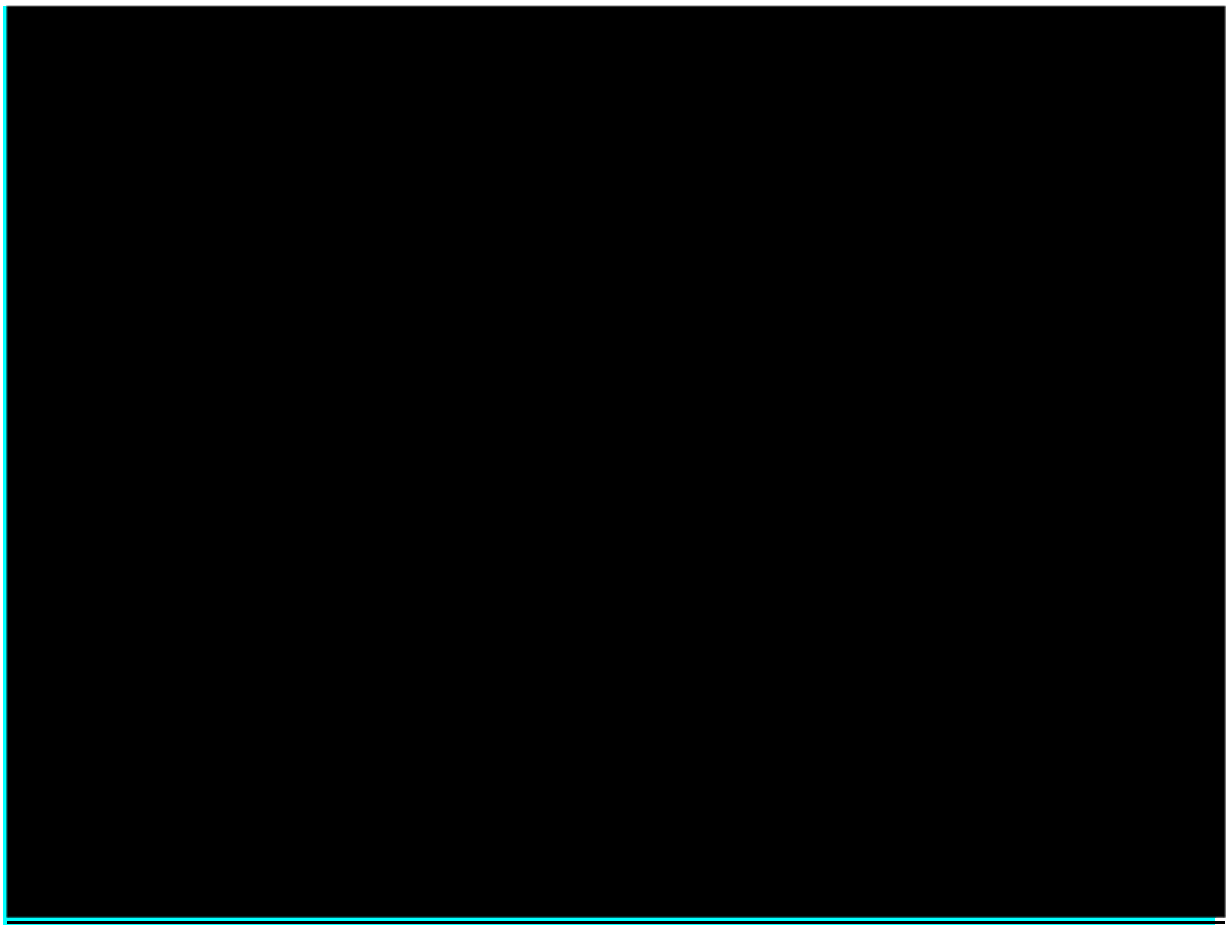
Figure 45: TTNLT parametric extrapolation for O-Benda in the rituximab refractory population



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; TTNLT, time to next antilymphoma treatment.

The curves are adjusted to ensure that the long-term TTNLT estimates are never predicted to be higher than OS. The final TTNLT curves used for the model base case are presented in Figure 46, demonstrating the curves at each stage of adjustment. As the O-Benda data is assumed to apply to both the R^2 and O-Benda arms, only the extrapolated curves for O-Benda are visible.

Figure 46: TTNLT: Final curves – R^2 versus O-Benda



Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; OS, overall survival; R^2, lenalidomide plus rituximab; TTNLT, time to next antilymphoma treatment.
Note: R^2 and O-Benda are assumed to have the same efficacy so only one curve can be seen.

Time on treatment

Data for ToT were used to establish the proportion of patients on treatment per cycle to calculate the overall drug costs. Where possible, ToT KM data were used and extrapolated using parametric distributions. As all induction treatments and maintenance treatments have a maximum duration, all extrapolations were capped at this time.

R^2 versus R-CHOP/R-CVP

ToT data for R^2 versus R-CHOP/R-CVP were taken from the same efficacy source as PFS and OS, i.e. matched AUGMENT population to the R-CHOP/R-CVP cohort from HMRN. The proportion of patients who were still on treatment over time was extracted and fitted with parametric survival curves. Due to the shape of the ToT KMs, the parametric curves produced poor fits to the data, largely over-estimating or under-estimating the actual proportion of patients on treatment (see Figure 47 and Figure 48). Consequently, and since induction and maintenance treatment is for a fixed period, the KM data were used directly in the model to inform the proportion of patients on treatment.

In clinical practice, R-CHOP and R-CVP induction is given for a maximum of eight 21-day cycles (168 days).[130] Rituximab maintenance is then given to patients who respond 3 months after the induction dose (91 days) for a maximum of 2 years (731 days).[129] Therefore, induction drug costs are applied to patients on treatment for up to 168 days, then maintenance costs are applied to patients on treatment between 259 days and 990 days.

Table 34 presents the AIC and BIC fit statistics for the parametric distributions for completeness.

Table 34: ToT: AIC and BIC – R^2 versus R-CHOP/R-CVP

Distribution	R^2		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	1376.26	1379.43	**882.61**	**884.76**
Generalized gamma	1183.35	1192.86	884.96	891.39
Gompertz	**1162.58**	**1168.92**	883.73	888.02
Log-logistic	1299.24	1305.58	894.76	899.05
Log-normal	1364.50	1370.84	901.89	906.18
Weibull	1241.75	1248.10	884.61	888.90

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: KM data used directly for the base case.
Bold = statistically best fit

Figure 47: ToT parametric curves for R^2 in the non-rituximab refractory population

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.

Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration.

Figure 48: ToT parametric curves for R- CHOP/R-CVP in the non-rituximab refractory population

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration. R-CHOP and R-CVP ToT is capped at 990 days.

R^2 versus obinutuzumab + bendamustine

In order to calculate the ToT for R^2 in comparison to O-Benda, the R-refractory population from MAGNIFY (censored at the end of the induction phase), was matched to the O-Benda patient population from GADOLIN. The ToT data was extracted from the matched data-set.

Due to the shape of the ToT KMs, the parametric curves produced poor fits to the data, largely under-estimating the actual proportion of patients on treatment (see Figure 49). Consequently, and due to the fact that induction treatment is for a fixed

period, the KM data were used directly in the model to inform the proportion of patients on treatment.

Table 35 presents the AIC and BIC fit statistics for the parametric distributions for completeness.

Table 35: ToT: AIC and BIC – R^2 versus O-Benda – R^2

Distribution	AIC	BIC
Exponential	835.34	837.91
Generalized gamma	**805.08**	**812.80**
Gompertz	810.70	815.85
Log-logistic	844.76	849.90
Log-normal	854.17	859.32
Weibull	828.17	833.32

Key: AIC, Akaike information criterion; Benda, bendamustine; BIC, Bayesian information criterion; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Bold = statistically best fit.

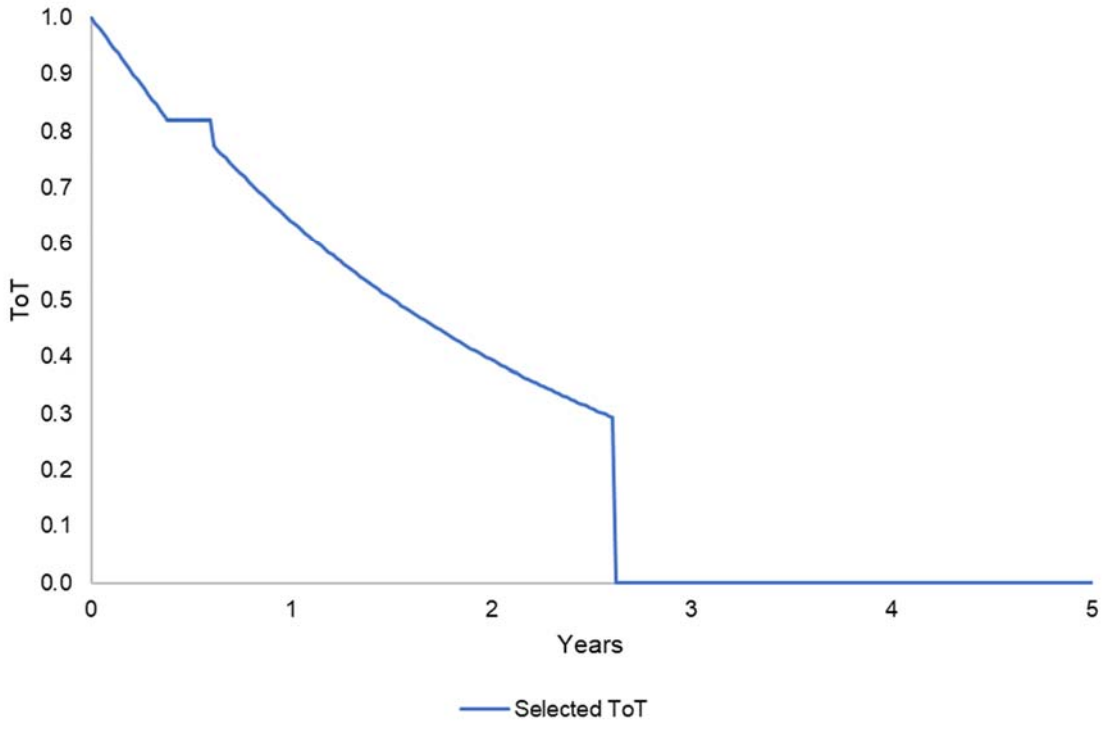
Figure 49: ToT parametric curves for R^2 in the rituximab refractory population

Key: Benda, bendamustine; KM, Kaplan–Meier; O, obinutuzumab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration. Induction phase in MAGNIFY is capped at 336 days; therefore, any patients still on treatment after 336 days were censored.

ToT curves for O-Benda are not available, therefore, information reported in the literature has been used to construct ToT curves. In GADOLIN, it is reported that 78.3% and 81.9% of patients received all scheduled doses of bendamustine and obinutuzumab, respectively, for the O-Benda combination. In addition, 77.5% of patients in the O-Benda arm go on to receive O-maintenance.[135] Thus, to derive a ToT curve from this information, the induction phase is assumed to linearly decrease over time from starting treatment (100%) to the end of induction at 168 days (81.9%, maximum of the O-Benda combination). The curve then jumps to 77.5% at the start of the maintenance treatment and is assumed to exponentially decrease using the median duration of O-maintenance reported in GADOLIN (521 days) for up to 2 years from starting maintenance where treatment costs are capped.

Figure 50 presents the total ToT curve for induction and maintenance used to cost for O-Benda + O-maintenance in the model.

Figure 50: ToT projection for O-Benda in the rituximab refractory population



Key: Benda, bendamustine; O, obinutuzumab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration.

Safety

AEs of treatments were included to account for the additional costs incurred due to treatment toxicities. Grade 3/4 AEs with incidence of greater than 2% in either treatment arm was considered. Two percent was selected as the cut-off as this is in-line with previous submissions.[48] This cut-off also ensured that all the important AEs were costed. The data used to inform the AEs for R^2 were dependent on the model population; AEs collected in AUGMENT were used for the non-R-refractory population, and AEs from the R-refractory population in MAGNIFY were used to inform the R-refractory population. AEs for O-Benda were taken from the GADOLIN study induction phases.[64, 135]

Due to the lack of safety data from HMRN, and limited reporting of AEs within the literature for relapsed/refractory FL, AEs for RCHOP & RCVP were derived from the RELEVANCE study.[5] RELEVANCE is a Phase III study comparing R^2 with R-chemotherapy (R-CHOP, R-CVP and R-Benda) for patients with previously untreated FL. RELEVANCE was used in the base case and incidence was adjusted relative to the incidence of R^2 in AUGMENT compared with R^2 in RELEVANCE:

Comparator AE incidence = ($AE_{COMPARATOR}$ incidence in RELEVANCE/AE_{R2} incidence in RELEVANCE) x AE_{R2} incidence in AUGMENT.

This meant that the difference in AE incidence between the treatments from a head-to-head study in a similar indication were used to derive the differences if such treatments were in the AUGMENT study. If any reported AEs for the comparators were greater than 2% incidence, they were also included for R^2. Any AEs reported in AUGMENT that were used in the model, but were not reported for the comparator, were assumed 0% incidence for the comparator and not costed for. This is conservative, meaning that more AEs are costed within the R^2 arm. An alternative source of AEs for the R-chemotherapies can be selected for scenario analysis using AEs from the relapsed/refractory FL setting. These are based on reported AEs from van Oers (for R-CHOP).[63] No relapsed/refractory studies were found for R-CVP; therefore, for this scenario R-CVP AE incidence is assumed to be the same as R-

CHOP. The incidence of the AEs used in the base case is summarized in Table 36 and Table 37.

Table 36: Grade 3/4 AE incidence: non-rituximab refractory population

AE	R^2 (n=176)		R-CHOP (%)	R-CVP (%)
	n	%		
Neutropenia	88	50.0%	90.3%	85.3%
Leukopenia	12	6.8%	30.8%	16.6%
Anaemia	8	4.5%	4.5%	4.5%
Pneumonia	6	3.4%	NR	NR
Lymphocyte count decreased	6	3.4%	NR	NR
Lymphopenia	5	2.8%	NR	NR
Febrile neutropenia	5	2.8%	9.3%	5.0%
White blood cell count decreased	5	2.8%	NR	NR
Diarrhoea	5	2.8%	1.6%	5.5%
Thrombocytopenia	4	2.3%	2.0%	0.0%
Hypokalaemia	4	2.3%	NR	NR
Pulmonary embolism	4	2.3%	NR	NR
Infusion related reaction	4	2.3%	0.5%	0.0%
Nausea and emesis	0	0.0%	1.4%	3.8%
Allergic reaction	1	0.6%	NR	NR
Hypotension	1	0.6%	NR	NR
Fatigue	2	1.1%	3.2%	0.0%
Alopecia	NR	NR	0.8%	0.0%
Abdominal pain	2	1.1%	1.2%	0.0%
Acute kidney injury	2	1.1%	NR	NR
Source	AUGMENT		RELEVANCE[5] (adjusted)[a]	

Key: AE, adverse event; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; NR, not reported; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a] Comparator AE incidence = ($AE_{COMPARATOR}$ incidence in RELEVANCE/AE_{R2} incidence in RELEVANCE) x AE_{R2} incidence in AUGMENT. R-other used the incidence from all three R-chemotherapies in RELEVANCE.

Table 37: Grade 3/4 AE incidence: rituximab refractory population

AE	R^2 (n=128)		O-Benda (n=204)	
	n	%	n	%
Neutropenia	54	42.2%	56	27.5%
Leukopenia	9	7.0%	0	0.0%

AE	R^2 (n=128)		O-Benda (n=204)	
	n	%	n	%
Anaemia	4	3.1%	11	5.4%
Pneumonia	4	3.1%	3	1.5%
Lymphocyte count decreased	4	3.1%	NR	NR
Lymphopenia	4	3.1%	0	0.0%
Febrile neutropenia	3	2.3%	8	3.9%
White blood cell count decreased	5	3.9%	NR	NR
Diarrhoea	3	2.3%	2	1.0%
Thrombocytopenia	7	5.5%	21	10.3%
Hypokalaemia	3	2.3%	2	1.0%
Infusion related reaction	NR	NR	18	8.8%
Nausea and emesis	NR	NR	2	1.0%
Hypotension	3	2.3%	1	0.5%
Fatigue	7	5.5%	3	1.5%
Sepsis	2	1.6%	2	1.0%
Abdominal pain	3	2.3%	NR	NR
Acute kidney injury	3	2.3%	NR	NR
Source	MAGNIFY (induction phase)		GADOLIN[135] (induction phase)	

Key: AE, adverse event; Benda, bendamustine; O, obinutuzumab; NR, not reported; R^2, lenalidomide plus rituximab.

AE incidence for maintenance treatment and autologous stem cell transplant (ASCT) are also considered. The main AEs considered for maintenance are neutropenia and infection. The incidence of AEs for rituximab maintenance have been taken from van Oers 2010 and obinutuzumab maintenance from GADOLIN.[64, 136] Grade 3/4 neutropenia and infection were the only AEs reported in van Oers for rituximab maintenance, and Grade 3/4 neutropenia and infections were the only AEs from GADOLIN that had incidence above 2% in the O-maintenance period.

AEs that were considered in the NG52 NHL guidelines were used within the model. Febrile neutropenia was identified by the guideline committee as the most likely to result in significant costs.[47] Leger et al. (2006)[137] reported that 98.3% of patients undergoing ASCT were treated for febrile neutropenia; this rate was included within the model.

Table 38: Post-induction AE incidence

AE	R-maintenance (n=167)		O-maintenance (n=143)		Post ASCT (n=60)	
	n	%	n	%	n	%
Neutropenia (Grade 3/4)	19	11.5%	17	10.8%	NA	NA
Infection (Grade 3/4)	16	9.7%	5	3.2%	NA	NA
Febrile neutropenia	NA	NA	NA	NA	59	98.3%
Source	Van Oers 2010[136]		GADOLIN[135]		NG52 guidelines[47] (Leger et al., 2006)[137, 138]	
Key: AE, adverse event; ASCT, autologous stem cell transplant; O, obinutuzumab; NA, not applicable; R, rituximab.						

B.3.4. Measurement and valuation of health effects

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

Data from the health state utility questionnaire EQ-5D-3L were collected in the AUGMENT study at screening, after every three cycles during treatment (Day 1 Cycle 4; Day 1 Cycle 7; Day 1 Cycle 10), and at the end of treatment, regardless of the cause of discontinuation. Following discontinuation, the assessments were completed every 6 months until progressed disease (PD), and 6 months after PD/relapse.

The EQ-5D is a standardized measure of health status developed by the EuroQol Group to provide a simple, generic measure of health for clinical and economic appraisal. The EQ-5D-3L descriptive system comprises the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels: no problems (1), some problems (2) and extreme problems (3).[139, 140]

To estimate utility score within the observed EQ-5D data, the EQ-5D responses were converted to an EQ-5D index score using the UK time trade-off (TTO) method.[141] Each combination of dimensions and levels can be converted to an EQ-5D index score. Patients who answered '1' to all five dimensions have a 'perfect' utility score of 1. For dimensions that a patient answered '2' or '3' (i.e. has some problems or

extreme problems), a utility decrement is applied to that dimension, as shown in Equation 1 below.

Equation 1: Calculation of EQ-5D index score (UK tariff)

$$\text{EQ} - \text{5D index} = 1 - 0.069\,\text{MO2} - 0.314\,\text{MO3} - 0.104\,\text{SC2} - 0.214\,\text{SC3} - 0.036\,\text{UA2} -$$

$$0.094\,\text{UA3} - 0.123\,\text{PD2} - 0.386\,\text{PD3} - 0.071\,\text{AD2} - 0.236\,\text{AD3} - 0.081\,\text{N2} - 0.269\,\text{N3}$$

Key: AD, anxiety/depression; MO, mobility; N2, one or more questions reported as a 2 or 3; N3, one or more questions answered with a 3; PD, pain/discomfort; SC, self-care; UA, usual activity.
Note: The number following the codes indicates a level 2 or level 3 response.

To derive utility scores for the economic model, analysis was conducted in three stages:

- **Exploratory analysis**: utility summaries were performed to determine the effect on utility of each covariate individually and to identify covariates to be considered for regression analysis. In addition to the effect of randomized treatment and health state (progressed off or on treatment), the following prognostic factors were explored using descriptive summaries:
  - Previous rituximab exposure
  - Follicular Lymphoma International Prognostic Index (FLIPI) group
  - Age group
  - Refractory to last therapy
  - Number of prior therapies

These prognostic factors were suggested by the clinicians at the UK advisory board as the most important factors to consider for patients with FL or MZL[133] and were collected in the AUGMENT study.

- **Mixed effects modelling**: mixed effects models were derived to estimate utilities adjusted for covariates, and for repeated measures within subjects
  - A random effect for patients was also included in the models to adjust for the correlation between multiple observations from the same patient. The model selection process explored simple models (including only health state and randomized treatment group) through to the most complex models, including all

covariates present in the model. This process used backwards stepwise selection to select the most appropriate model

- **Selection of final utility mixed effects model**: Appropriate mixed effects utility models were selected as inputs for the economic model.

The full methods and results of the exploratory analysis and mixed effects modelling are shown in Appendix Q.

Table 39 presents the summary of the number of patients in the analysis by treatment, as well as the number of observations. Exploratory analysis was carried out including all patients with an observation for utility. Observations with unknown accompanying progression status were removed for the fitting of mixed effects models. On average, patients receiving R^2 had more observations per person than patients receiving R-mono.

Table 39: Number of patients and EQ-5D observations by treatment

Treatment arm	Patients in ITT population	Patients with baseline utility	Patients with post-baseline utility	Post-baseline observations	Mean number of post-baseline observations
Pooled	358	350	345	2146	6.22
R^2	178	172	169	1117	6.609
R-mono	180	178	176	1029	5.847
Key: ITT, intention to treat; mono, monotherapy; R, rituximab.					

Table 40 presents a summary of baseline utility by treatment. The table indicates that baseline utility is similar between R^2 and R-mono (0.825 versus 0.848, respectively).

Table 40: Baseline utility by treatment – observed utility

Treatment arm	n [patients]	Mean (SD)	Median (range)
Pooled	350	0.837 (0.197)	0.848 (-0.126, 1)
R^2	172	0.825 (0.216)	0.848 (-0.126, 1)
R-mono	178	0.848 (0.178)	0.85 (0.159, 1)
Key: mono, monotherapy; R, rituximab; R^2, lenalidomide plus rituximab; SD, standard deviation.			

Table 41 presents the final mixed effects model after the covariate selection process. The variables remaining after this process were:

- Health state (progression-free vs progressed))
- Next anti-lymphoma treatment
- Treatment
- Baseline utility
- Previous rituximab exposure
- Refractory to last prior regimen
- Number of prior therapies

The health state coefficient value indicates that pre-progression has a utility increment of 0.026 compared with the post-progression. The regression model shows that patients receiving their next anti-lymphoma treatment incur a decrement of 0.03 compared to those who have not, which is significant at the 5% level. The R^2 arm has a utility increment of 0.011 compared with the R-mono arm; this is the smallest coefficient and is not statistically significant in the model (p=0.423). No previous exposure to rituximab has a utility increment of 0.028 compared with previous exposure. Refractory to last therapy has a utility decrement of 0.036 compared with not refractory to last therapy. One prior therapy has a utility increment of 0.034 compared with >1 prior therapy.

Table 41: Parsimonious mixed effects model

Parameters	Estimate	SE	p-value
Intercept	0.326	0.035	<0.001
Health state: Pre-progression	0.026	0.010	0.006
Health state: Progressed on treatment	-0.030	0.013	0.023
Randomized treatment arm: R^2	0.011	0.014	0.423
Baseline utility	0.557	0.037	<0.001
Previous rituximab exposure: no	0.028	0.019	0.142
Refractory to last therapy: yes	-0.036	0.020	0.080
Number of prior therapies: 1	0.034	0.014	0.017
Key: R^2, lenalidomide plus rituximab; SE, standard error.			

Table 42 summarizes the final utility estimates using the regression equation. As the difference between treatments from the model is minimal and not statistically significant, the same utility values based on R^2 are used in the model.

Table 42: Estimated least square means estimates from the final regression model

Health state	R^2	R-mono
Pre-progression	0.847	0.840
Post-progression off treatment	0.821	0.813
Post-progression on treatment	0.791	0.784

Key: mono, monotherapy; R, rituximab.
Note: Utility estimates derived from treatment specific patient characteristics in AUGMENT.

Mapping

EQ-5D values were collected directly from AUGMENT; therefore, no mapping was required.

Health-related quality of life studies

An SLR of the published literature was conducted to identify all relevant utility studies for the treatment of adult patients with R/R FL or MZL. A total of 38 studies were included from 53 publications, including 12 HTAs and 1 observational study. All studies consisted of FL patients, either treatment-naïve or previously treated. In the studies assessing the treatment-naïve population, utility data were reported for the sub-group of patients who had disease progression during treatment. The process of study identification, search strategies and a description of the included utility studies is presented in Appendix H.

Many of the studies found within the literature review refer to the same set of utility values derived from a single study. Only the abstract is available, so the study itself was not included within the SLR due to not reporting any utilities; however, information on this study has been gathered from many economic evaluations for FL patients. Wild et al. (2006) is a study that included 222 patients aged 18 years and over with histologically confirmed FL and an ECOG performance status of 0–2.[127] Details of this study are presented in Appendix H.

Appendix H details the studies found in the SLR to be of most use to the analysis. Table 43 summarizes the mean utility values from the most relevant studies for the economic model in comparison with the mean utility values from AUGMENT. As the reported patient demographics between the studies are similar to AUGMENT; utilities should be comparable.

The mean utility value for the PF state is generally consistent between the studies, with the exception of Pereira et al., which is lower in comparison. Additionally, sample size is small, and as patient characteristics are not reported, it is unclear if this has an impact on the reported value. The mean utility for the PP state is higher in AUGMENT compared with the other studies. The aggregated utility values reported in GADOLIN and Wild et al. are within a similar range; however, the stand-alone progressed utility from Wild et al. and relapsed utility from Pereira et al. are lower. At an advisory board, mean utility values from Wild et al. and GADOLIN were presented compared with AUGMENT to get a general understanding of which values were more appropriate. Clinicians felt that Wild et al. could be limited due to the age of the study and that GADOLIN would provide a more robust comparison with AUGMENT utilities. Clinicians agreed that, although the PP utility seems high from GADOLIN, it would be inaccurate to assume that death is imminent in an indolent lymphoma population, and this level of quality of life could therefore be considered appropriate for many patients.[133]

The difference between PF and PP off-treatment utility from AUGMENT (0.026) is smaller than the difference from Wild et al. (0.069) and GADOLIN (0.064)); however, the difference seen in Pereira et al. (0.27) is much bigger. Given the smaller sample size, inconsistencies and lack of reporting on methods and patients, Pereira et al. was not considered for use in the economic model. The final utility values used in the analysis or as scenarios are discussed below.

Table 43: HRQL studies mean utilities compared to AUGMENT utilities

Category	AUGMENT[a]	Wild et al., 2006[127]	Pereira et al., 2010[142]	GADOLIN[48]
Study demographics[b]				
Sample size	358	222[c]	21	413
Female	52.0%	59.2%	NR	46.4%
Mean age	61.89	60.4	NR	62.0
ECOG, 0-1 2	98.9% 1.1%	95.9% 0.9%	NR	94.7% 4.6%
Stage I-II III-IV	27.1% 72.9%	30.3% 60.6%	NR	14.3% 79.7%
Health state utilities				
PF	0.847	0.805 (0.018)	0.72 (0.250)	0.822 (0.010 – on treatment) 0.807 (0.012 – off treatment)
PP	0.821 (off treatment) 0.791 (on treatment)	0.736 (aggregated) 0.62 (0.06 – relapsed disease)	0.45 (0.431)	0.758 (0.024)

Key: NR, not reported; PF, progression-free; PP, post-progression.
Notes: [a]R^2 utilities used for reference
[b]Some values may not add to 100% due to missing data
[c]Patient demographics reported in Pettengall et al., 2007[24], demographics available for 218 patients.

Adverse reactions

The impact of Grade 3 and 4 AEs has been explored in the cost-effectiveness analysis. Utility decrements for each of the AEs included in the analysis (described in Section B.3.3) were sourced from a targeted review of the literature or used in previous appraisals. AE utility decrements are applied in the model for the expected duration of each AE, the data for which were sourced from the literature. When an expected duration estimate could not be sourced, mean duration was assumed to be the maximum of the available duration estimates. A similar assumption was made in ID1379.[123] Table 44 summarizes the AE utility decrements, durations and sources used in the cost-effectiveness analysis.

Table 44: AE utility decrements and durations

Adverse event	Disutility	Duration (days)	Source for disutility	Source for duration
Neutropenia	0.090	15.09	Nafees et al. (2008)[143]	TA306[144]
Leukopenia	0.119	13.96	TA513[52] (assumed to be the same as anaemia)	TA306[144]
Anaemia	0.119	16.07	Swinburn et al. (2010)[145]	TA306[144]
Pneumonia	0.200	14.00	Beusterien et al. (2010)[146]	TA306[144]
Lymphocyte count decreased	0.100	34.00	Stein et al. (2018)[147]	Assumed maximum of all Grade 3/4 AEs
Lymphopenia	0.100	34.00	Stein et al. (2018)[147]	TA306[144]
Febrile neutropenia	0.150	7.14	Lloyd et al. (2006)[148]	TA306[144]
White blood cell count decreased	0.100	34.00	Stein et al. (2018)[147]	Assumed maximum of all Grade 3/4 AEs
Diarrhoea	0.048	34.00	Nafees et al. (2008)[143]	Assumed maximum of all Grade 3/4 AEs
Thrombocytopenia	0.108	23.23	Tolley et al. (2013)[149]	TA306[144]
Hypokalaemia	0.124	34.00	TA423[150]	Assumed maximum of all Grade 3/4 AEs
Pulmonary embolism	0.124	34.00	TA423[150]	Assumed maximum of all Grade 3/4 AEs
Infusion-related reaction	0.195	34.00	Tolley et al. (2013)[149]	Assumed maximum of all Grade 3/4 AEs
Nausea and emesis	0.048	6.00	Nafees et al. (2008)[143]	TA306[144]
Allergic reaction	0.098	34.00	Hannouf et al. (2012)[151]	Assumed maximum of all Grade 3/4 AEs
Hypotension	0.057	8.00	Hannouf et al. (2012)[151]	TA306[144]
Fatigue	0.073	31.50	Nafees et al. (2008)[143]	TA306[144]
Alopecia	0.045	34.00	Nafees et al. (2008)[143]	Assumed maximum of all Grade 3/4 AEs
Infection	0.195	34.00	Tolley et al. (2013)[149]	Assumed maximum of all Grade 3/4 AEs

Adverse event	Disutility	Duration (days)	Source for disutility	Source for duration
Sepsis	0.267	34.00	Hannouf et al. (2012)[151]	Assumed maximum of all grade ¾ AEs
Abdominal pain	0.069	17.00	Doyle et al. (2008)[152]	TA306[144]
Acute kidney injury	0.270	29.75	TA306[144]	TA306[144]
Key: AEs, adverse events.				

The disutility of each AE is multiplied by the duration and probability of each AE per treatment arm, resulting in a one-off QALY decrement per treatment. Table 45 summarizes the final utility decrement per treatment used in the model.

Table 45: Overall QALY decrement per treatment

Treatment	One-off QALY loss due to AEs
Non-R-refractory	
R^2	0.0052
R-CHOP	0.0058
R-CVP	0.0046
R-refractory	
R^2	0.0054
O-Benda	0.0043
Post-induction	
R-maintenance	0.0022
O-maintenance	0.0010
ASCT	0.0029
Key: AE, adverse event; ASCT, autologous stem-cell transplant; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; O, obinutuzumab; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab.	

In the non-R-refractory setting, it is projected that R^2 has a higher utility loss than R-CVP. This is because some AEs that were reported in AUGMENT and considered for R^2 were not reported in RELEVANCE and therefore assumed 0% incidence for R-CVP. This biases against R^2: if the AEs that were not reported for RELEVANCE were not considered for R^2, the disutility for R^2 would reduce to 0.0033, which is less than the disutility projected for R-CVP. As noted in Section B.2.10. an RCT examining R^2 versus R-chemo in untreated FL patients reported that there were

seemingly fewer AEs for R^2 compared with R-chemo.[99] Furthermore, R^2 is not associated with the standard toxicities associated with chemotherapy treatments, and so the associated utility loss for R^2 may be expected to be lower than that of the R-chemo regimens, as was seen for R^2 versus R-CHOP.

For the R-refractory population, the majority of the incidence reported from MAGNIFY for R^2 was higher than the incidence reported from O-Benda in GADOLIN. Therefore, the disutility associated with R^2 was greater than estimated for O-Benda. As there are no data to compare the AEs within the same study, it is uncertain how these AEs would compare in practice. Nonetheless, clinical opinion considers R^2 to be less aggressive than O-Benda and so naively comparing data from MAGNIFY to GADOLIN could be considered a conservative approach.

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

For the base case, utilities derived from the AUGMENT EQ-5D regression model have been used to directly inform the health states in the model for all treatments, with literature utilities from Wild et al. (2006) tested in scenario analysis. The values derived from the AUGMENT study are based directly on a relevant patient population and measure the health states as per the economic model using EQ-5D, which is NICE's preferred measure for utility values.

The disease characteristics used to derive the utility values from the regression model change depending on the population; therefore, each population has a slightly different utility value. As the treatment covariate between R^2 and R-mono in AUGMENT was minimal and non-significant (see Table 41), the same utilities have been applied to each treatment arm within the pairwise comparisons. The different disease characteristics used to inform the regression model and the overall utility value per health state are summarized in Table 46.

Table 46: Disease characteristics used to inform the utilities and utility values

Covariate	R^2 vs R-CHOP/R-CVP	R^2 vs O-Benda
Baseline utility[a]		0.837
% rituximab naïve	31.7%	0.0%
% refractory to last prior regimen[b]	15.6%	93.9%
% with 1 prior regimen	88.9%	51.2%

Source	HMRN[124]	GADOLIN (O-Benda arm FL population)[135]
Health state	**R^2 vs R-CHOP/R-CVP**	**R^2 vs O-Benda**
Progression-free	0.863	0.814
Progressed (off treatment)	0.837	0.787
Progressed (on treatment)	0.808	0.758

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab.
Note: [a]Baseline utility was derived from AUGMENT pooled population. [b]Refractory to last prior regimen was not reported from HMRN so values used for R-CHOP and R-CVP were assumed to be the same % as the pooled population in AUGMENT.

Age-related utility decrements have also been included in the model base case to account for the natural decline in quality of life associated with age. This was done by estimating the utility values of the general population at each age and creating a utility multiplier based upon the algorithm by Ara and Brazier (2010).[153] This multiplier is applied in each cycle throughout the model time horizon. The algorithm used to estimate the multiplier is shown below:

*General population utility value = 0.9508566 + 0.0212126*male – 0.0002587*age – 0.0000332*age^2*

The general population baseline age is estimated from the same starting age from each of the pairwise comparisons (see Table 25). These result in baseline general population utilities of 0.803 for R-CHOP/R-CVP analysis and 0.815 for the O-Benda analysis. The utilities derived from the AUGMENT regression model (Table 46) exceed the age-matched general population utility values for the PF and PP off-treatment health states. This could be considered implausible, however it is important to note that the utilities algorithm derived from the general population will include multiple comorbidities, whereas study candidates could have fewer comorbidities. Also, clinical opinion suggests that it is not unreasonable to assume that patients who are progression-free have similar utilities to the general population, and the impact on quality of life for patients with disease progression but who do not start treatment immediately would be minimal in the period between determination of progression and initiation of any subsequent therapy. However, to account for the uncertainties on these utility values, multiple scenarios using literature values or

utility adjustments have been conducted. The following scenarios have been considered and are presented in Section B.3.8:

- Utility values collected from the Wild et al. study[127]
  - Progression-free utility of 0.805 (SE: 0.018)
  - Progressed (off-treatment) utility of 0.7363 (assuming combined health states of active disease – newly diagnosed/relapsed, see Appendix H)
  - Progressed (on-treatment) utility value of 0.62 (assuming single health state active disease – relapsed, see Appendix H)
- Utility values capped at general population. This scenario adjusts any utility values that are above the baseline general population utility to equal the general population utility.

Table 47 summarizes the utility values used in the base case analysis.

Table 47: Summary of utility values for cost-effectiveness analysis

State	Utility value	Reference in submission	Justification
Progression-free	Vs R-CHOP/CVP: 0.863 Vs O-Benda: 0.814	Section B.3.4, page 177 and 182	EQ-5D values derived from a relevant patient population and model specific health states.
Post-progression (off treatment)	Vs R-CHOP/CVP: 0.837 Vs O-Benda: 0.787		
Post-progression (on-treatment)	Vs R-CHOP/CVP: 0.808 Vs O-Benda: 0.758		

State	Utility value	Reference in submission	Justification
Neutropenia	-0.0037	Section B.3.4, page 180	Identified through targeted published literature search or assumed equivalent to published estimate for a similar AE.
Leukopenia	-0.0045		
Anaemia	-0.0052		
Pneumonia	-0.0077		
Lymphocyte count decreased	-0.0093		
Lymphopenia	-0.0093		
Febrile neutropenia	-0.0029		
White blood cell count decreased	-0.0093		
Diarrhoea	-0.0045		
Thrombocytopenia	-0.0069		
Hypokalaemia	-0.0115		
Pulmonary embolism	-0.0115		
Infusion related reaction	-0.0182		
Upper respiratory tract infection	-0.0008		
Nausea and emesis	-0.0091		
Allergic reaction	-0.0012		
Hypotension	-0.0063		
Fatigue	-0.0042		
Alopecia	-0.0182		
Infection	-0.0237		
Sepsis	-0.0032		
Abdominal pain	-0.022		
Acute kidney injury	-0.0037		

Key: AE, adverse event; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab.

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

In line with the NICE reference case, the perspective on costs in all cost-effectiveness analyses is that of the NHS and PSS in England. An SLR for healthcare resource use and cost data relevant to this submission is reported in Appendix I.

Intervention and comparators' costs and resource use

Table 48 summarizes the costs and associated healthcare resource use of each treatment in the analysis.

Table 48: Unit costs associated with the technology in the economic model

Items	R^2	R-CHOP	R-CVP	O-Benda	Reference in submission
Induction treatment cost per cycle	£8,848 cycle 1; £5,338 cycles 2-5; £4,169 cycles 6-12. [redacted] cycle 1 with PAS; [redacted] cycles 2-5 with PAS; [redacted] cycles 6-12 with PAS	£1,216 cycles 1-8	£1,200 cycles 1-8	£9,995 cycle 1; £3,371 cycles 2-6	Table 54
Induction administration cost per cycle	£1,348 cycle 1; £335 cycles 2-5	£400 cycles 1-8	£400 cycles 1-8	£1,413 cycle 1; £738 cycles 2-6	Table 55
Maintenance treatment cost per cycle	NA	£1,345 (SC) £1,170 (IV)	£1,345 (SC) £1,170 (IV)	£3,312	Table 54
Maintenance administration per cycle	NA	£273 (SC) £338 (IV)	£273 (SC) £338 (IV)	£400	Table 55

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; IV, intravenous; NA, not applicable ; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab; SC, subcutaneous.

Drug acquisition costs

The unit drug costs for each treatment and its source are summarized in Table 49.

Table 49: Unit costs of each treatment

Treatment	Size	Cost per pack	Source
Lenalidomide	21 x 2.5 mg tablets	£3,426.00 ([redacted] with PAS)	MIMS (Revlimid)[154]
	21 x 5 mg tablets	£3,570.00 ([redacted] with PAS)	
	21 x 10 mg tablets	£3,780.00 ([redacted] with PAS)	
	21 x 15 mg tablets	£3,969.00 ([redacted] with PAS)	
	21 x 20 mg tablets	£4,168.50 ([redacted] with PAS)	
Rituximab	2 x 100 mg vials	£349.25	MIMS (MabThera)[155]
	1 x 500 mg vial	£873.15	
	1 x 1,400 mg SC inj	£1,344.65	
	2 x 100 mg vials	£314.33	MIMS (Rixathon®)[156]
	1 x 500 mg vial	£785.84	
Cyclophosphamide	1 x 1,000 mg vial	£13.47	eMIT[157]
	1 x 2,000 mg vial	£27.50	
	1 x 500 mg vial	£8.31	
Doxorubicin	1 x 10 mg vial	£4.48	
	1 x 200 mg vial	£15.59	
	1 x 50 mg vial	£17.78	
Vincristine	5 x 1 mg vials	£11.59	
	5 x 2 mg vials	£17.82	
	5 x 5 mg vials	£99.00	
Prednisolone	28 x 1 mg tablets	£0.17	
	28 x 2.5 mg tablets	£0.59	
	30 x 20 mg tablets	£4.17	
	56 x 25 mg tablets	£20.25	
	28 x 5 mg tablets	£0.27	
Bendamustine	5 x 100 mg vials	£75.13	
	5 x 25 mg vials	£26.32	
Obinutuzumab	1 x 1,000 mg vial	£3,312.00	MIMS (Gazyvaro®)[158]

Key: eMIT, electronic market information tool; MIMS, Monthly Index of Medical Specialities; PAS, patient access scheme; SC, subcutaneous; inj, injection

The dosing schedule for each treatment is outlined in Section B.3.2, Table 27. Dose reductions can occur for lenalidomide; therefore, for the base case, dosing data have been taken directly from AUGMENT or MAGNIFY to align the drug costs with the efficacy data. Specifically, the observed number of patients on each dosage of lenalidomide at every cycle was combined with unit drug costs to give a weighted cost per cycle. This is multiplied by the proportion of patients eligible for treatment (have not yet failed treatment) who receive treatment in that cycle. In contrast to using mean relative dose intensities (RDIs), this method accurately captures the impact of treatment reductions or missed treatment cycles over time on costs. This method was used in another lenalidomide submission to accurately capture drug costs.[159] The number of patients eligible for treatment for each cycle is calculated based on the ToT data (ToT KM curves, presented in Section B.3.3) and the mean treatment cycle length also derived directly from the clinical study data. For each patient, either the maximum of the number of cycles expected to have been completed given their time on treatment or the maximum number of treatment cycles recorded in the data was determined. Using the maximum number of cycles each patient was expected to have received, the total number of patients eligible for treatment for each cycle was calculated. Furthermore, to align with the costing method applied for lenalidomide, rituximab costs for the R^2 treatment arm were also multiplied by the proportion of patients eligible for treatment who receive treatment in each cycle. These calculations are presented in Appendix R for both AUGMENT and MAGNIFY. Scenario analyses were conducted that using the mean RDIs. A value of [redacted] was used in the scenario as mean RDI for rituximab in the R^2 arm of AUGMENT; for lenalidomide, this was [redacted][66] The values [redacted] and [redacted] were used for rituximab and lenalidomide, respectively, for the scenario in the R-refractory population from MAGNIFY.

To similarly align the costing of the comparators to the study dosing methods described above for R^2, it was assumed that the proportion of patients eligible for treatment who receive treatment in each cycle as per the rituximab monotherapy arm of AUGMENT would apply to all comparator arms of the model. Alternatively, a mean dose intensity value of 87.5% was assumed in line with the range sourced by the ERG for a previous submission[123], and this was applied across all individual chemotherapies within the R-chemotherapy comparator regimens. As dose

reductions are not recommended for rituximab[160], no dose intensity value was applied to rituximab within R-chemotherapy combinations or R-maintenance. This is also expected to reflect the efficacy data of these therapies, sourced from real-world evidence and therefore reflective of clinical practice and SmPC recommendations. The dosing schedule for obinutuzumab was also not adjusted for dose intensity as dose reductions are neither recommended in the SmPC nor were allowed within the GADOLIN study which informs the O-Benda efficacy.[64, 132] Dose intensity for bendamustine within the O-Benda regimen, on the other hand, was included based on the ratio between average planned dose and actual cumulative dose reported from the GADOLIN study and TA472.[48, 64] This resulted in a mean dose intensity of 92.6% for the bendamustine element of O-Benda.

For treatments dependent on patients' BSA, IPD from AUGMENT are used, with the method of moments technique to calculate the average number of vials that would be required to satisfy one administration of treatment.[161] The method of moments first derives a log-normal distribution for the patient BSA within the study based upon the mean and standard deviation of the BSA measured at baseline. It then uses the log-normal distribution to predict what proportion of patients requires each number of vials to administer the required dose. This method assumes that patients only receive whole vials (no vial sharing), and thus accounts for drug wastage. Other methods are included within scenario analysis; dose banding is explored based on the national dose banding schedule[162], as well as assuming no wastage by using the minimum cost per mg for each treatment. The average numbers of vials estimated for each dose per treatment from the method of moments and dose banding technique are presented in Table 50. The average number of vials is then used to calculate the average cost per cycle for each treatment.

Table 50: Average vials per treatment

Regimen	Treatment	Size	Average vials from method of moments	Average vials from dose banding
R^2	Rituximab (375 mg/m^2)	100 mg vial	2.35	2.05
		500 mg vial	1.02	1.01
R-chemo	Rituximab (375 mg/m^2)	100 mg vial	2.35	2.05
		500 mg vial	1.02	1.01
R-chemo	Cyclophosphamide (750 mg/m^2)	500 mg vial	0.74	0.70
		1,000 mg vial	1.25	1.28
		2,000 mg vial	0.00	0.00
R-chemo	Doxorubicin (50 mg/m^2)	10 mg vial	0.00	0.00
		50 mg vial	0.00	0.00
		200 mg vial	1.00	1.00
R-chemo	Vincristine (1.4 mg/m^2)	1 mg vial	0.00	0.00
		2 mg vial	1.00	1.00
		5 mg vial	0.00	0.00
O-Benda	Bendamustine (90 mg/m^2)	25 mg vial	0.51	0.40
		100 mg vial	1.77	1.80

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; IV, intravenous; NA, not applicable ; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab; SC, subcutaneous.

For oral therapies other than lenalidomide, the least waste and most efficient pack size of tablets was used. For prednisolone, the 20 mg pack was considered the most efficient based on cost per tablet and dose schedule, respectively. As lenalidomide was costed with the assumption that the cost of a full pack of lenalidomide tablets is incurred per cycle, the cost of any potential wastage due to missed doses/unused tablets is fully accounted for in the model.

According to the SmPC for lenalidomide, all patients should receive tumour lysis syndrome prophylaxis (allopurinol, rasburicase or equivalent as per institutional guidelines) during the first week of the first cycle or for a longer period if clinically indicated.[4] According to clinical opinion, allopurinol would be used in practice for tumour lysis syndrome prophylaxis, and so a weekly cost of administration of allopurinol was included in the first treatment cycle for patients receiving lenalidomide. Data used to calculate the cost of allopurinol are available in Table 51.

Table 51: Tumour lysis syndrome prophylaxis regimen drug costs

Treatment	Size	Cost per pack	Dose per administration	Administrations per week	Cost per week	Source
Allopurinol	28 x 100 mg tablets	£1.72	100 mg	7	£0.43	MIMS (Allopurinol)[163]; SmPC (Allopurinol)[164]
Key: MIMS, Monthly Index of Medical Specialities; SmPC, summary of product characteristics.						

Granulocyte colony stimulating factor (G-CSF) is taken to reduce the duration of neutropenia and incidence of febrile neutropenia while taking immunomodulatory drugs. Data on the usage of G-CSF for lenalidomide was taken from the AUGMENT CSR. The percentage of patients receiving G-CSF during treatment was converted into an instantaneous rate based on the total treatment period of 12 four-weekly cycles of lenalidomide. A cost per 4-weekly cycle was then derived by multiplying the cost per dose of G-CSF by the percentage of patients receiving G-CSF per 4-weekly cycle. This method was based on that used in another lenalidomide submission to accurately capture G-CSF costs.[159] The data used are given in Table 52.

Table 52: G-CSF regimen drug costs per cycle

Treatment	Size	Pack cost	Dosing (u/kg)	Dose applied	Cost per dose	Source
Filgrastim	5 x 30,000,000iu	£263.52	500,000	37,542,416	£65.95	MIMS (Filgrastim)[165]; SmPC (Filgrastim)[166]
Treatment	**Patients receiving G-CSF (%)**	**Instantaneous rate**	**Patients receiving G-CSF per 4-weekly cycle (%)**		**Cost per cycle**	**Source**
R^2	35.8%	0.92%	3.63%		£2.39	AUGMENT CSR[66]
Key: CSR, clinical study report; G-CSF, granulocyte colony stimulating factor; iu, international units.						

Rituximab maintenance is given to some patients who respond to rituximab chemotherapy induction treatment at a dose of 375 mg/m^2 by intravenous therapy (IV) or 1,400 mg by subcutaneous formulation every 3 months after the last dose of

induction therapy for a maximum of 2 years or until disease progression.[129, 160] Clinical opinion suggests that greater than 90% of R-maintenance is given subcutaneously; therefore, for the model base case it is assumed 100% subcutaneous usage with 90% tested in scenario analysis.

Obinutuzumab maintenance is given to patients who achieve a partial or complete response to obinutuzumab induction treatment, or have stable disease at a dose of 1,000 mg once every 2 months for a maximum of 2 years or disease progression.[132] The cost per treatment cycle of each of these treatments has been calculated the same way as for the induction treatments, i.e. using the ToT curves (see Section B.3.3).

Given that the treatment pathway in AUGMENT does not include rituximab maintenance and that clinicians suggested that rituximab maintenance would unlikely to be offered post R^2, maintenance is not considered for patients who received R^2 in the model.[133]

Table 53: Proportion of patients who receive maintenance

Induction treatment	% patients receiving maintenance treatment	Source
R-CHOP/R-CVP	[redacted]	HMRN [124]
O-Benda	77.5%	GADOLIN[64]
Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; HMRN, Haematological Malignancy Research Network; O, obinutuzumab; R, rituximab.		

The cost per treatment cycle of each component of treatment regimen is presented in Table 54.

Table 54: Cost per treatment cycle

Regimen	Treatment	Cost per cycle
R^2	Lenalidomide	£4,168.50 [redacted] (with PAS)
	Rituximab	£4,679.93 cycle 1; £1,169.98 cycles 2-5
R-CHOP	Rituximab	£1,169.98
	Cyclophosphamide	£23.05
	Doxorubicin	£15.59

Regimen	Treatment	Cost per cycle
	Vincristine	£3.56
	Prednisone	£3.48
R-CVP	Rituximab	£1,169.98
	Cyclophosphamide	£23.05
	Vincristine	£3.56
	Prednisolone	£3.48
O-Benda	Obinutuzumab	£9,936.00 cycle 1; £3,312.00 cycles 2-6
	Bendamustine	£58.74
R-maintenance (IV)	Rituximab	£1,169.98
R-maintenance (SC)	Rituximab	£1,344.65
O-maintenance	Obinutuzumab	£3,312.00

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; IV, intravenous; O, obinutuzumab; PAS, patient access scheme; R, rituximab; R^2, lenalidomide plus rituximab; SC, subcutaneous.

As discussed in Section B.3.3, total drug costs use the duration of treatment to estimate the proportion of patients on treatment per model cycle.

Administration costs

Drug administration costs are based on NHS reference costs tariffs, additional pharmacy costs for the preparation of the infusion, and NHS transport costs. For rituximab or obinutuzumab combination chemotherapies, a cost of £374.52 (SB14Z – deliver complex chemotherapy, including prolonged infusional treatment) was applied at first administration of each cycle, and a cost of £312.34 (SB15Z – deliver subsequent elements of chemotherapy cycle) for each subsequent administration per cycle, in line with administrations tariffs applied in the treatment-specific NICE submissions.[48, 50] For simpler chemotherapies, such as rituximab (when in combination with lenalidomide, an oral therapy, in the R^2 arm), a cost of £309.22 (SB13Z – deliver more complex parental chemotherapy at first attendance) is applied for first administration of each cycle and £312.34 (SB15Z) for subsequent administrations. No administration costs are applied to oral therapies.

A cost of £312.34 (SB15Z) per administration was applied for rituximab maintenance (IV), and for obinutuzumab maintenance a cost of £374.52 (SB14Z) was applied – these are in line with the assumptions given in TA243 and TA472, respectively.[48, 50]

In TA513, the Committee commented that obinutuzumab's infusions take longer than rituximab infusions and are likely to have higher administration costs; therefore, applying a lower tariff for rituximab maintenance is consistent with these comments.[52] For rituximab maintenance, a cost of £247.74 (SB12Z - deliver simple parenteral chemotherapy at first attendance) was applied per administration.

The ERG for TA472 considered that all chemotherapies should be administered in a day case setting and that bendamustine would be administered over 2 days; administration assumptions in the model have been made consistent with these comments.[48]

Pharmacy costs are applied for all infusion treatments assuming a 15-minute infusion preparation time based on TA243[50, 167], and £48 per hour for hospital-based scientific and professional staff (Band 6) from Personal Social Services Research Unit (PSSRU) costs.[168] In TA243, it was also assumed that 30% of patients require NHS transport at a cost of £39.24.[50] The ERG from TA472 believed that this cost should be included with the administration costs in the base case; therefore, this cost has been inflated to 2018 costs and applied to all administrations in the model.

Table 55 summarizes all the administration costs used within the model and presents which treatments these costs are applied to.

Table 55: Administration costs applied in the model

Administration type	Cost	Source	R-CHOP/CVP O-Benda		R-mono (in R^2)		Maintenance	
			Day 1	Days 2+	Day 1	Days 2+	R-main	O-main
Intravenous R- or O-chemotherapy, first administration	£374.52	NHS Reference Cost (2017/18) SB14Z: Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance (Daycase)[169]	✓					✓
Other intravenous chemotherapy, first administration	£309.22	NHS Reference Cost (2017/18) SB13Z: Deliver more Complex Parenteral Chemotherapy at First Attendance (Daycase)[169]			✓			
Subsequent intravenous chemotherapy administration	£312.34	NHS Reference Cost (2017/18) SB15Z: Deliver Subsequent Elements of a Chemotherapy Cycle (Daycase)[169]		✓		✓	✓ (IV)	
Maintenance, first administration (SC)	£247.74	NHS Reference Cost (2017/18) SB12Z: Deliver Simple Parenteral Chemotherapy at First Attendance (Daycase)[169]					✓ (SC)	
Pharmacy preparation cost	£12.00	TA243[50] PSSRU 2018[168]	✓	✓	✓	✓	✓	✓
NHS transportation	£13.43	TA243[50]	✓	✓	✓	✓	✓	✓

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; IV, intravenous; main, maintenance; mono, monotherapy; NHS, National Health Service; O, obinutuzumab; PSSRU, Personal Social Services Research Unit; R, rituximab; R^2, lenalidomide plus rituximab; SC, subcutaneous.

Treatment-specific monitoring

In order to monitor the dose-limiting toxicities of neutropenia and thrombocytopenia from lenalidomide, the SmPC suggests that a complete blood cell count, including white blood cell count with differential count, platelet count, haemoglobin and haematocrit, should be performed weekly for the first 3 weeks of Cycle 1, every 2 weeks during Cycles 2–4, and then at the start of each cycle thereafter.[133] In order to account for these treatment-specific tests, the cost of a full blood count is added to each of the specified treatment cycles for lenalidomide for each visit. The cost of a full blood count was taken from Papaioannou et al. (2012)[167] and uplifted to 2018 costs using PSSRU inflation indices[168], resulting in a cost of £6.28 per full blood count.

Health-state unit costs and resource use

Table 56 presents the costs that are included in each of the health states.

Table 56: List of health states and associated costs in the economic model

Health states	Items	Value	Reference in submission
PF (on treatment)	Drug acquisition	**R²** Cycle 1: [redacted] Cycles 2-5: [redacted] Cycles 6-12: [redacted] [redacted] **R-CVP:** £1,200 per cycle **O-Benda** Cycle 1: £9,995 Cycles 2-6: £3,371	Table 48, page 187
	Drug administration	**R²** Cycle 1: [redacted] Cycles 2-5: [redacted] [redacted] [redacted] **O-Benda** Cycle 1: £1,413 Cycles 2-6: £738	Table 48, page 187

Health states	Items	Value	Reference in submission
	Maintenance/ASCT	R-maintenance: £1,345 (SC), £1,170 (IV) O-maintenance: £3,312 ASCT: £35,558	Table 48, page 187 Table 59, page 202
	Disease monitoring	£254.95 per month	Table 57, page 200
	Adverse events	£1,832 (R^2 non R-refractory) £3,604 (R-CHOP) £2,754 (R-CVP) £1,773 (R^2 R-refractory) £1,376 (O-Benda) £370 (R-maintenance) £253 (O-maintenance) £6,336 (ASCT)	Table 61, page 204
PF (off-treatment)	Disease monitoring	£83.09 per month	Table 57, page 200
PP (off-treatment)	Disease monitoring	£58.04 per month	Table 57, page 200
PP (on-treatment)	Disease monitoring	£232.17 per month	Table 57, page 200
	Subsequent treatments	£5,195 (R^2) £8,371 (R-CHOP/CVP/O-Benda)	Table 62, page 206
Death	Terminal care	£6,362	Page 206

Key: ASCT, autologous stem cell transplant; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; IV, intravenous; O, obinutuzumab; PF, progression-free; PP, post-progression; R, rituximab; R^2, lenalidomide plus rituximab; SC, subcutaneous.

Disease monitoring

Disease monitoring resource use costs are assumed to be similar to those presented in previous FL NICE submissions[48, 50, 52] and ESMO guidelines.[14] The disease monitoring resource use is split by health state: progression-free (split further into three periods: induction phase, maintenance phase and follow-up phase) and post-progression.

The three defined periods for the progression-free resource use outcomes are different per treatment depending on the duration of induction therapy and whether it is followed by maintenance treatment. The induction phase is rounded to the nearest

whole month. For example, R-CHOP is given for eight 21-day cycles; therefore, the induction phase is 6 months ([12/365.25] months x 21 days x 8 cycles = 5.52 months). Patients that receive maintenance and are therefore still within the PF (on-treatment) health state after the induction phase incur the maintenance treatment disease monitoring costs.

During the induction phase, it is assumed that patients have monthly haematologist visits and diagnostic tests with a CT scan every 6 months. The frequency of the follow-up visits is based on the ESMO guidelines, which reduce to every 3–4 months following the completion of induction therapy. Therefore, the maintenance phase assumes a frequency of every 3 months, and the post-maintenance follow-up phase assumes a frequency of every 4 months.[14] CT scans are assumed to reduce to once annually in the maintenance phase and no scans in the post-maintenance follow-up phase. In the PD state, patients are assumed to have higher frequency of visits and, therefore, have monthly haematology visits and monthly diagnostic tests.

Resource use information for MZL patients is limited. However, similar tests and frequencies are suggested in the MZL ESMO guidelines.[15] Therefore, it is assumed that these are the same as the FL disease monitoring and no additional costs have been considered.

Costs for each category are based on NHS reference costs or costs reported in TA243 inflated to 2018 costs. Total monthly resource use costs are summarized in Table 57.

Table 57: Disease monitoring resource use frequencies and costs

Item		Progression-free monthly frequency			Progressed monthly frequency	Unit cost	Source for cost
		Induction	Maintenance	Follow-up			
Haematologist led		1	0.33	0.25	1	£164.80	NHS Reference Costs 2017/18: 303 Clinical Haematology consultant led, Non-Admitted Face-to-Face Attendance, Follow-up[169]
Diagnostic tests	FBC	1	0.33	0.25	1	£6.28	TA243[50] inflated to 2018 costs
	patient history/physical exam					£6.21	
	Full profile (U&E, LFT, calcium)					£17.10	
	Serum IgG, IgA, IgM and electrophoresis					£25.10	
	LDH test					£12.69	
CT scans		0.17	0.04	0	0	£136.70	NHS Reference Costs 2017/18: RD27Z Computerised Tomography Scan of more than Three Areas[169]
Total monthly cost		**£254.95**	**£83.09**	**£58.04**	**£232.17**		

Key: CT, computed tomography; FBC, full blood count; Ig, immunoglobin; LDH, lactate dehydrogenase; LFT, liver function tests; NHS, National Health Service; TA, technology appraisal; U&E, urea and electrolytes.

Stem cell transplant

Patients who are considered fit and young enough (approximately less than 70 years) and who relapse early but who are not refractory to induction therapy would be considered for consolidation with ASCT, most likely in a second- or third-line setting.[47, 133] Given that ASCT is part of the clinical pathway for some patients, costs, AEs and disutilities associated with these have also been considered within the model. As consolidation with ASCT is dependent on a number of factors for each patient, the proportion of patients considered for ASCT in the model is reflective of the individual efficacy source. As R-CVP is generally not used as an induction regimen prior to ASCT, the proportion of patents who have ASCT was taken from the R-CVP efficacy from HMRN and not the pooled efficacy, to reflect the costs associated with R-CVP in clinical practice. R-CHOP, on the other hand, may be used as an induction therapy in ASCT candidates, and therefore this proportion is taken from the individual R-CHOP cohort from HMRN. The proportion of patients who receive ASCT after induction therapy (but before progression) is summarized in Table 58. As with maintenance treatment, ASCT was not offered to patients after R^2 within the AUGMENT study protocol, and clinicians suggested it was unlikely that ASCT would be offered post R^2 in clinical practice.[133] Consequently, ASCT is not considered a relevant treatment post R^2 in the model.

Table 58: Proportion of patients who receive SCT

Induction treatment	ASCT (%)	Source
R^2	0.00%	AUGMENT[66]
R-CHOP	[redacted]	HMRN[124]
R-CVP	0.00%	HMRN[124]
O-Benda	0.00%	GADOLIN[64]

Key: ASCT, autologous stem cell transplant; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; HMRN, Haematological Malignancy Research Network; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab.

The cost of ASCT is based on the costs assumed in NHL guidance 2016.[47] The cost of ASCT is assumed to be £34,000 based upon the tariff utilized by the transplanting haematologist on the guidance committee.[47] Alternative costs using NHS reference costs were considered to under-estimate the true costs and were only considered in

scenario analysis. In the model base case, the NHL guidance cost is inflated to 2018 costs with the NHS reference cost (considered to under-estimate the true costs) used in scenario analysis. These costs are summarized in Table 59.

Table 59: Costs associated with SCT

Type of SCT	Cost	Source
ASCT	£35,558.15	NHL guidelines[47] uplifted to 2018 costs[168]
	£18,520.20	NHS reference costs 2017/18: SA26A elective inpatient[169]
Key: ASCT, autologous stem cell transplant; NHL, non-Hodgkin's lymphoma; NHS, National Health Service.		

Adverse reaction unit costs and resource use

As discussed in Section B.3.3, the AEs considered are those Grade 3–4 AEs occurring in ≥2% of patients. The unit costs associated with the management of these AEs were sourced from NHS reference costs 2017/18, NICE guidelines or previous NICE appraisals. Table 60 summarizes the costs associated with each AE.

Table 60: Adverse event costs included in the model

Adverse event	Cost per event	Source
Neutropenia	£1,892.59	NHS reference costs 2017/18: SA08J NEL[169]
Leukopenia	£3,414.95	NHS reference costs 2017/18: SA31E NEL[169]
Anaemia	£2,995.67	NHS reference costs 2017/18: Weighted average of SA03G to SA03H, NEL[169]
Pneumonia	£2,526.61	NHS reference costs 2017/18: Weighted average of DZ11K to DZ11V, NEL[169]
Lymphocyte count decreased	£382.38	NHS reference costs 2017/18: Weighted average of SA08G to SA08J, day case[169]
Lymphopenia	£382.38	NHS reference costs 2017/18: Weighted average of SA08G to SA08J, day case[169]
Febrile neutropenia	£6,511.63	NICE guidelines NG52; Appendix A[47]; inflated to 2018
White blood cell count decreased	£382.38	NHS reference costs 2017/18: Weighted average of SA08G to SA08J, day case[169]
Diarrhoea	£1,507.73	NHS reference costs 2017/18: FD01J NEL[169]
Thrombocytopenia	£2,754.86	NHS reference costs 2017/18: Weighted average of SA12G to SA12K, NEL[169]
Hypokalaemia	£339.40	NHS reference costs 2017/18: Weighted average of KC05H to KC05N, day case[169]

Adverse event	Cost per event	Source
Pulmonary embolism	£1,329.92	NHS reference costs 2017/18: Weighted average of DZ09J to DZ09Q, across NEL, NES and day case[169]
Infusion-related reaction	£618.19	NHS reference costs 2017/18: SA31E NES[169]
Nausea and emesis	£618.19	NHS reference costs 2017/18: SA31E NES[169]
Allergic reaction	£395.24	NHS reference costs 2017/18: Weighted average of WH05Z across NEL, NES and day case[169]
Hypotension	£2,169.10	ERG report, TA306 2013[144]; inflated to 2018
Fatigue	£93.06	ERG report, TA306 2013[144]; inflated to 2018
Alopecia	£0.00	Assume no hospital episodes related to this AE, and no direct costs are incurred as in LRiG estimate rev. TA162, TA175 (TA374)[170]
Infection	£1,570.07	NHS reference costs 2017/18: Weighted average of WH07A to WH07G across NEL, NES and day case[169]
Sepsis	£2,829.68	NHS reference costs 2017/18: Weighted average of WJ06A to WJ06J across NEL[169]
Abdominal pain	£623.23	NHS reference costs 2017/18: Weighted average of FD05A and FD05B across NEL, NES and day case[169]
Acute kidney injury	£2,673.79	NHS reference costs 2017/18: Weighted average of LA07H to LA07P, NEL[169]
Key: ERG, Evidence Review Group; NEL, non-elective inpatient long-stay; NES, non-elective inpatient short-stay; NHS, National Health Service; SCT, stem cell transplant; TA, technology appraisal.		

The unit cost of each AE is applied to the incidence rate for each treatment, which is applied as a one-off upfront cost to each treatment arm in the model. The cost of AEs for maintenance treatment and ASCT are first multiplied by the proportion of patients who receive such treatments in each treatment arm (Table 53 and Table 58). The total cost of AEs per treatment is presented in Table 61.

Table 61: Total AE cost per treatment

Treatment	Total cost
Non-R-refractory	
R^2	£1,831.71
R-CHOP	£3,604.13
R-CVP	£2,753.56
R-refractory	
R^2	£1,773.94
O-Benda	£1,376.18
Post-induction	
R-maintenance	£369.95
O-maintenance	£253.32
ASCT	£6,400.93

Key: AE, adverse event; ASCT, autologous stem-cell transplant; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; O, obinutuzumab; R, rituximab; R^2, lenalidomide plus rituximab.

Miscellaneous unit costs and resource use

Subsequent treatments

Subsequent treatments were included in the model as an average cost per patient which is applied as a one-off cost to those patients entering the PP (on-treatment) health state derived using the TTNLT data. The average subsequent treatment cost was based on the same efficacy source used to derive the clinical survival for each treatment arm where possible (see Section B.3.3). For the non-R-refractory population, subsequent treatments from AUGMENT were used to derive the cost for patients on the R^2 arm. For R-CHOP and R-CVP, data from HMRN using the total subsequent treatment data from the pooled R-chemotherapies has been used in order to draw from a larger sample size. For the R-refractory population, subsequent treatment data from MAGNIFY are too immature and data from the GADOLIN study are not presented. Therefore, subsequent treatment data used for the non-R-refractory population have been used. This is considered a viable assumption given that in clinical practice, treatment options are essentially the same for R-refractory patients but expected outcomes would differ. Scenarios have also been presented, which assumes the same subsequent treatment cost between treatment arms.

Total treatment costs have been calculated from the per cycle cost and mean duration or taken from the literature. The mean duration of treatment was taken from HMRN, with AUGMENT mean durations used in scenario analysis. Subsequent treatment costs and the proportion of patients who received each treatment are presented in Table 62.

Table 62: Subsequent treatments and costs

Subsequent treatment	AUGMENT n (%) n=178	HMRN n (%) n=129	Mean duration (days)		Total treatment cost	Total administration cost	Cost source
			AUGMENT	HMRN (base case)			
R-mono	4 (2.2%)		89.5		£13,767	£3,901	MIMS[156]
R-Benda	6 (3.4%)		145.3		£5,079	£3,196	MIMS and eMIT[156, 157]
R-CHOP	3 (1.7%)		99.0		£5,060	£1,664	See Section B.3.5
R-CVP	0 (0%)		21.0		£6,539	£2,179	See Section B.3.5
Other R-chemo	10 (5.6%)		92.8		£13,366	Assumed to be included in treatment cost	TA137[121] uplifted to 2018 costs[168]
O-Benda	3 (1.7%)		121.8		£20,669	£3,923	See Section B.3.5
Bendamustine	8 (4.5%)		124.0		£266	£2,480	eMIT[157]
Other chemotherapy	7 (3.9%)		45.7		£3,855	Assumed to be included in treatment cost	TA137[121] uplifted to 2018 costs[168]
Targeted therapies	12 (6.7%)		223.2		£8,701	£0	Idelalisib cost used - MIMS[171]
Radiotherapy	10 (5.6%)		NA		£1,932	NA	TA137[121] uplifted to 2018 costs[168]
Other	6 (3.4%)		63.2		£3,855	Assumed to be included in treatment cost	TA137[121] uplifted to 2018 costs[168]
ASCT	2 (1.1%)		NA		£35,558	NA	See Section B.3.5
Total weighted cost (R^2)					**£3,053**	**£401**	
Total weighted cost (comparator)					**£7,712**	**£660**	

Key: ASCT, autologous stem cell transplant; Benda, bendamustine; chemo, chemotherapy; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; eMIT, electronic market information tool; HMRN, Haematological Malignancy Research Network; MIMS, Monthly Index of Medical Specialities; mono, monotherapy; NA, not applicable; NR, not reported; R, rituximab; R^2, lenalidomide plus rituximab; TA, technology appraisal.
Notes: [a]Percentages include multiple lines therefore total may be over 100%.

Terminal care

A one-off, end-of-life cost was applied to patients at the point of dying to reflect the cost of terminal care. The end-of-life cost was calculated based on the average cost derived from the Round et al. (2015) modelling study, which estimated the cost of cancer care during the final phases of life.[172] The study presented the end-of-life cost from health, social or informal care services for breast, colorectal, lung or prostate cancer individually in England and Wales. In the model, the average health and social care costs are used in the base case and uplifted to 2018 costs using indices from PSSRU.[168] This resulted in a cost of £6,361.77 per patient upon death.

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

Summary of base case analysis inputs

A summary of all base case parameters and distributions are provided in Appendix S.

Assumptions

The key assumptions of the economic analysis are described in Table 63.

Table 63: Summary of model assumptions

Topic	Assumption	Justification/reason
Population	Model uses pooled FL and MZL population, and so efficacy and inputs taken from FL populations are assumed applicable to the MZL population.	Similar outcomes for relapsed/refractory FL and MZL are reported in the literature presenting outcomes by histology (see Section B.1.3). Clinical opinion suggests that patients with MZL are treated the same as patients with FL within the relapsed/refractory setting. There is a lack of data reported for MZL patients alone.
Efficacy	Treatment effect is assumed until 5 years.	Five years was chosen based on the previous appraisals in the same patient population. Lack of evidence to suggest an appropriate time; therefore, other time points tested in scenario analysis.
	R-CHOP and R-CVP are assumed to have the same efficacy.	Clinical opinion suggested that similar outcomes would be expected between the two treatments in the relapsed/refractory setting. HMRN data show that endpoints for OS and PFS are similar.

Topic	Assumption	Justification/reason
	R^2 assumed same OS as O-Benda.	Unconfounded induction phase data is only available for 1-year using the MAGNIFY study for R^2. KM curves overlap between the OS MAGNIFY data versus GADOLIN.
	R^2 assumed same TTNLT as O-Benda.	Due to lack of TTNLT data for R^2, and the similarity in OS and PFS data, TTNLT has been assumed to be the same
	After the maximum follow-up for PFS in MAGNIFY, R^2 is assumed to have the same PFS hazard per cycle as O-Benda.	PFS KM curves for R^2 and O-Benda appear to diverge from initiation but converge as time progresses, suggesting the relative treatment effect is non-constant. To best reflect this, the observed R^2 data was utilized to maximum follow-up, beyond which the comparator hazard was applied to extrapolate.
	After the maximum follow-up for PFS in AUGMENT, R^2 is assumed to have the same PFS hazard per cycle as R-CHOP/CVP	PFS KM curves for R^2 and RCHOP/RCVP appear to diverge from initiation but converge as time progresses, suggesting the relative treatment effect is non-constant. To best reflect this, the observed R^2 data was utilized to maximum follow-up, beyond which the comparator hazard was applied to extrapolate.
	O-Benda ToT assumed to have an exponential distribution	Lack of data reported for the duration of O-Benda treatment. Therefore, assumptions were required to calculate drug costs using information available.
	AEs were assumed to be 0% for comparators if not reported in literature.	No other evidence to suggest what incidence would be compared to R^2; therefore, conservative assumption used.
Utilities	Equal health state utilities were assumed between treatment comparisons.	The utility regression model using data from AUGMENT did not show a significant difference in utility between the R^2 and R-mono treatment arms, and data were not available to compare all treatments based upon treatment effects. Literature data are used in scenario analysis.
Dosing	The IPD from AUGMENT used to account for actual versus scheduled treatment dosing were assumed to be the same for the comparators.	No data were available for the comparators; however, dose interruptions may be possible in clinical practice. Therefore, IPD from the R-mono arm of AUGMENT were used to calculate the % eligible but not receiving treatment per cycle to accurately reflect costs.

Topic	Assumption	Justification/reason
	All patients receive subcutaneous injection for R-maintenance.	Clinical opinion suggests a large proportion of patients would have R-maintenance subcutaneously. A lower proportion of patients receiving R-maintenance subcutaneously was tested in scenario analysis.
	Pharmacy preparation costs and NHS transport was assumed to apply to all administrations.	These are in line with ERG preferred assumptions in a previous appraisal in the same disease area (TA472).[48]
Subsequent treatments	The distribution of subsequent treatments for the R-refractory population was assumed to be the same as the non-R-refractory population.	There are no data for the distribution of treatments for R-refractory patients. Clinician opinion suggests that the distribution of treatments for these two populations in practice are similar.

Key: AE, adverse event; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; IPD, individual patient data; KM, Kaplan-Meier; O, obinutuzumab; OS, overall survival; MZL, marginal zone lymphoma; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.

B.3.7. Base case results

Base case incremental cost-effectiveness analysis results

Table 64 present the base case incremental cost-effectiveness results for R^2 at the agreed PAS price. R^2 is shown to be cost-effective versus all comparators at the £30,000 willingness-to-pay (WTP) threshold. In comparison to O-Benda, the incremental costs are lower for R^2 with a very slight QALY loss due to AE disutility causing a large incremental cost-effectiveness ratio (ICER). However, as this ICER sits in the southwest quadrant of the cost-effectiveness plane, this ICER can be interpreted as the ICER of O-Benda versus R^2, and therefore R^2 is considered the more cost-effective treatment.

Table 64: Base case pairwise results (based on PAS price)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R² versus R-CHOP							
R-CHOP	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£11,471
R² versus R-CVP							
R-CVP	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,814
R² versus O-Benda							
O-Benda	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,960,557 (SW)

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; LYG, life years gained; O, obinutuzumab; PAS, patient access scheme; QALY, quality-adjusted life year; R, rituximab; R², lenalidomide plus rituximab; SW, southwest.

Table 65 presents the fully incremental analysis for the non-R-refractory population. R-CHOP is extendedly dominated by R-CVP, given that R-CVP has smaller AE costs and no patients who went onto ASCT.

Table 65: Base case fully incremental analysis – non-rituximab refractory population (with PAS)

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (strict dominance)	Incremental ICER (Extended dominance)
R-CVP	[redacted]	[redacted]				
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Strictly Dominated
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,814	£16,814

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; PAS, patient access scheme; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab.

B.3.8. Sensitivity analyses

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed within the cost-effectiveness model for 1,000 iterations. The mean incremental costs and QALYs from R^2 versus the comparators are displayed in Table 66. The visual results of the PSA runs are displayed in Figure 51–Figure 53. The results of the probabilistic analysis are similar to those of the deterministic analysis.

For the comparison with O-Benda, given that the efficacy is assumed equal, the PSA has less effect on the results. The ICER displayed still appears slightly different to the deterministic ICER; however, this is due to the estimate of an extremely small QALY loss due to AEs, which, if changed slightly, will cause a larger difference in the ICER. Given the sensitivity of the ICER to differential QALYs, the net monetary benefit (NMB) of R^2 and O-Benda for deterministic and probabilistic results are provided in Table 67 as an alternative comparison.

Table 66: Mean results of PSA (1,000 runs) and comparison with deterministic results

Technology	Total costs (£)		Total QALYs		ICER (£) versus baseline (QALYs)	
	PSA	Deterministic	PSA	Deterministic	PSA	Deterministic
R² versus R-CHOP						
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]		
R²	[redacted]	[redacted]	[redacted]	[redacted]	£13,443	£11,471
R² versus R-CVP						
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]		
R²	[redacted]	[redacted]	[redacted]	[redacted]	£20,896	£16,814
R² versus O-Benda						
O-Benda	[redacted]	[redacted]	[redacted]	[redacted]		
R²	[redacted]	[redacted]	[redacted]	[redacted]	£24,486,823 (SW)	£16,960,560 (SW)

Key: Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CI, confidence interval; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; O, obinutuzumab; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R², lenalidomide plus rituximab; SW, southwest.

Table 67: Mean results of PSA (1,000 runs) and comparison with deterministic net monetary benefit results – R² versus O-Benda

Technology	Total costs (£)		Total QALYs		NMB (£)	
	PSA	Deterministic	PSA	Deterministic	PSA	Deterministic
R² versus O-Benda						
O-Benda	[redacted]	[redacted]	[redacted]	[redacted]	£41,721	£41,847
R²	[redacted]	[redacted]	[redacted]	[redacted]	£68,169	£68,206

Key: Benda, bendamustine; CI, confidence interval; NMB, net monetary benefit; O, obinutuzumab; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R², lenalidomide plus rituximab.

Figure 51: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus R-CHOP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 52: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus R-CVP

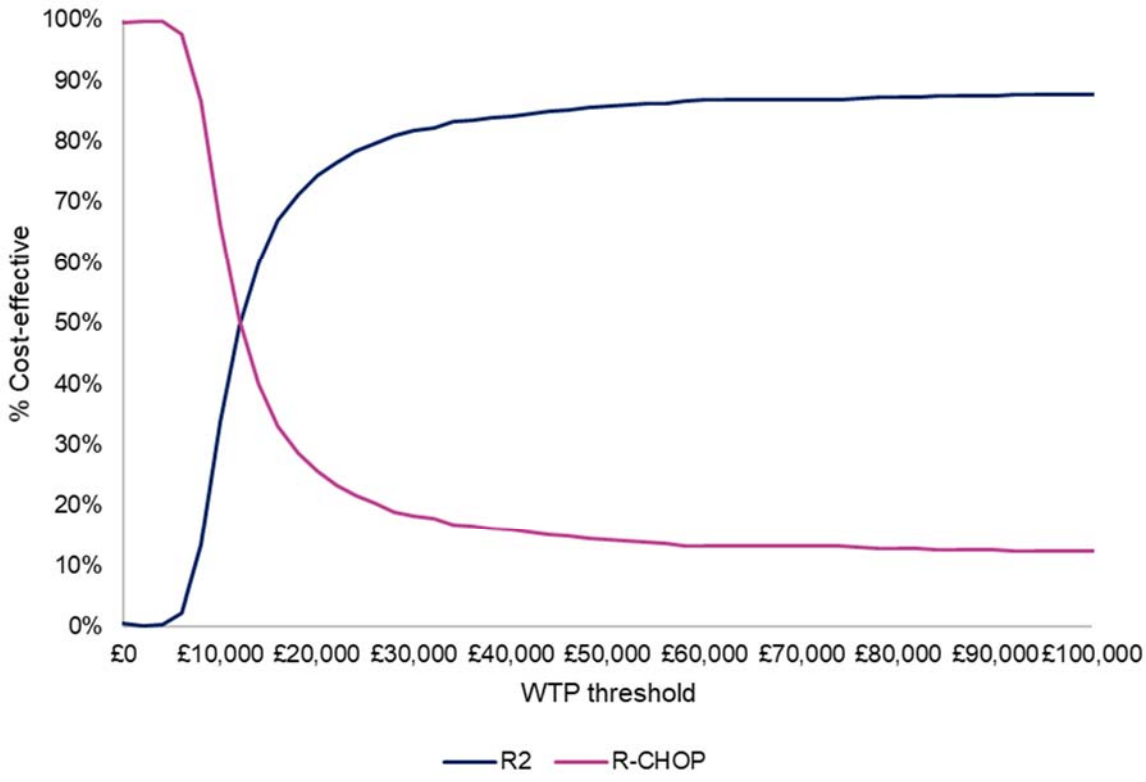
Key: CVP, cyclophosphamide, vincristine and prednisolone; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 53: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus O-Benda

Key: Benda, bendamustine; CVP, cyclophosphamide, vincristine and prednisolone; O, obinutuzumab; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

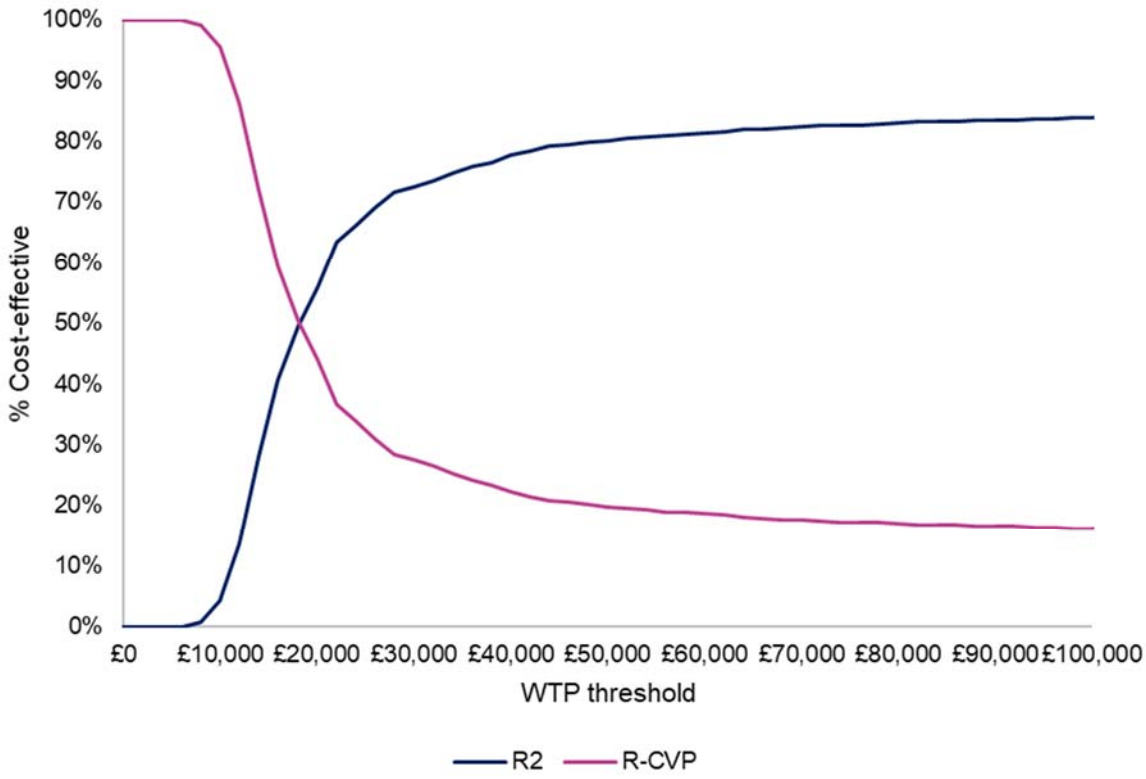
Figure 54–Figure 56 present the cost-effectiveness acceptability curves for R^2 versus comparators based on the 1,000 PSA iterations at different willingness-to-pay (WTP) thresholds. At the £30,000 WTP threshold, the probability of R^2 being cost-effective is 81.7%, 72.4% and 100% compared to R-CHOP, R-CVP and O-Benda, respectively.

Figure 54: Cost-effectiveness acceptability curve – R^2 vs R-CHOP



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 55: Cost-effectiveness acceptability curve – R^2 vs R-CVP



Key: CVP, cyclophosphamide, vincristine and prednisolone; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 56: Cost-effectiveness acceptability curve – R^2 vs O-Benda

Key: Benda, bendamustine; O, obinutuzumab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

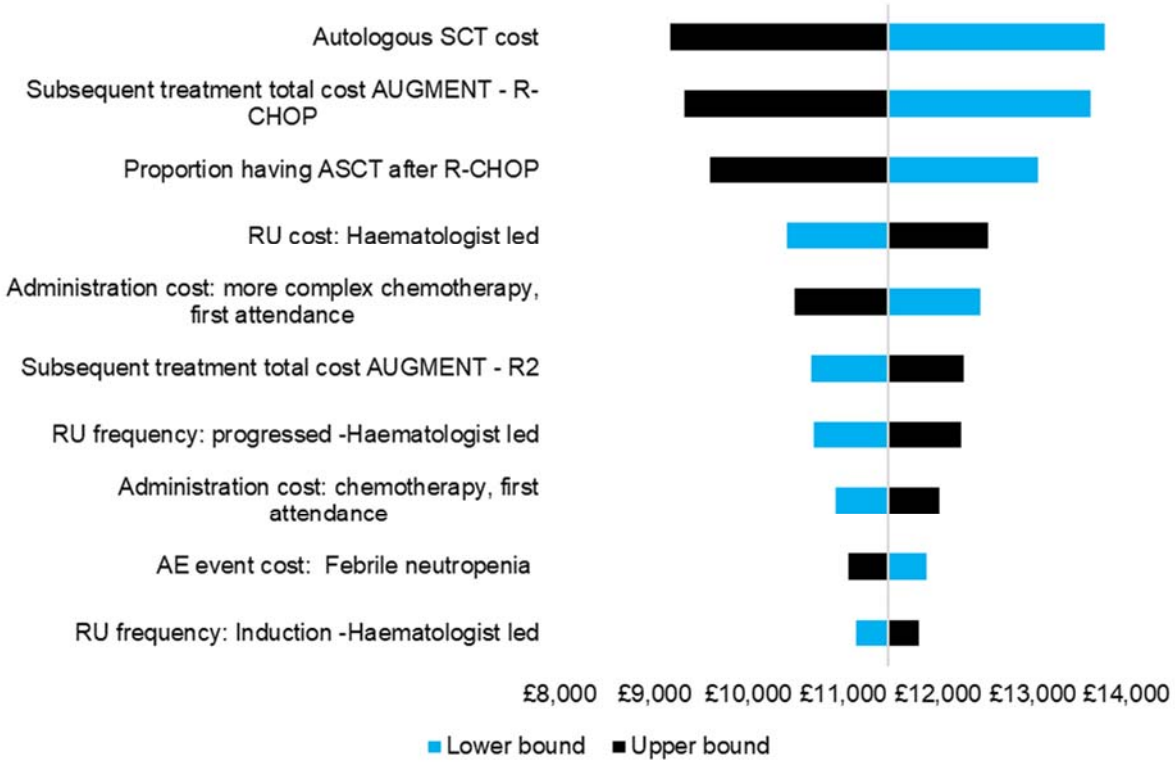
Deterministic sensitivity analysis

Figure 57–Figure 59 present the tornado diagrams showing the parameters with the greatest impact on the results with descending sensitivity from one-way sensitivity analysis (OWSA), when their values were set to their upper and lower limits of the confidence intervals reported in Section B.3.6.

The parameters that had the largest impact on the ICER for R^2 versus R-CHOP was the cost of SCT, proportion of patients who receive SCT, subsequent treatment costs for R-CHOP and resource use. The ICER ranged from £9,177 to £13,766, demonstrating that values tested at their upper and lower bounds still produced ICERs below £30,000. Similarly, with the R-CVP comparison, the parameters that had the largest impact were subsequent treatment costs (including ASCT cost, which is modelled as a subsequent therapy in addition to consolidation) and resource use, with all ICERs remaining under £30,000. For O-Benda, as the mean ICER is within

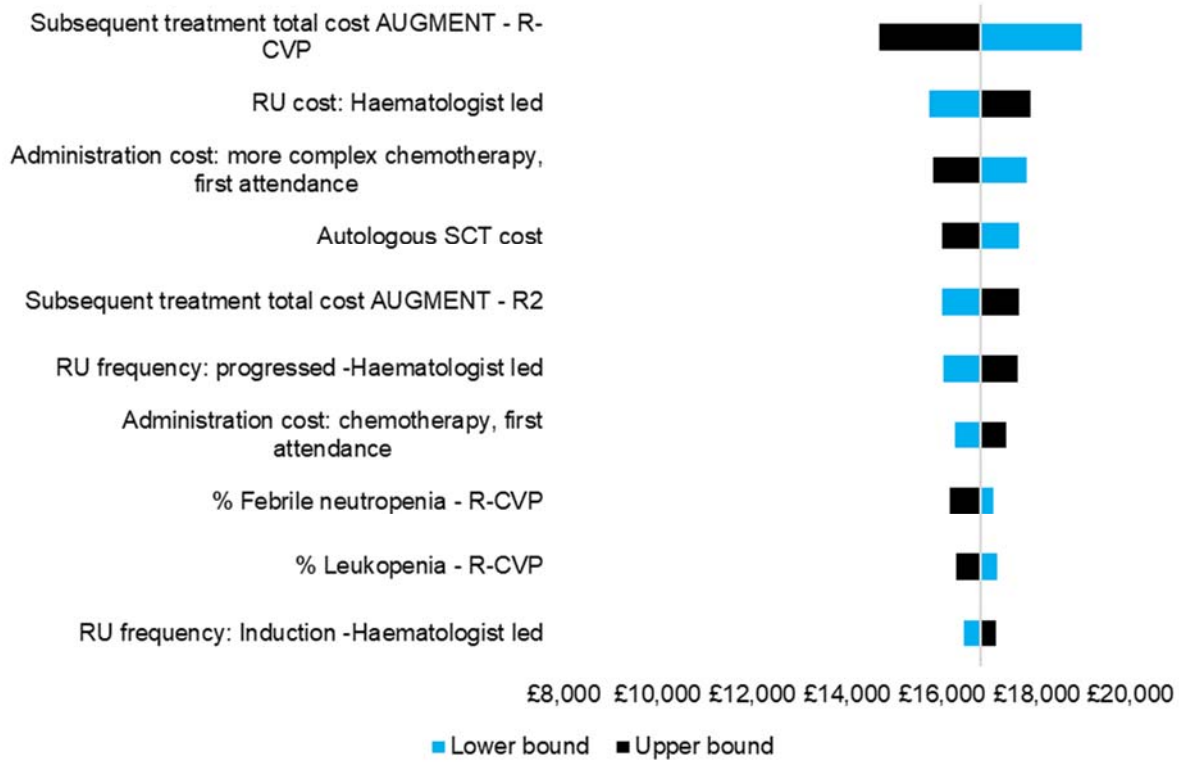
the southwest quadrant of the cost-effectiveness plane, the OWSA was ran on the NMB instead of the ICER. Similar parameters had the largest effect on the NMB, but all results resulted in R^2 being more cost effective than O-Benda.

Figure 57: Tornado diagram showing OWSA results on ICER – R^2 vs R-CHOP



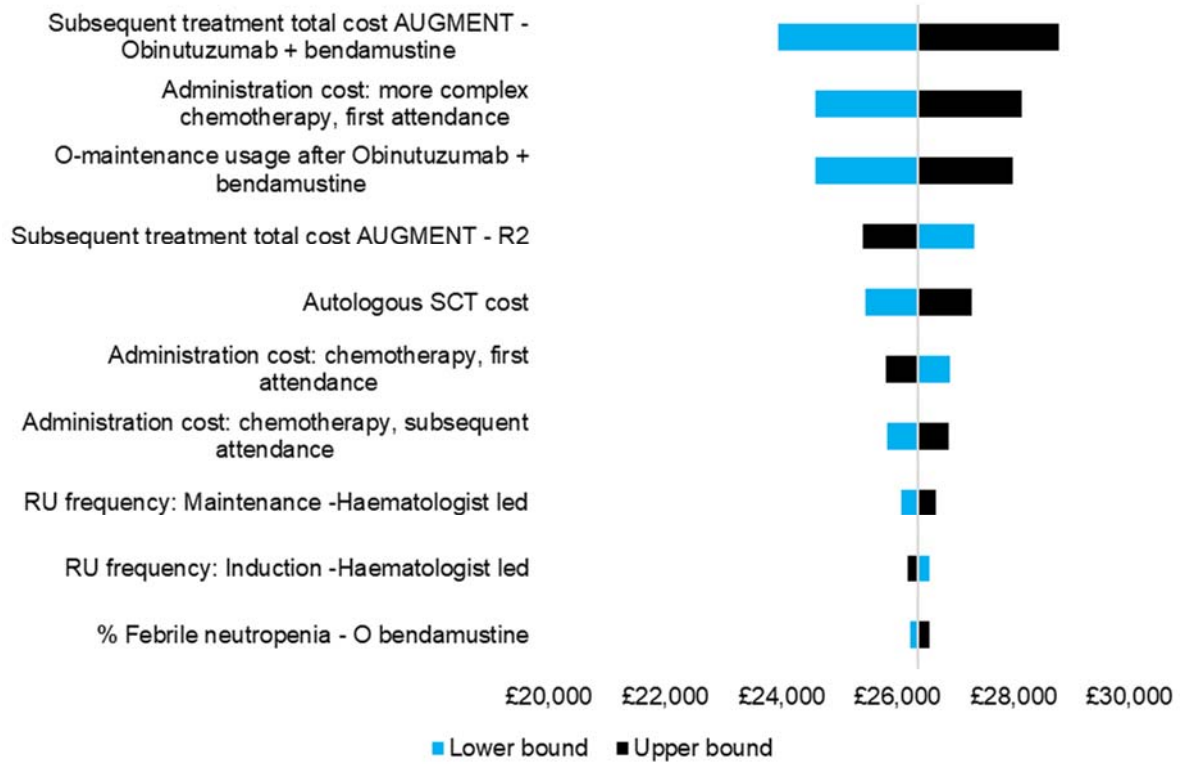
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use.

Figure 58: Tornado diagram showing OWSA results on ICER – R^2 vs R-CVP



Key: CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 59: Tornado diagram showing OWSA results on NMB – R^2 vs O-Benda



Key: Benda, bendamustine; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; O, obinutuzumab; OWSA, one-way sensitivity analysis; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Scenario analysis

Table 68–Table 70 present the scenario analysis performed to assess structural uncertainty within the model. The majority of the scenarios tested resulted in ICERs below the £30,000 threshold for the comparison with R-CHOP and R-CVP with the exception of a 5-year time horizon for R^2 versus R-CVP. Given the indolent nature of lymphoma, a 5-year time horizon is considered too short to capture any life-time benefit. For the comparison with O-Benda, most scenarios were in line with the base case whereby R^2 was the more cost-effective treatment, with the exception of R^2 OS extrapolations using the Gompertz distribution. The Gompertz distribution results in implausible survival rates assuming that approximately 95% of patients have died by 5 years, at which time the treatment waning effect is implemented.

Table 68: Results of scenario analysis (with PAS) – R^2 vs R-CHOP

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Time horizon	40 years	5 years	[redacted]	[redacted]	£29,070
		10 years	[redacted]	[redacted]	£16,202
		20 years	[redacted]	[redacted]	£12,306
No half cycle correction	Applied	Not applied	[redacted]	[redacted]	£11,412
Discount rate for costs	3.50%	0.0%	[redacted]	[redacted]	£12,424
		6.0%	[redacted]	[redacted]	£11,016
Discount rate for QALYs	3.50%	0.0%	[redacted]	[redacted]	£8,174
		6.0%	[redacted]	[redacted]	£14,074
Source of adverse event frequencies: literature	1L trial	Literature	[redacted]	[redacted]	£13,822
Apply comparator hazard to R^2 arms: after	5 years	3 years	[redacted]	[redacted]	£13,746
		10 years	[redacted]	[redacted]	£10,183
Use lenalidomide trial RDI	IPD	RDI	[redacted]	[redacted]	£9,128
Use rituximab trial RDI (R^2)	IPD	RDI	[redacted]	[redacted]	£11,153
Use comparator arm RDIs	IPD	RDI	[redacted]	[redacted]	£11,379
Vial costing option	Method of moments	Dose banding	[redacted]	[redacted]	£11,452
		No wastage	[redacted]	[redacted]	£11,526
Exclude wastage of prednisone	Include wastage	Exclude wastage	[redacted]	[redacted]	£11,472
Use of rituximab biosimilar in clinical practice	100%	50.0%	[redacted]	[redacted]	£11,493
		75.0%	[redacted]	[redacted]	£11,482
Use of subcutaneous rituximab during maintenance	100%	90.0%	[redacted]	[redacted]	£11,494
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied	[redacted]	[redacted]	£14,783

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Source for subsequent treatment mean durations	HMRN	AUGMENT			£12,349
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18			£14,276
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.			£13,117
Do not apply age-adjusted utility values	Applied	Not applied			£10,791
Do not apply adverse event disutility values	Applied	Not applied			£11,492
Cap utilities at general population equivalent utility	Not applied	Applied			£11,977
Distribution for R^2 ToT	KM	Exponential			£4,398
		GenGamma			£10,404
		Gompertz			£11,923
		Log-logistic			£10,881
		Log-normal			£8,312
		Weibull			£10,764
Distribution for R-CHOP ToT	KM	Exponential			£11,927
		GenGamma			£11,352
		Gompertz			£11,603
		Log-logistic			£11,992
		Log-normal			£12,420
		Weibull			£11,918

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Distribution for R^2 PFS	KM+comparator hazard	Exponential	[redacted]	[redacted]	£13,190
		GenGamma	[redacted]	[redacted]	£14,257
		Gompertz	[redacted]	[redacted]	£14,891
		Log-logistic	[redacted]	[redacted]	£13,747
		Log-normal	[redacted]	[redacted]	£12,462
		Weibull	[redacted]	[redacted]	£14,811
Distribution for R-CHOP PFS	GenGamma	Exponential	[redacted]	[redacted]	£12,434
		GenGamma	[redacted]	[redacted]	£11,471
		Gompertz	[redacted]	[redacted]	£10,129
		Log-logistic	[redacted]	[redacted]	£10,282
		Log-normal	[redacted]	[redacted]	£10,223
		Weibull	[redacted]	[redacted]	£10,933
Distribution for R^2 OS	Weibull	Exponential	[redacted]	[redacted]	£10,839
		GenGamma	[redacted]	[redacted]	£11,700
		Gompertz	[redacted]	[redacted]	£12,726
		Log-logistic	[redacted]	[redacted]	£11,172
		Log-normal	[redacted]	[redacted]	£9,994
		Weibull	[redacted]	[redacted]	£11,471
Distribution for R-CHOP OS	Weibull	Exponential	[redacted]	[redacted]	£13,216
		GenGamma	[redacted]	[redacted]	£10,255
		Gompertz	[redacted]	[redacted]	£9,845
		Log-logistic	[redacted]	[redacted]	£10,332
		Log-normal	[redacted]	[redacted]	£9,983
		Weibull	[redacted]	[redacted]	£11,471

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Distribution for R^2 TTNLT	Log-normal	Exponential	[redacted]	[redacted]	£11,614
		GenGamma	[redacted]	[redacted]	£11,680
		Gompertz	[redacted]	[redacted]	£11,711
		Log-logistic	[redacted]	[redacted]	£11,630
		Log-normal	[redacted]	[redacted]	£11,471
		Weibull	[redacted]	[redacted]	£11,722
Distribution for R-CHOP TTNLT	Log-normal	Exponential	[redacted]	[redacted]	£11,327
		GenGamma	[redacted]	[redacted]	£11,466
		Gompertz	[redacted]	[redacted]	£11,443
		Log-logistic	[redacted]	[redacted]	£11,331
		Log-normal	[redacted]	[redacted]	£11,471
		Weibull	[redacted]	[redacted]	£11,187
Distribution for TTNLT using unstratified curves	Stratified, Log-normal	Exponential	[redacted]	[redacted]	£11,696
		GenGamma	[redacted]	[redacted]	£11,393
		Gompertz	[redacted]	[redacted]	£11,458
		Log-logistic	[redacted]	[redacted]	£10,989
		Log-normal	[redacted]	[redacted]	£11,455
		Weibull	[redacted]	[redacted]	£11,438

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; NHS, National Health Service; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; ToT, time on treatment; TTNLT, time to next antilymphoma treatment.

Table 69: Results of scenario analysis (with PAS) – R^2 vs R-CVP

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Time horizon	40 years	5 years			£45,317
		10 years			£24,514
		20 years			£18,120
No half cycle correction	Applied	Not applied			£16,746
Discount rate for costs	3.50%	0.0%			£17,768
		6.0%			£16,357
Discount rate for QALYs	3.50%	0.0%			£11,976
		6.0%			£20,636
Source of adverse event frequencies: literature	1L trial	Literature			£18,393
Apply comparator hazard to R^2 arms: after	5 years	3 years			£20,471
		10 years			£14,656
Use lenalidomide trial RDI	IPD	RDI			£14,467
Use rituximab trial RDI (R^2)	IPD	RDI			£16,496
Use comparator arm RDIs	IPD	RDI			£16,723
Vial costing option	Method of moments	Dose banding			£16,794
		No wastage			£16,819
Exclude wastage of prednisone	Include wastage	Exclude wastage			£16,815
Use of rituximab biosimilar in clinical practice	100%	50.0%			£16,836
		75.0%			£16,825
Use of subcutaneous rituximab during maintenance	100%	90.0%			£16,837
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied			£20,131

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Source for subsequent treatment mean durations	HMRN	AUGMENT	[redacted]	[redacted]	£17,693
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18	[redacted]	[redacted]	£17,810
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.	[redacted]	[redacted]	£19,230
Do not apply age-adjusted utility values	Applied	Not applied	[redacted]	[redacted]	£15,816
Do not apply adverse event disutility values	Applied	Not applied	[redacted]	[redacted]	£16,820
Cap utilities at general population equivalent utility	Not applied	Applied	[redacted]	[redacted]	£17,557
Distribution for R^2 ToT	KM	Exponential	[redacted]	[redacted]	£9,731
		GenGamma	[redacted]	[redacted]	£15,746
		Gompertz	[redacted]	[redacted]	£17,266
		Log-logistic	[redacted]	[redacted]	£16,223
		Log-normal	[redacted]	[redacted]	£13,650
		Weibull	[redacted]	[redacted]	£16,106
Distribution for R-CVP ToT	KM	Exponential	[redacted]	[redacted]	£17,270
		GenGamma	[redacted]	[redacted]	£16,694
		Gompertz	[redacted]	[redacted]	£16,947
		Log-logistic	[redacted]	[redacted]	£17,335
		Log-normal	[redacted]	[redacted]	£17,760
		Weibull	[redacted]	[redacted]	£17,261

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Distribution for R^2 PFS	KM+comparator hazard	Exponential	[redacted]	[redacted]	£18,623
		GenGamma	[redacted]	[redacted]	£19,746
		Gompertz	[redacted]	[redacted]	£20,413
		Log-logistic	[redacted]	[redacted]	£19,209
		Log-normal	[redacted]	[redacted]	£17,856
		Weibull	[redacted]	[redacted]	£20,329
Distribution for R-CVP PFS	GenGamma	Exponential	[redacted]	[redacted]	£17,842
		GenGamma	[redacted]	[redacted]	£16,814
		Gompertz	[redacted]	[redacted]	£15,415
		Log-logistic	[redacted]	[redacted]	£15,567
		Log-normal	[redacted]	[redacted]	£15,486
		Weibull	[redacted]	[redacted]	£16,249
Distribution for R^2 OS	Weibull	Exponential	[redacted]	[redacted]	£15,745
		GenGamma	[redacted]	[redacted]	£17,201
		Gompertz	[redacted]	[redacted]	£18,935
		Log-logistic	[redacted]	[redacted]	£16,309
		Log-normal	[redacted]	[redacted]	£14,318
		Weibull	[redacted]	[redacted]	£16,814
Distribution for R-CVP OS	Weibull	Exponential	[redacted]	[redacted]	£19,774
		GenGamma	[redacted]	[redacted]	£14,747
		Gompertz	[redacted]	[redacted]	£14,051
		Log-logistic	[redacted]	[redacted]	£14,879
		Log-normal	[redacted]	[redacted]	£14,286
		Weibull	[redacted]	[redacted]	£16,814
Distribution for R^2 TTNLT	Log-normal	Exponential	[redacted]	[redacted]	£17,002

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
		GenGamma			£17,096
		Gompertz			£17,135
		Log-logistic			£17,028
		Log-normal			£16,814
		Weibull			£17,151
Distribution for R-CVP TTNLT	Log-normal	Exponential			£16,741
		GenGamma			£16,799
		Gompertz			£16,756
		Log-logistic			£16,659
		Log-normal			£16,814
		Weibull			£16,556
Distribution for TTNLT using unstratified curves	Stratified, Log-normal	Exponential			£17,158
		GenGamma			£16,738
		Gompertz			£16,772
		Log-logistic			£16,317
		Log-normal			£16,793
		Weibull			£16,859

Key: CVP, cyclophosphamide, vincristine and prednisolone; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; NHS, National Health Service; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; ToT, time on treatment; TTNLT, time to next antilymphoma treatment.

Table 70: Results of scenario analysis (with PAS) – R^2 vs O-Benda

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Time horizon	40 years	5 years			£19,586,588 (SW)
		10 years			£17,899,253 (SW)
		20 years			£17,078,850 (SW)
No half cycle correction	Applied	Not applied			£16,986,895 (SW)
Discount rate for costs	3.50%	0.0%			£18,190,301 (SW)
		6.0%			£16,296,284 (SW)
Discount rate for QALYs	3.50%	0.0%			£15,168,854 (SW)
		6.0%			£18,023,349 (SW)
Source of adverse event frequencies: literature	1L trial	Literature			£16,960,557 (SW)
Apply comparator hazard to R^2 arms: after	5 years	3 years			£16,960,557 (SW)
		10 years			£16,960,557 (SW)
Use lenalidomide trial RDI	IPD	RDI			£18,655,104 (SW)
Use rituximab trial RDI (R^2)	IPD	RDI			£17,182,648 (SW)
Use comparator arm RDIs	IPD	RDI			£17,090,908 (SW)
Vial costing option	Method of moments	Dose banding			£17,198,580 (SW)
		No wastage			£17,254,199 (SW)
Exclude wastage of prednisone	Include wastage	Exclude wastage			£16,960,475 (SW)
Source of R^2 dosing schedule (rituximab): MAGNIFY	SmPC	MAGNIFY			£16,925,967 (SW)
Use of rituximab biosimilar in clinical practice	100%	50.0%			£16,689,611 (SW)
		75.0%			£16,825,084 (SW)
Use of subcutaneous rituximab during maintenance	100%	90.0%			£16,960,557 (SW)

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied	[redacted]	[redacted]	£14,442,218 (SW)
Source for subsequent treatment mean durations	HMRN	AUGMENT	[redacted]	[redacted]	£16,299,503 (SW)
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18	[redacted]	[redacted]	£16,246,999 (SW)
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.	[redacted]	[redacted]	£7,319,346 (SW)
Do not apply age-adjusted utility values	Applied	Not applied	[redacted]	[redacted]	£16,662,545 (SW)
Do not apply adverse event disutility values	Applied	Not applied	[redacted]	[redacted]	£21,038,562 (SW)
Cap utilities at general population equivalent utility	Not applied	Applied	[redacted]	[redacted]	£16,960,557 (SW)
Distribution for R^2 ToT	KM	Exponential	[redacted]	[redacted]	£18,787,157 (SW)
		GenGamma	[redacted]	[redacted]	£17,658,726 (SW)
		Gompertz	[redacted]	[redacted]	£16,617,577 (SW)
		Log-logistic	[redacted]	[redacted]	£17,577,875 (SW)
		Log-normal	[redacted]	[redacted]	£18,487,535 (SW)
		Weibull	[redacted]	[redacted]	£17,607,459 (SW)
Distribution for R^2 PFS	KM+comparator hazard	Exponential	[redacted]	[redacted]	£2,443,048 (SW)
		GenGamma	[redacted]	[redacted]	Dominant
		Gompertz	[redacted]	[redacted]	Dominant
		Log-logistic	[redacted]	[redacted]	Dominant
		Log-normal	[redacted]	[redacted]	Dominant
		Weibull	[redacted]	[redacted]	£1,664,836 (SW)

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Distribution for O-Benda PFS	Log-normal	Exponential			£17,152,915 (SW)
		GenGamma			£37,732,766 (SW)
		Gompertz			£22,461,100 (SW)
		Log-logistic			£11,835,229 (SW)
		Log-normal			£16,960,557 (SW)
		Weibull			£10,071,634 (SW)
Distribution for R^2 OS	Equal to comparator	Exponential			Dominant
		GenGamma			Dominant
		Gompertz			£10,316 (SW)
		Log-logistic			£49,668 (SW)
		Log-normal			Dominant
		Weibull			£30,593 (SW)
Distribution for O-Benda OS	Exponential	Exponential			£16,960,557 (SW)
		GenGamma			£16,933,370 (SW)
		Gompertz			£16,933,212 (SW)
		Log-logistic			£16,938,610 (SW)
		Log-normal			£16,935,461 (SW)
		Weibull			£16,949,260 (SW)
Distribution for R^2 TTNLT	Log-normal	Exponential			£1,278,172 (SW)
		GenGamma			£56,386,246 (SW)
		Gompertz			Dominant
		Log-logistic			£3,391,741 (SW)
		Log-normal			£16,960,557 (SW)
		Weibull			£2,231,995 (SW)

Parameter	Base case	Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Distribution for O-Benda TTNLT	Log-normal	Exponential	[redacted]	[redacted]	Dominant
		GenGamma	[redacted]	[redacted]	£10,019,581 (SW)
		Gompertz	[redacted]	[redacted]	£3,078,660 (SW)
		Log-logistic	[redacted]	[redacted]	Dominant
		Log-normal	[redacted]	[redacted]	£16,960,557 (SW)
		Weibull	[redacted]	[redacted]	Dominant

Key: Benda, bendamustine; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; NHS, National Health Service; O, obinutuzumab; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; SW, southwest; ToT, time on treatment; TTNLT, time to next antilymphoma treatment.

Scenario: R^2 versus R-CHOP/R-CVP using ITC (van Oers)-based data

Table 71 presents the results of the scenario comparing R^2 to R-CHOP/CVP using van Oers-based evidence. Further details of how this scenario is derived in the model is discussed in Appendix T.

Table 71: Scenario analysis: R^2 versus R-CHOP/CVP: trial based (with PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R-CHOP /CVP	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£10,880

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PAS, patient access scheme; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab.

Summary of sensitivity analyses results

The probabilistic results are consistent with the deterministic results for comparisons of R^2 to R-CHOP, R-CVP and O-Benda. The OWSA identified parameters that had the biggest impact on the ICER or NMB and qualified the impacts of taking extreme values of each parameter on the cost-effectiveness results. The OWSA showed that the cost-effectiveness results were not overly sensitive to these parameters, with all ICERs consistently remaining cost effective versus the comparator.

A wide range of scenario analyses were performed on key model assumptions and alternative choices to test the robustness of base case results. The majority of the results remained under the £30,000 threshold, with only implausible scenarios resulting in a greater ICER.

The cost-effectiveness acceptability curve based on 1,000 runs estimates that the probability of R^2 being cost-effective at the £30,000 WTP threshold is 81.7%, 72.4% and 100% compared to R-CHOP, R-CVP and O-Benda, respectively.

B.3.9. Subgroup analysis

In line with the final scope, no subgroups were modelled within the economic evaluation.

B.3.10. Validation

Validation of cost-effectiveness analysis

As described in Section B.1.3 an advisory board consisting of six clinicians and two UK economic experts was held to validate the following key aspects:[133]

- The model structure and its appropriateness to reflect the clinical pathway
- The approach used to inform the efficacy of R^2 to the comparators
- Extrapolation of survival beyond the trial period
- The use and clinical validity of utilities derived from AUGMENT versus those in the literature
- Subsequent treatment usage

The efficacy data sources were used to validate the models survival projections. Furthermore, the cost-effectiveness model itself was quality assured by a health economist not involved in the model building who reviewed the model for coding errors, inconsistencies and plausibility of inputs. The model was also subject to stress testing of extreme scenarios to test for known modelling errors and questioning of assumptions.

Validation

Validation compared the PFS, OS and TTNLT KM data from the efficacy source with the PFS, OS and TTNLT outputs from the model. All modelled outcomes appear to be consistent with the observed data with the exception PFS in Year 1 for R-CHOP/R-CVP, which under-estimated the observed PFS. However, this becomes more aligned with observed data from Year 2 onwards. Median PFS for R-CHOP and R-CVP is also slightly over estimated from the modelled outputs.

In addition to comparing the endpoints with the observed data individually, the endpoints were also validated to ensure that they were producing sensible extrapolations and did not cross the curves of other endpoints at implausible time

points. The distribution of patients over time for the non-R-refractory population demonstrates the effect that TTNLT has on the OS. Given the assumption of equal PFS hazard after the final available data point, there is a larger gap between progression and patients receiving their next treatment with R^2 than R-CHOP/R-CVP, which demonstrates the impact of TTNLT, determined to be a better reflection of clinically meaningful progression based on the guidance of the clinical community. The base case choices of the curves that estimate the long-term outcomes have been chosen so that no implausible crossing of curves appears within treatment estimates.

B.3.11. Interpretation and conclusions of economic evidence

The economic analysis is based on a de novo economic model with a structure designed to reflect relapsed/refractory FL and MZL as simply as possible while still capturing the relevant outcomes. The structure employed is consistent with previous appraisal model structures in the same disease area.

The model incorporates the most relevant efficacy and safety data and uses statistical techniques to demonstrate the comparative efficacy of R^2 with R-CHOP, R-CVP and O-Benda, accounting for any differences in patient populations that were possible to adjust for. The inclusion in the analysis of a real-world data source, the UK HMRN, addresses some of the limitations of the available literature-based evidence for specific comparators. This allowed for comparisons to be reflective of UK clinical practice with comparative efficacy adjusted for so that R^2 reflected a UK real-world cohort. The use of TTNLT in the model meant that the results used a more clinically appropriate endpoint that is more meaningful in terms of costs, quality of life and patient survival. The model makes use of results from trial-based utility and safety analysis with parameters tested in scenario analysis to assess the uncertainty in these results. Key model assumptions were validated with clinicians, and model outputs were validated to ensure that implausible projections were not selected.

The main limitation of the model is the lack of direct comparative efficacy with key comparators. However, analysis has been conducted utilising the most appropriate available, with all adjustments appropriate being made in order to reduce the heterogeneity between populations. Additionally, there is a lack of evidence available

for the MZL population. However, clinicians agreed that these patients would be treated in the same way as FL patients. A targeted literature review confirmed that endpoints relevant in MZL are consistent with endpoints from the FL population, justifying the assumption that these populations could be pooled. Due to the indolent nature of the disease, the immaturity of the OS data is a further limitation.

Despite these limitations, the model demonstrated that R^2 was cost effective versus the comparators of relevance. A benefit associated with R^2 that is hard to capture within the cost-effective framework is the clinical need for additional treatment options. Patients are prescribed treatment based on their prior therapy, prior response, age and general fitness, and some patients will receive multiple rounds of treatment over many years. Existing treatment options leave substantial unmet need across the full spectrum of relapsed/refractory FL and MZL patient populations, resulting in a high clinical need for additional treatment options in these disease settings. R^2 represents a novel treatment modality offering a cost-effective alternative to repeating patients' exposure to current chemotherapy-based treatment options and their associated toxicities, with benefit anticipated through changing both mechanism of action and toxicity profile.

B.4. References

1. Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0069) March 2019 2019. (Updated: June 2019) Data on file.
2. Celgene Ltd. Kim Linton clinical validation. 12 June 2019 2019. Data on file.
3. Chiu H, Trisal P, Bjorklund C, et al. Combination lenalidomide-rituximab immunotherapy activates anti-tumour immunity and induces tumour cell death by complementary mechanisms of action in follicular lymphoma. *Br J Haematol*. 2019; 185(2):240-53.
4. Celgene. Draft Summary of Product Characteristics (SPC). Revlimid® (Lenalidomide). 2019. Accessed: 19/02/2019.
5. Morschhauser F, Fowler NH, Feugier P, et al. Rituximab plus Lenalidomide in Advanced Untreated Follicular Lymphoma. *New England Journal of Medicine*. 2018; 379(10):934-47.
6. Rancea M, Will A, Borchmann P, et al. Sixteenth biannual report of the Cochrane Haematological Malignancies Group: focus on Non-Hodgkin's lymphoma. *J Natl Cancer Inst*. 2014; 106(8).
7. Haematological Malignancy Research Network (HMRN). QuickStats. 2019. Available at: https://www.hmrn.org/statistics/quick. Accessed: 30 April 2019.
8. Cancer Research UK. Low grade NHL. 2018. (Updated: 23 January 2018) Available at: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/low-grade. Accessed: 30 April 2019.
9. Office of National Statistics (ONS). Cancer Registration Statistics, England, 2017. 2019. (Updated: 26 April 2019) Available at: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland/2017/finalreferencetables.xlsx. Accessed: 1 May 2019.
10. Lymphoma Action. MALT lymphoma (extranodal marginal zone lymphoma). 2018. Available at: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-08/LYMweb0218MALT2017v1.2.pdf. Accessed: 1 May 2019.
11. Lymphoma Action. Splenic marginal zone lymphoma. 2018. Available at: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-06/LYMweb0220SplenicMZ2017v1.2.pdf. Accessed: 1 May 2019.
12. Lymphoma Action. Nodal marginal zone lymphoma. 2019. Available at: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-06/LYMweb0219NodalMZ2017v1.2.pdf. Accessed: 1 May 2019.
13. Kahl BS and Yang DT. Follicular lymphoma: evolving therapeutic strategies. *Blood*. 2016; 127(17):2055-63.
14. 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*. 2016; 27(suppl 5):v83-v90.
15. Dreyling M, Thieblemont C, Gallamini A, et al. ESMO Consensus conferences: guidelines on malignant lymphoma. part 2: marginal zone lymphoma, mantle cell lymphoma, peripheral T-cell lymphoma. *Ann Oncol*. 2013; 24(4):857-77.
16. Cancer Research UK. Follicular lymphoma. 2019. (Updated: 11 February 2019) Available at: https://about-cancer.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/follicular-

lymphoma?_ga=2.13534143.986479336.1555421727-16612518.1549888859. Accessed: 1 May 2019.
17. McNamara C, Davies J, Dyer M, et al. Guidelines on the investigation and management of follicular lymphoma. *Br J Haematol*. 2012; 156(4):446-67.
18. HMRN. Incidence. 2019. Available at: https://www.hmrn.org/statistics/incidence. Accessed: 1 May 2019.
19. HMRN. Patient's age and treatment for haematological malignancy: a report from the Haematological Malignancy Research Network (HMRN). 2014. (Updated: June 2014) Available at: https://www.hmrn.org/Download.aspx?target=document&id=41. Accessed: 1 May 2019.
20. HMRN. Survival: Follicular lymphoma. 2019. Available at: https://www.hmrn.org/statistics/survival. Accessed: 1 May 2019.
21. Cancer Research UK. Non-Hodgkin lymphoma: Survival. 2019. (Updated: 19 February 2018) Available at: https://about-cancer.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/survival. Accessed: 1 May 2019.
22. Provencio M, Sabin P, Gomez-Codina J, et al. Impact of treatment in long-term survival patients with follicular lymphoma: A Spanish Lymphoma Oncology Group registry. *PLoS One*. 2017; 12(5):e0177204.
23. Johnson PW, Rohatiner AZ, Whelan JS, et al. Patterns of survival in patients with recurrent follicular lymphoma: a 20-year study from a single center. *J Clin Oncol*. 1995; 13(1):140-7.
24. Pettengell R, Donatti C, Hoskin P, et al. The impact of follicular lymphoma on health-related quality of life. *Ann Oncol*. 2008; 19(3):570-6.
25. 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*. 2019; 184(4):660-3.
26. El-Zimaity H YM. Marginal Zone B-cell Lymphoma. *Medscape*. 2014.
27. Olszewski AJ and Castillo JJ. Survival of patients with marginal zone lymphoma: analysis of the Surveillance, Epidemiology, and End Results database. *Cancer*. 2013; 119(3):629-38.
28. Denlinger NM, Epperla N and William BM. Management of relapsed/refractory marginal zone lymphoma: focus on ibrutinib. *Cancer Manag Res*. 2018; 10:615-24.
29. Andorsky DJ, Kolibaba KS, Assouline S, et al. An open-label phase 2 trial of entospletinib in indolent non-Hodgkin lymphoma and mantle cell lymphoma. 2019; 184(2):215-22.
30. de Vos S, Swinnen LJ, Wang D, et al. Venetoclax, bendamustine, and rituximab in patients with relapsed or refractory NHL: a phase Ib dose-finding study. *Ann Oncol*. 2018; 29(9):1932-8.
31. Dreyling M SA, Mollica L, Leppä S, Follows GA, Lenz G, Kim W, et al. Copanlisib in patients with relapsed or refractory indolent b-cell lymphoma (CHRONOS-1). *Haematological Oncology*. 2017; 35(S2):119; Abstract 08.
32. 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*. 2017; 28(9):2169-78.
33. Flinn IW, Kahl BS, Leonard JP, et al. Idelalisib, a selective inhibitor of phosphatidylinositol 3-kinase-delta, as therapy for previously treated indolent non-Hodgkin lymphoma. *Blood*. 2014; 123(22):3406-13.

34. Gopal AK, Kahl BS, de Vos S, et al. PI3Kδ Inhibition by Idelalisib in Patients with Relapsed Indolent Lymphoma. *New England Journal of Medicine*. 2014; 370(11):1008-18.
35. Kahl BS, Bartlett NL, Leonard JP, et al. Bendamustine is effective therapy in patients with rituximab-refractory, indolent B-cell non-Hodgkin lymphoma: results from a Multicenter Study. *Cancer*. 2010; 116(1):106-14.
36. Rummel MJ, Al-Batran SE, Kim SZ, 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.
37. Witzig TE, Wiernik PH, Moore T, et al. Lenalidomide oral monotherapy produces durable responses in relapsed or refractory indolent non-Hodgkin's Lymphoma. *J Clin Oncol*. 2009; 27(32):5404-9.
38. Wang HI, Roman E, Crouch S, et al. A Generic Model for Follicular Lymphoma: Predicting Cost, Life Expectancy, and Quality-Adjusted-Life-Year Using UK Population-Based Observational Data. *Value Health*. 2018; 21(10):1176-85.
39. Walker MS, Stepanski EJ, Reyes C, et al. Symptom Burden and Quality of Life in Patients with Follicular Lymphoma undergoing Maintenance Treatment with Rituximab Compared with Observation. *Ther Adv Hematol*. 2011; 2(3):129-39.
40. Lymphoma Action. Follicular lymphoma. 2019. Available at: https://lymphoma-action.org.uk/types-lymphoma-non-hodgkin-lymphoma/follicular-lymphoma. Accessed: 1 May 2019.
41. Oerlemans S, Issa DE, van den Broek EC, et al. Impact of therapy and disease-related symptoms on health-related quality of life in patients with follicular lymphoma: results of the population-based PHAROS-registry. *Eur J Haematol*. 2014; 93(3):229-38.
42. Kang D, Cho J, Kim IR, et al. Health-Related Quality of Life in Non-Hodgkin Lymphoma Survivors: A Prospective Cohort Study. *Cancer Res Treat*. 2018; 50(4):1051-63.
43. Cheung MC, Imrie KR, Friedlich J, et al. The impact of follicular (FL) and other indolent non-Hodgkin's lymphomas (NHL) on work productivity-a preliminary analysis. *Psychooncology*. 2009; 18(5):554-9.
44. National Institute of Health and Care Excellence (NICE). Treating follicular lymphoma. 2016. Available at: https://pathways.nice.org.uk/pathways/non-hodgkins-lymphoma#path=view%3A/pathways/non-hodgkins-lymphoma/treating-follicular-lymphoma.xml&content=view-node%3Anodes-first-line-treatment-for-early-stage-disease. Accessed: 13 February 2019.
45. National Institute for Health and Care Excellence (NICE). Treating mucosa-associated lymphoid tissue lymphoma. 2019. Available at: https://pathways.nice.org.uk/pathways/non-hodgkins-lymphoma/managing-non-hodgkins-lymphoma#content=view-index&path=view%3A/pathways/non-hodgkins-lymphoma/treating-mucosa-associated-lymphoid-tissue-lymphoma.xml. Accessed: 3 May 2019.
46. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®): B-Cell Lymphomas. 2019. (Updated: 18 June 2019) Available at: https://www.nccn.org/professionals/physician_gls/pdf/b-cell.pdf. Accessed: 18 July 2019.
47. National Institute for Health and Care Excellence (NICE). NG52: Non-Hodgkin's lymphoma: diagnosis and management. 2016. Available at: https://www.nice.org.uk/guidance/ng52. Accessed: 25 April 2019.

48. National Institute for Health and Care Excellence (NICE) TA472 Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab. 2017. Available at: https://www.nice.org.uk/guidance/ta472. Accessed: 05 January 2019.
49. Roche Registration GmbH. Summary of product characteristics. Gazyvaro 1,000 mg concentrate for solution for infusion. 10 December 2018 2018. Available at: https://www.ema.europa.eu/en/documents/product-information/gazyvaro-epar-product-information_en.pdf. Accessed: 5 May 2019.
50. National Institute for Health and Care Excellence (NICE). TA243: Rituximab for the first-line treatment of stage III-IV follicular lymphoma. 2012. Available at: https://www.nice.org.uk/guidance/ta243. Accessed: 26 April 2019.
51. National Institute for Health and Care Excellence (NICE). Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma. 2011. Available at: https://www.nice.org.uk/guidance/ta226/resources/rituximab-for-the-firstline-maintenance-treatment-of-follicular-nonhodgkins-lymphoma-pdf-82600313486533. Accessed: 5 May 2019.
52. National Institute for Health and Care Excellence (NICE) TA513 Obinutuzumab for untreated advanced follicular lymphoma. 2018. Available at: https://www.nice.org.uk/guidance/ta513. Accessed: 05 January 2019.
53. Lymphoma Research Foundation. Marginal Zone Lymphoma. 2018. (Updated: February 2018) Available at: https://www.lymphoma.org/wp-content/uploads/2018/07/MarginalZoneLymphomaF_Fact-Sheet_2018.pdf. Accessed: 5 May 2019.
54. National Institute for Health and Care Excellence (NICE). Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma: Final scope. 2018.
55. Fowler N. Frontline strategy for follicular lymphoma: are we ready to abandon chemotherapy? *Hematology Am Soc Hematol Educ Program*. 2016; 2016(1):277-83.
56. Casulo C and Barr PM. How I treat early-relapsing follicular lymphoma. *Blood*. 2019; 133(14):1540-7.
57. Conconi A TC, Cascione L, Torri V, Kiesewetter B, Margiotta-Casaluci G, Gaidano G, et al. Early progression of disease (POD24) as survival predictor in MALT lymphoma. *J Clin Oncol*. 2019; 37(15_suppl):7548.
58. 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. *J Clin Oncol*. 2019; 37(14):1188-99.
59. Celgene. MAGNIFY - updated response rate tables. 2018.
60. Celgene. SLR of the clinical efficacy of pharmacological therapies for relapsed/refractory indolent non-Hodgkin lymphoma. (Version 4) 30 May 2019 2019. Data on file.
61. Andorsky D, Coleman M, Yacoub A, et al. Response rate to lenalidomide plus rituximab (R2) as independent of number of prior lines of therapy: Interim analysis of initial phase of MAGNIFY phase IIIb study of R2 followed by maintenance in relapsed/refractory indolent NHL. *Journal of Clinical Oncology*. 2018; 36(15_suppl):7516-.
62. Leonard JP, Jung SH, Johnson J, et al. Randomized Trial of Lenalidomide Alone Versus Lenalidomide Plus Rituximab in Patients With Recurrent Follicular Lymphoma: CALGB 50401 (Alliance). *J Clin Oncol*. 2015; 33(31):3635-40.

63. van Oers MH, Klasa R, Marcus RE, 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.
64. 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*. 2016; 17(8):1081-93.
65. Tuscano JM, Dutia M, Chee K, et al. Lenalidomide plus rituximab can produce durable clinical responses in patients with relapsed or refractory, indolent non-Hodgkin lymphoma. *Br J Haematol*. 2014; 165(3):375-81.
66. Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007) 2018.
67. 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*. 1997; 15(3):1110-7.
68. Jaffe ES. The 2008 WHO classification of lymphomas: implications for clinical practice and translational research. *Hematology Am Soc Hematol Educ Program*. 2009:523-31.
69. Hainsworth JD. Rituximab as first-line and maintenance therapy for patients with indolent non-Hodgkin's lymphoma: interim follow-up of a multicenter phase II trial. *Semin Oncol*. 2002; 29(1 Suppl 2):25-9.
70. Ardeshna KM, Qian W, Smith P, et al. Rituximab versus a watch-and-wait approach in patients with advanced-stage, asymptomatic, non-bulky follicular lymphoma: an open-label randomised phase 3 trial. *Lancet Oncol*. 2014; 15(4):424-35.
71. Taverna C, Martinelli G, Hitz F, et al. Rituximab Maintenance for a Maximum of 5 Years After Single-Agent Rituximab Induction in Follicular Lymphoma: Results of the Randomized Controlled Phase III Trial SAKK 35/03. *J Clin Oncol*. 2016; 34(5):495-500.
72. Ghielmini M, Schmitz SF, Cogliatti SB, et al. Prolonged treatment with rituximab in patients with follicular lymphoma significantly increases event-free survival and response duration compared with the standard weekly x 4 schedule. *Blood*. 2004; 103(12):4416-23.
73. Coiffier B, Osmanov EA, Hong X, et al. Bortezomib plus rituximab versus rituximab alone in patients with relapsed, rituximab-naive or rituximab-sensitive, follicular lymphoma: a randomised phase 3 trial. *Lancet Oncol*. 2011; 12(8):773-84.
74. Celgene. A Phase 3b randomized study of lenalidomide (CC-5013) plus rituximab maintenance therapy followed by lenalidomide single-agent maintenance versus rituximab maintenance in subjects with relapsed/refractory follicular, marginal zone or mantle cell lymphoma. (MAGNIFY interim report (22 Nov 2017): CC-5013-NHL-008) 2017.
75. Celgene. Lenalidomide Plus Rituximab Followed by Lenalidomide Versus Rituximab Maintenance for Relapsed/Refractory Follicular, Marginal Zone or Mantle Cell Lymphoma. (MAGNIFY). 2019. (Updated: 15 March 2019) Available at: https://clinicaltrials.gov/ct2/show/NCT01996865. Accessed: 05 July 2019.

76. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – Demographic and Baseline Characteristics for IEE population. 04 June 2019 2018. Data on file
77. Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007 (updated tables)) 2018.
78. Leonard JP, Trneny M, Izutsu K, et al. HEALTH-RELATED QUALITY OF LIFE (HRQOL) IN RELAPSED/REFRACTORY (R/R) INDOLENT NHL IN THE PHASE 3 AUGMENT TRIAL OF RITUXIMAB (R) PLUS LENALIDOMIDE (R2) VERSUS R PLUS PLACEBO. 15-ICML International conference on malignant lymphoma. Lugano, Switzerland. 18-22 June 2019 2019. TBC.
79. Celgene. MAGNIFY - updated PFS tables. 10 Aug 2018 2018.
80. Celgene. MAGNIFY - updated PFS KM curve. 10 Aug 2018 2018.
81. Leonard JP, Trneny M, Izutsu K, et al. AUGMENT phase III study: Lenalidomide/rituximab (R2) improved efficacy over rituximab/placebo in relapsed/refractory follicular patients irrespective of POD24 status 15-ICML International Conference on Malignant Lymphoma. Lugano, Switzerland. 18–22 June, 2019 2019.
82. Celgene. MAGNIFY - updated subgroup analyses (ORR). 10 Aug 2018 2018. Data on file.
83. Celgene. MAGNIFY - updated subgroup analyses (CRR). 10 Aug 2018 2018. Data on file.
84. Clarostat. Statistical Report: Indirect Treatment Comparisons in Relapsed/Refractory Follicular Lymphoma and Marginal Zone Lymphoma. 11 June 2019 2019. Data on File.
85. Dias S WN, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2016. (Updated: September 2016) Available at: http://nicedsu.org.uk/wp-content/uploads/2017/05/TSD2-General-meta-analysis-corrected-2Sep2016v2.pdf. Accessed: 6 May 2019.
86. Guyot P, Ades AE, Ouwens MJ and Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012; 12:9.
87. Phillippo D, Ades A, Dias S, et al. NICE DSU Technical Support Document 18: Methods for Population-adjusted Indirect Comparisons in Submissions to NICE. 2016. Available at: https://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf. Accessed: 02 May 2019.
88. Rummel M, Kaiser U, Balser C, et al. Bendamustine plus rituximab versus fludarabine plus rituximab for patients with relapsed indolent and mantle-cell lymphomas: a multicentre, randomised, open-label, non-inferiority phase 3 trial. *Lancet Oncol*. 2016; 17(1):57-66.
89. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004; 104(5):1258-65.
90. Federico M, Bellei M, Marcheselli L, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol*. 2009; 27(27):4555-62.
91. 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*. 2015; 33(23):2516-22.
92. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982; 5:649–65.
93. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health*. 2012; 15(6):940-7.
94. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative Effectiveness Without Head-to-Head Trials: A Method for Matching-Adjusted Indirect Comparisons Applied to Psoriasis Treatment with Adalimumab or Etanercept. *Pharmacoeconomics*. 2010; 28(10):935-45.
95. Collett D. *Modelling Survival Data in Medical Research*, Third Edition ed. London: Chapman and Hall, 2015.
96. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Stat Med*. 2016; 35(30):5642-55.
97. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – Safety update. 7 March 2019 2018. Data on File.
98. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – TEAE summary. 13 June 2019 2018. Data on file.
99. Zucca E, Rondeau S, Vanazzi A, et al. Short regimen of rituximab plus lenalidomide in follicular lymphoma patients in need of first-line therapy. *Blood*. 2019.
100. Sarkozy C, Maurer MJ, Link BK, et al. Cause of Death in Follicular Lymphoma in the First Decade of the Rituximab Era: A Pooled Analysis of French and US Cohorts. *J Clin Oncol*. 2019; 37(2):144-52.
101. Olszewski AJ, Reagan JL and Castillo JJ. Late infections and secondary malignancies after bendamustine/rituximab or RCHOP/RCVP chemotherapy for B-cell lymphomas. *Am J Hematol*. 2018; 93(1):E1-E3.
102. Moher D, Liberati A, Tetzlaff J and Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009; 6(7):e1000097.
103. Guzauskas GF, Masaquel A, Reyes C, et al. 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. *Journal of Medical Economics*. 2018; 21(10):960-7.
104. Haukaas FS, Ohna A and 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. *Applied Health Economics and Health Policy*. 2018; 16(4):569-77.
105. All Wales Medicines Strategy Group (AWMSG). AWMSG secretariat assessment report. Idelalisib (Zydelig®) 100 mg and 150 mg film-coated tablets. 2017. Available at: http://www.awmsg.org/. Accessed: 14 March 2019.
106. Guzauskas GF, Masaquel A, Thuresson PO, et al. Is obinutuzumab plus chemotherapy followed by obinutuzumab monotherapy for previously untreated follicular lymphoma patients cost-effective in the US? *Blood*. 2017; 130.
107. Scottish Medicines Consortium (SMC). Idelalisib, 100mg and 150mg film-coated tablets (Zydelig®). 2015. Available at: https://www.scottishmedicines.org.uk/. Accessed: 14 March 2019.

108. Blommestein H, Issa D, Pompen M, et al. Cost-effectiveness of rituximab as maintenance treatment for relapsed follicular lymphoma: results of a population based study. *Eur J Haematol*. 2014; 92(5):398-406.
109. Pink J, Lane S and Hughes DA. Mechanism-based approach to the economic evaluation of pharmaceuticals: Pharmacokinetic/pharmacodynamic/pharmacoeconomic analysis of rituximab for follicular lymphoma. *Pharmacoeconomics*. 2012; 30(5):413-29.
110. Vandekerckhove S, Lamotte M, Iannazzo S and Chirita O. Cost utility analysis of bortezomib in a biomarker positive subgroup of relapsed and refractory folicular lymphoma. *Value in Health*. 2012; 15(7):A427.
111. Soini EJ, Martikainen JA and 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*. 2011; 22(5):1189-97.
112. Deconinck E, Miadi-Fargier H, Pen CL and Brice P. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: Long-term economic evaluation. *Pharmacoeconomics*. 2010; 28(1):35-46.
113. Papadakis K, Follows GA, Boyer J, et al. Cost effectiveness analysis of rituximab maintenance in patients with untreated high tumour burden follicular lymphoma after response to immunochemotherapy: A UK national healthcare services perspective. *Blood*. 2010; 116(21).
114. Capote FJ, Conde E, Fernández Rañada JM, et al. Rituximab cost analysis for maintenance treatment of patients with follicular lymphoma. *Farm Hosp*. 2008; 32(1):25-34.
115. Chiattone C, Saggia MG and Santos EV. Cost-utility analysis of maintenance therapy with rituximab for follicular lymphoma in brazil according to local HTA guidelines. *Annals of Oncology*. 2008; 19(S8):viii231.
116. Hayslip JW and Simpson KN. Cost-effectiveness of extended adjuvant rituximab for US patients aged 65-70 years with follicular lymphoma in second remission. *Clin Lymphoma Myeloma*. 2008; 8(3):166-70.
117. Hornberger J, Reyes C, Lubeck D and Valente N. Economic evaluation of rituximab plus cyclophosphamide, vincristine and prednisolone for advanced follicular lymphoma. *Leukemia and Lymphoma*. 2008; 49(2):227-36.
118. Kasteng F, Erlanson M, Hagberg H, et al. Cost-effectiveness of maintenance rituximab treatment after second line therapy in patients with follicular lymphoma in Sweden. *Acta Oncologica*. 2008; 47(6):1029-36.
119. Scottish Medicines Consortium (SMC). Rituximab 10mg/ml concentrate for infusion (MabThera®). 2006. Available at: https://www.scottishmedicines.org.uk/. Accessed: 14 March 2019.
120. Scottish Medicines Consortium (SMC). Ibritumomab tiuxetan (Zevalin®). 2005. Available at: https://www.scottishmedicines.org.uk/. Accessed: 14 March 2019.
121. National Institute for Health and Care Excellence (NICE). TA137: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. 2008. Available at: https://www.nice.org.uk/guidance/ta137. Accessed: February 2019.
122. Wang H, Aas E, Smith A, et al. An economic evaluation model for follicular lymphoma (FL): Predicting treatment cost, life expectancy and quality-adjusted life year of different scenarios using UK population based observational data. *Value Health*. 2017; 20(9):A425-A6.
123. National Institute for Health and Care Excellence (NICE). ID1379: Idelalisib for treating follicular lymphoma refractory to 2 treatments. 2018. Available at:

https://www.nice.org.uk/guidance/indevelopment/gid-ta10290. Accessed: 25 April 2019.
124. Haematological Malignancy Research Network (HMRN). Clinical Management and Outcome of Follicular Lymphoma (FL) with a Focus on Relapsed/Refractory Disease. 17 July 2019. Data on File.
125. Woods B, Sideris E, Palmer S, et al. NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care: a Critical review. 2017. Available at: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf. Accessed: 28 May 2019.
126. Office for National Statistics (ONS). National Life Tables: England and Wales 2015-2017. 2018. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandandwalesreferencetables. Accessed: 10 May 2019.
127. Wild D, Walker M, Pettengell R and Lewis G. PCN62 Utility elicitation in patients with follicular lymphoma. *Value in Health*. 2006; 9(6):A294.
128. Andorsky DJ, Franson PJ, Mouro JL, et al. MAGNIFY: A randomized, phase 3b trial in patients with relapsed/refractory (R/R) non-Hodgkin lymphoma (NHL) investigating lenalidomide plus rituximab (R2) induction followed by maintenance R2 followed by lenalidomide versus R2 induction followed by rituximab (R) maintenance. *Journal of Clinical Oncology*. 2015; 33(15_suppl):TPS8606-TPS.
129. F. Hoffmann-La Roche. Summary of Product Characteristics (SPC). Mabthera® (rituximab). 2019. Available at: https://www.medicines.org.uk/emc/product/3801/smpc. Accessed: 08 April 2019.
130. NHS: South East Cancer Network. CHOP 21 +/- Rituximab for Non-Hodgkin's Lymphoma (NHL) 2011. Available at: http://www.londoncanceralliance.nhs.uk/media/36668/NHL_CHOP-R_Protocol_V1_1.pdf. Accessed: 08 April 2019.
131. NHS: South East Cancer Network. CVP plus Rituximab (R-CVP) for Lymphoma 2011. Available at: http://www.londoncanceralliance.nhs.uk/media/36671/NHL_CVP-R_Protocol_V1.0.pdf. Accessed: 08 April 2019.
132. F. Hoffmann-La Roche. Summary of Product Characteristics (SPC). Gazyvaro® (obinutuzumab). 2019. Available at: https://www.medicines.org.uk/emc/product/3279/smpc. Accessed: 08 April 2019.
133. Celgene. UK Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Draft meeting report) 13 March 2019. (Updated: 13 May 2019) Data on file.
134. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluaions Alongside Clinical Trials - Extrapolation With Patient-Level Data. 2011. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed: 10 May 2019.
135. Cheson BD, Chua N, Mayer J, et al. Overall Survival Benefit in Patients With Rituximab-Refractory Indolent Non-Hodgkin Lymphoma Who Received Obinutuzumab Plus Bendamustine Induction and Obinutuzumab Maintenance in the GADOLIN Study. *Journal of Clinical Oncology*. 2018; 36(22):2259-66.
136. 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*. 2010; 28(17):2853-8.

137. Leger C, Sabloff M, McDiarmid S, et al. Outpatient autologous hematopoietic stem cell transplantation for patients with relapsed follicular lymphoma. *Ann Hematol*. 2006; 85(10):723-9.
138. Robinson SP, Canals C, Luang JJ, et al. The outcome of reduced intensity allogeneic stem cell transplantation and autologous stem cell transplantation when performed as a first transplant strategy in relapsed follicular lymphoma: an analysis from the Lymphoma Working Party of the EBMT. *Bone Marrow Transplant*. 2013; 48(11):1409-14.
139. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16(3):199-208.
140. EuroQol Group. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument. 2015. Available at: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L_UserGuide_2015.pdf. Accessed: 13 May 2019.
141. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997; 35(11):1095-108.
142. Pereira C, Negreiro F and Silva C. PSY43 Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. *Value in Health*. 2010; 13(7):A468.
143. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes*. 2008; 6:84.
144. National Institute for Health and Care Excellence (NICE). TA306: Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma. 2014. Available at: https://www.nice.org.uk/guidance/ta306. Accessed: 25 April 2019.
145. Swinburn P, Lloyd A, Nathan P, et al. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin*. 2010; 26(5):1091-6.
146. Beusterien KM, Davies J, Leach M, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes*. 2010; 8:50.
147. Stein EM, Yang M, Guerin A, et al. Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health Qual Life Outcomes*. 2018; 16(1):193.
148. Lloyd A, Nafees B, Narewska J, et al. Health state utilities for metastatic breast cancer. *British journal of cancer*. 2006; 95(6):683-90.
149. Tolley K, Goad C, Yi Y, et al. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ*. 2013; 14(5):749-59.
150. National Institute for Health and Care Excellence (NICE). TA423: Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens (TA250). 2016. Available at: https://www.nice.org.uk/guidance/ta423. Accessed: 25 April 2019.
151. Hannouf MB, Sehgal C, Cao JQ, et al. Cost-effectiveness of adding cetuximab to platinum-based chemotherapy for first-line treatment of recurrent or metastatic head and neck cancer. *PLoS One*. 2012; 7(6):e38557.
152. Doyle S, Lloyd A and Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer*. 2008; 62(3):374-80.
153. Ara R and Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health*. 2010; 13(5):509-18.

154. MIMS. Revlimid. 2019. Available at: https://www.mims.co.uk/drugs/cancer/antineoplastics/revlimid. Accessed: 25 April 2019.
155. MIMS. MabThera. 2019. Available at: https://www.mims.co.uk/drugs/cancer/antineoplastics/mabthera. Accessed: 27 February 2019.
156. MIMS. Rixathon. 2019. Available at: https://www.mims.co.uk/drugs/cancer/antineoplastics/rixathon. Accessed: 27 February 2019.
157. GOV.UK. Drugs and pharmaceutical electronic market information tool (eMIT). 2019. Available at: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. Accessed: 25 April 2019.
158. MIMS. Gazyvaro. 2019. Available at: https://www.mims.co.uk/drugs/cancer/antineoplastics/gazyvaro. Accessed: 27 February 2019.
159. National Institute for Health and Care Excellence (NICE). ID474: Multiple myeloma (newly diagnosed) - lenalidomide. 2019. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-tag429. Accessed: 28 May 2019.
160. F. Hoffmann-La Roche. Summary of Product Characteristics (SPC). Mabthera® (rituximab). 2019. Available at: https://www.medicines.org.uk/emc/product/5333/smpc. Accessed: 26 April 2019.
161. Hatswell AJ, Porter J, Lee D, et al. The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. *Value Health*. 2016; 19(8):1055-8.
162. NHS England. National dose banding tables. 2019. Available at: https://www.england.nhs.uk/commissioning/spec-services/npc-crg/group-b/b02/. Accessed: 27 February 2019.
163. MIMS. Allopurinol. 2019. Available at: https://www.mims.co.uk/drugs/musculoskeletal-disorders/gout/allopurinol. Accessed: 31/05/2019.
164. Accord Healthcare Limited. Summary of Product Characteristics (SPC). Allopurinol 100mg Tablets. 2019. Available at: https://www.medicines.org.uk/emc/product/6007/smpc. Accessed: 31/05/2015.
165. MIMS. Neupogen. 2019. Available at: https://www.mims.co.uk/drugs/cardiovascular-system/anaemia-neutropenia/neupogen. Accessed: 31/05/2019.
166. Amgen Ltd. Summary of Product Characteristics (SPC). Neupogen 30 MU (0.3 mg/ml) solution for injection. 2019. Available at: https://www.medicines.org.uk/emc/product/3000/smpc. Accessed: 31/05/2019.
167. Papaioannou D, Rafia R, Rathbone J, et al. Rituximab for the first-line treatment of stage III-IV follicular lymphoma (review of Technology Appraisal No. 110): a systematic review and economic evaluation. *Health Technol Assess*. 2012; 16(37):1-253, iii-iv.
168. Curtis L and Burns A. PSSRU: Unit Costs of Health and Social Care. 2018. Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/. Accessed: 26 April 2018.
169. NHS Improvement. 2017/18 reference cost data. 2018. Available at: https://improvement.nhs.uk/resources/reference-costs/. Accessed: 26 April 2019.

170. National Institute for Health and Care Excellence (NICE). TA374: Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy. 2015. Available at: https://www.nice.org.uk/guidance/ta374. Accessed: 30 April 2019.
171. MIMS. Zydelig. 2019. Available at: https://www.mims.co.uk/drugs/cancer/antineoplastics/zydelig. Accessed: 01 June 2019.
172. Round J, Jones L and Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliative medicine*. 2015; 29(10):899-907.

B.5. Appendices

Appendix C: Summary of product characteristics (SmPC) and European public assessment report (EPAR)

Appendix D: Identification, selection and synthesis of clinical evidence

Appendix E: Subgroup analysis

Appendix F: Adverse reactions

Appendix G: Published cost-effectiveness studies

Appendix H: Health-related quality-of-life studies

Appendix I: Cost and healthcare resource identification, measurement and valuation

Appendix J: Clinical outcomes and disaggregated results from the model

Appendix K: Checklist of confidential information

Appendix L: Clinical guidelines for management of follicular lymphoma and marginal zone lymphoma

Appendix M: AUGMENT and MAGNIFY trial summaries

Appendix N: Further primary endpoint analyses for AUGMENT

Appendix O: AUGMENT and MAGNIFY additional endpoints

Appendix P: Health-related quality of life in AUGMENT

Appendix Q: Utility analysis

Appendix R: Trial dosing

Appendix S: Base case parameters and distributions

Appendix T: R2 vs R-CHOP/R-CVP using ITC (van Oers)

Appendix U: Parametric curves

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Clarification questions

August 2019

File name	Version	Contains confidential information	Date
ID1374 Lenalidomide ERG Clarification questions v0.2 to PM for company [noACIC]	**V0.2**	**no**	**08/08/2019**

Notes for company

Highlighting in the template

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

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

Section A: Clarification on effectiveness data

Literature searches

A1. Appendix D.1 states that update searches were conducted from September 2017 to 04 April 2019. Please clarify whether this was the case, or whether update searches were conducted from database inception.

This was incorrectly stated as an updated search from September 2017 to 04 April 2019. This was a de novo Clinical SLR conducted to replace the older Clinical SLR (with a cut off of September 2017), as some changes were made to the protocol and search strategies were made more extensive. All searches were conducted from database inception.

A2. The PRISMA flowchart in Figure 1, Appendix D provides details of the updated search results, but there are no details of the original September 2017 search results. Please provide details of the original September 2017 search results.

The PRISMA flowchart in Figure 1, Appendix D provides details of the full search, from scratch up till 04 April 2019, and not just an update from September 2017 to 04 April 2019. Therefore, an additional PRISMA flowchart is not required.

A3. Please describe the rationale for searching ClinicalTrials.gov for "studies with results", rather than all studies. Please explain why ClinicalTrials.gov was not searched for the update searches in April 2019 (Appendix D.1).

Clinicaltrials.gov was searched in the Clinical SLR (April 2019 cut off) with the intention of identifying relevant ongoing studies, and not intended to retrieve data for those records (in line with what is requested in Section 2.11 of the NICE STA guidance).[1] The search was not restricted to “studies with results” (see search strategy below). 302 trials were retrieved from the search. After screening for relevance 24 were flagged for inclusion, of which 9 had already been included in the SLR as they had been published in either a peer-reviewed article or conference abstract. This leaves 15 trials that were not identified from other sources and may be likely to publish results within the next 12 months.

Clinicaltrials.gov search

Search strategy:

Condition or disease: ("Relapsed" OR "Refractory") AND ("follicular lymphoma" OR "FL" OR "marginal zone lymphoma" OR "MZL")

Other terms: none

Study results: all studies

Status: any

Age: any

Sex: any

Accepts healthy volunteers: no

Additional criteria: last updated between 01/01/2017 and 04/04/2019

Expert search strategy:

("Relapsed" OR "Refractory") AND ("follicular lymphoma" OR "FL" OR "marginal zone lymphoma" OR "MZL") [DISEASE] AND INFLECT ("01/01/2017" : "04/04/2019") [LAST-UPDATE-POSTED]

Hits = 302

A4. Please provide full details for the searches of conference proceedings and HTA organisation websites referred to in Appendix D.1, G.1, H.1 and I.1, including the specific resources searched, the search strategies, search terms used, and results.

Please see below for details of conference proceedings searched and identified in the **Clinical SLR**

Database	Search terms	# hits
European Hematology Association (EHA): 2015-2018	• Follicular lymphoma • Marginal zone lymphoma	2018: 61+19 = 80 2017: 56+22 = 78 2016: 58+17 = 75 2015: 49+14 = 63
International Conference on Malignant Lymphomas (ICML): 2013, 2015, 2017	• Follicular lymphoma • Marginal zone lymphoma	2017: 265 + 46 = 311 2015: 138 (see details below) 2013: 190 (see details below)
American Society for Clinical Oncology (ASCO): 2015-2018	• Follicular lymphoma • Marginal zone lymphoma	2018: 540+8 = 548 2017: 476+19 = 495 2016: 419+7 = 426 2015: 464+15 = 479
American Society of Hematology (ASH) Annual Meeting: 2014- 2018	• Follicular lymphoma • Marginal zone lymphoma	2018: 262+68 = 330 2017: 245+62 = 307 2016: 233+53 = 288 2015: 183+64 = 247 2014: 183+67 = 250
European Society of Medical Oncology (ESMO): 2014-2018	• Follicular lymphoma • Marginal zone lymphoma	2018: 74+7 = 81 2017: 77+4 = 81 2016: 68+4 = 72 2015: 10+0 = 10 2014: 77+2 = 79
National Institute for Health and Care Excellence (NICE)	• Follicular lymphoma • Marginal zone lymphoma	FL: 29 MZL: 4
Canadian Agency for Drugs and Technologies in Health (CADTH)	• Follicular lymphoma • Marginal zone lymphoma	FL: 28 MZL: 13
Therapeutic Goods Administration Australia (TGA)	• Follicular lymphoma • Marginal zone lymphoma	FL: 222 MZL: 38

Details conference search
EHA

- Keyword follicular lymphoma/marginal zone lymphoma in advanced search
- Filter by event (e.g. 20th conference) and filter on abstracts in content types

ICML 2013 (Control+F search)

- Oral presentations: 40+6 = 46
- Poster presentations: 63+14 = 74
- Publications: 56+14= 70

ICML 2015 (Control+F search)
- Oral presentations: 31+3 = 34
- Poster presentations: 33+7 = 40
- Publications: 53+11= 64

ICML 2017: one abstract book, control+F search

ASCO:
- Keyword follicular lymphoma/marginal zone lymphoma in advanced search
- In advanced search, check the box ASCO annual meeting and the year of interest

ASH
- Click on the link for the correct abstract book
- In the "search this issue" search bar, type follicular lymphoma / marginal zone lymphoma

ESMO:
- Click on the link for the correct abstract book
- In the search bar, type follicular lymphoma/marginal zone lymphoma

Please see below for details of conference proceedings searched and identified in the **Economic SLRs**

Table 1: Conference search details (searched on February 2019)

Conference name and year	Website	Search terms	Included
International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual Meeting 2017, 2018	http://www.ispor.org/heor-resources/presentations-database/search	• Lymphoma • Follicular lymphoma • Marginal zone lymphoma • MZL • MALT • Brill-Symmers • Brill Symmers • Nodal Marginal • Extranodal Marginal • Splenic Marginal	0
ISPOR European Meeting – 2017, 2018	http://www.ispor.org/heor-resources/presentations-database/search		0
American Society of Hematology (ASH) – 2017, 2018	http://www.hematology.org/Annual-Meeting/Archive.aspx	• Lymphoma • Follicular lymphoma • Marginal zone lymphoma • MZL • MALT • Brill-Symmers • Brill Symmers • Nodal Marginal • Extranodal Marginal • Splenic Marginal	0
European Hematology Association (EHA) – 2017, 2018	https://ehaweb.org/congress/previous-congresses/		0
International Conference on	http://www.lymphcon.ch/icml/website/icml-abstracts-		0

Conference name and year	Website	Search terms	Included
Malignant Lymphoma (ICML) – 2015, 2017	books/icml-abstract-books-1981-2011.html	• Resource • Cost • Utility • EQ-5D, EQ 5D, EQ5D	
American Society of Clinical Oncology (ASCO) – 2017, 2018	https://meetinglibrary.asco.org/browse-meetings/		0

Table 2: HTA search details (February 2019)

HTA agencies	URL	Search terms	Included
All Wales Medicines Strategy Group (AWMSG)	http://www.awmsg.org/	• Lymphoma • Follicular lymphoma • Marginal zone lymphoma • MZL • MALT • Brill-Symmers • Brill Symmers • Nodal Marginal • Extranodal Marginal • Splenic Marginal	1
The Haute Autorité de santé (HAS)	https://www.has-sante.fr/portail/jcms/r_1455081/en/home-page?portal=r_1455081		0
Statens legemiddelverk (SLV)	https://legemiddelverket.no/English https://www.legemiddelsok.no/		0
National Institute for Health and Care Excellence (NICE)	https://www.nice.org.uk/		4
Scottish Medicines Consortium (SMC)	https://www.scottishmedicines.org.uk/		4

A5. A "targeted literature review" is referred to in the following sections: *B.1.3. Health condition and position of the technology in the treatment pathway: Aetiology, course and prognosis; B.3.4. Measurement and valuation of health effects: Adverse reactions; and B.3.11. Interpretation and conclusions of economic evidence*. Please provide full details of these targeted literature reviews.

For Section B.3.4, a targeted review was done in order to source inputs to inform the adverse event disutility values. This targeted review involved reviewing previous NICE technology appraisals in a similar disease area including TA472[2], TA137[3] and ID1379.[4] Adverse event disutility and duration values and sources were taken from these submissions, for those adverse events reported. For adverse events not reported in these submissions, the sources used for the other adverse events were reviewed or PubMed and Google Scholar searches were conducted to find alternative sources for these adverse events disutility's. Assumptions were made for durations unable to be sourced from the literature (assuming the maximum duration).

B.1.3 and B.3.11 both refer to a different targeted literature review to that mentioned above. The focus of this additional literature review was to identify clinical evidence reporting the prognosis of FL and MZL within the same studies.

In order to identify clinical studies which compared prognosis of previously treated FL and MZL, results of the original clinical SLR were first investigated. From this original clinical SLR, seven clinical studies were identified that provided information on the prognosis of previously treated FL and MZL, with a range of treatments, including those treatments that were not of specific interest to our decision problem for the submission.

Given that the original clinical SLR was conducted in September 2017, a targeted literature search in PubMed alone was also conducted to further identify relevant studies that may have been published since the SLR was conducted. Although the original SLR was later replaced with an updated SLR (in April 2019) for the NICE submission, this targeted review was conducted before then and as such, the below represents exactly what was carried out.

For this additional targeted review, PubMed was searched on 25 January 2019 using the following search string:

- "(follicular lymphoma) AND (marginal cell lymphoma) AND (relaps* OR refract*)"

The search results were restricted to articles published from 1 September 2017 onwards. This search returned 12 citations in total, but only two of these were relevant to our research question.

Therefore, a total of nine articles were retrieved from the SLR and the additional targeted literature search combined. Data on response rates, OS and PFS were extracted from the papers for FL and MZL, as available. A summary of these nine articles and extracted results is provided in Table 3.

Only two of the nine articles reported OS and/or PFS for both patients with FL and patients with MZL. The study by Andorsky and colleagues (2019) showed very similar median PFS between FL and MZL (5.7 vs. 5.5 months) and a similar PFS rate at 24 weeks (51.5% vs. 46.2%; Figure 1; Figure 2).[5] In the study by de Vos and colleagues (2018), OS was not reached for either patients with FL or those with MZL; however, on looking at the KM curves, the OS rate at 18 months was similar between the FL and MZL patients (Figure 3).[6] A similar pattern was seen for the PFS KM curve between patients with FL or MZL (Figure 4), although median PFS was not reached for FL and was 12.0 months for MZL.[6] When response rates were compared between FL and MZL across the nine studies identified, seven of the studies reported similar overall response rates between the histology groups. When all outcome results were considered, it seems reasonable to suggest that the prognosis of relapsed/refractory FL and MZL are similar, albeit using evidence from some small study populations, particularly for the MZL patient group.

Table 3. Comparing prognosis of FL and MZL patients in clinical studies of patients with relapsed/refractory FL and MZL

Citation	Patients (n)	Study type	Therapy	OS	PFS	Response
From review of Celgene's clinical SLR						
Witzig, 2009[7]	Relapsed/refractory indolent NHL (n=43) – no limit on number of prior therapies	Clinical trial	Lenalidomide	NR	NR by type	After median follow up of 4.4 months, response for FL (grade 1 or 2); and MZL: ORR: 27%; 0% CR: 9%; 0% PR: 18%; 0% SD: 32%; 67% PD: 41%; 33%
Dreyling 2017 (CHRONOS 1 Part A)[8]	Indolent lymphoma (n=20), relapsed or refractory to two or more prior lines of therapy (FL=16; MZL=3; SLL=1)	Clinical trial	Copanlisib	NR by type	NR by type	After median treatment duration of 22.7 weeks, response for FL; MZL: ORR: 40.0%; 66.7% CR: 13.3%; 0% Unconfirmed CR: 6.7%; 0% PR: 20.0%; 66.7% SD: 53.3%; 33.3% PD: 0%; 0% N/A: 6.7%; 0%
Dreyling 2017 (CHRONOS 1 Part B)[9]	Indolent B-cell non-Hodgkin lymphoma (FL, MZL, SLL and lymphoplasmacytoid/ Waldenström macroglobulinemia) – relapsed after, or refractory to, ≥2 prior lines of treatment (total n=142; FL=104; MZL=23)	Clinical trial	Copanlisib	NR	NR by type	After median treatment duration of 22 weeks, response for FL; MZL: ORR: 58.7%; 69.6% CR: 14.4%; 8.7% PR: 44.2%; 60.9%
Flinn, 2014[10]	Relapsed indolent non-Hodgkin	Clinical trial	Idelalisib	NR	NR by type	After a median treatment duration of 3.8 months, ORR was:

	lymphoma (n=64; FL=38; MZL=6), who had received ≥1 prior chemotherapy regimen and prior rituximab					FL: 45% MZL: 33%
Gopal, 2014[11]	Patients with indolent non-Hodgkin's lymphoma who had received ≥2 prior systemic therapies (rituximab and an alkylating agent), and had not had a response or had had a relapse within 6 months (n=125; FL=72; MZL=15)	Clinical trial	Idelalisib	NR by type	NR by type	After a median treatment duration of 6.6 months, ORR was: FL: 54% MZL: 47%
Kahl, 2010[12]	Rituximab-refractory, indolent B-cell lymphoma (n=100; FL=62; MZL=16)	Clinical trial	Bendamustine	NR	NR by type	After a median treatment duration of 6 cycles (18 weeks), and a median duration of follow up of 11.8 months, response rates for FL, lymph node MZL, extra-lymph node MZL were: ORR: 74%; 78%; 86% CR: 15%; 11%; 43% Unconfirmed CR: 5%; 0%; 0% PR: 55%; 67%; 43% SD: 15%; 22%; 14% PD: 10%; 0%; 0% Unknown: 2%; 0%; 0%
Rummel, 2005[13]	Patients with mantle cell or low-grade lymphomas who had have received at least one prior chemotherapy (1-3 were allowed, but prior	Single arm clinical trial	Bendamustine plus rituximab	NR by type	NR by type	After a treatment duration of 4 cycles (16 weeks), and a median duration of follow-up of 20 months, response rates for FL; MZL were: ORR: 96%; 83% CR: 71%; 67% PR: 25%; 17%

	ritux was not allowed) and were permitted to be refractory to previous treatment (n=63; FL=24; MZL=6)					
From additional PubMed search						
Andorsky, 2019[5]	Relapsed or refractory indolent non-Hodgkin lymphoma or mantle cell lymphoma (MCL) (FL=41; MZL=17)	Open label phase 2 trial	Entospletinib	NR	Median PFS: FL: 5.7 months (95% CI: 3.6–11.2 months; Figure 1) MZL: 5.5 months (95% CI 3.5–22.1; Figure 2) PFS at 24 weeks: FL: 51·5% (95% CI 32·8-67·4%); MZL: 46·2% (95% CI 18·5-70·2%)	After median treatment duration of 16.6 weeks, response rates in FL; MZL: ORR: 17.1% (95% CI 8.3–29.7%); 11.8% (95% CI 2.1–32.6%) CR: 0%; 0% SD: 51.2%; 70.6% PD: 26.8%; 11.8%
De Vos, 2018[6]	Relapsed/refractory non-Hodgkin's lymphoma (n=60; FL=32; MZL=6)	Phase 1b dose-finding study	Venetoclax, bendamustine, and rituximab	Median OS (Figure 3): FL: Not reached; MZL: Not reached	Median PFS (Figure 4): FL: not yet reached; MZL: 12.0 months (95% CI: 3.8–21.0)	After median treatment duration of 5.5 months, and median time on study of 7.4 months, response rates for FL; MZL: ORR: 75%; 100% CR: 38%; 50% PR: 38%; 50% SD: 6%; 0% PD: 9%; 0% Discontinued w/o assessment: 9%; 0%

Figure 1. Progression-free survival in patients with relapsed/refractory FL (n=41) in the Andorsky (2019) study

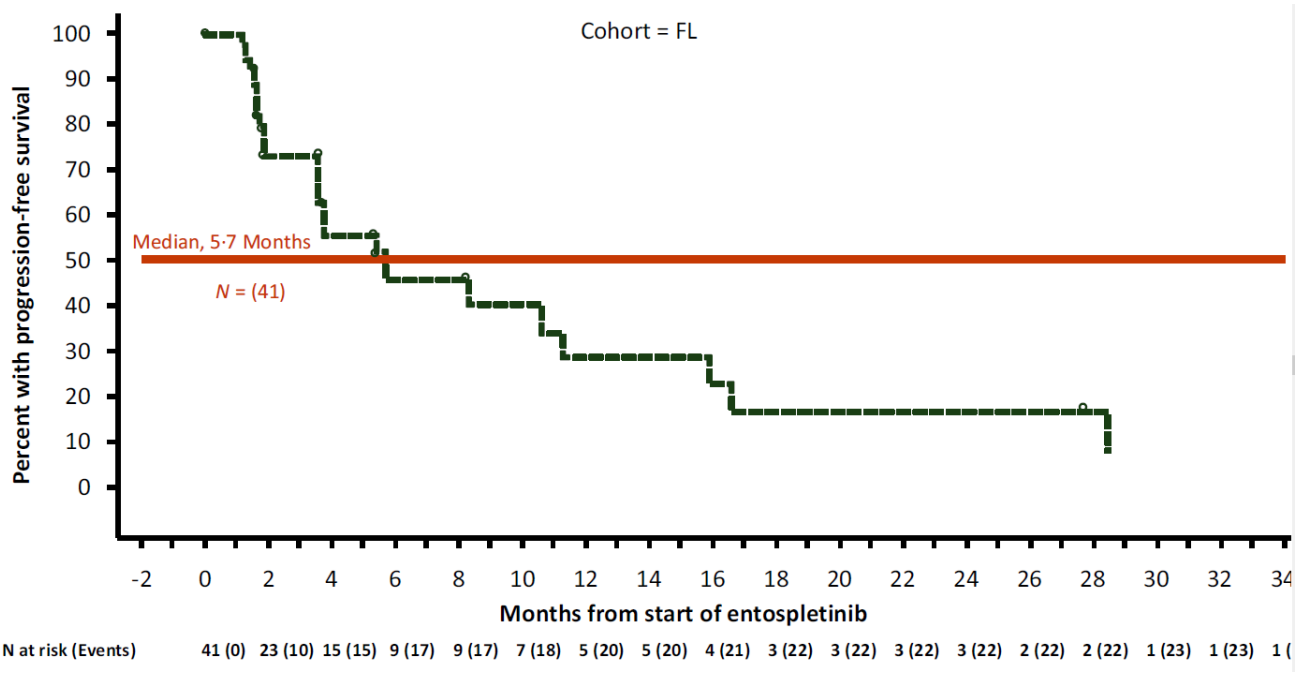


Figure 2. Progression-free survival in patients with relapsed/refractory MZL (n=17) in the Andorsky (2019) study

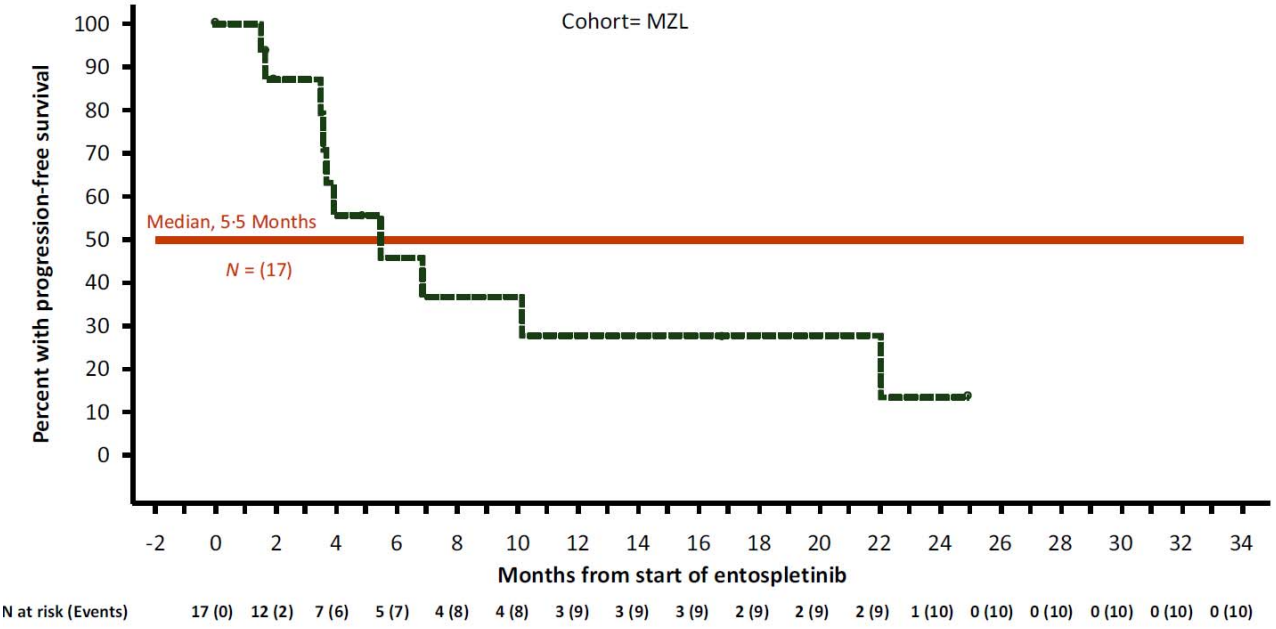


Figure 3. Overall survival for patients with relapsed/refractory indolent NHL, by histology in the de Vos (2018) study

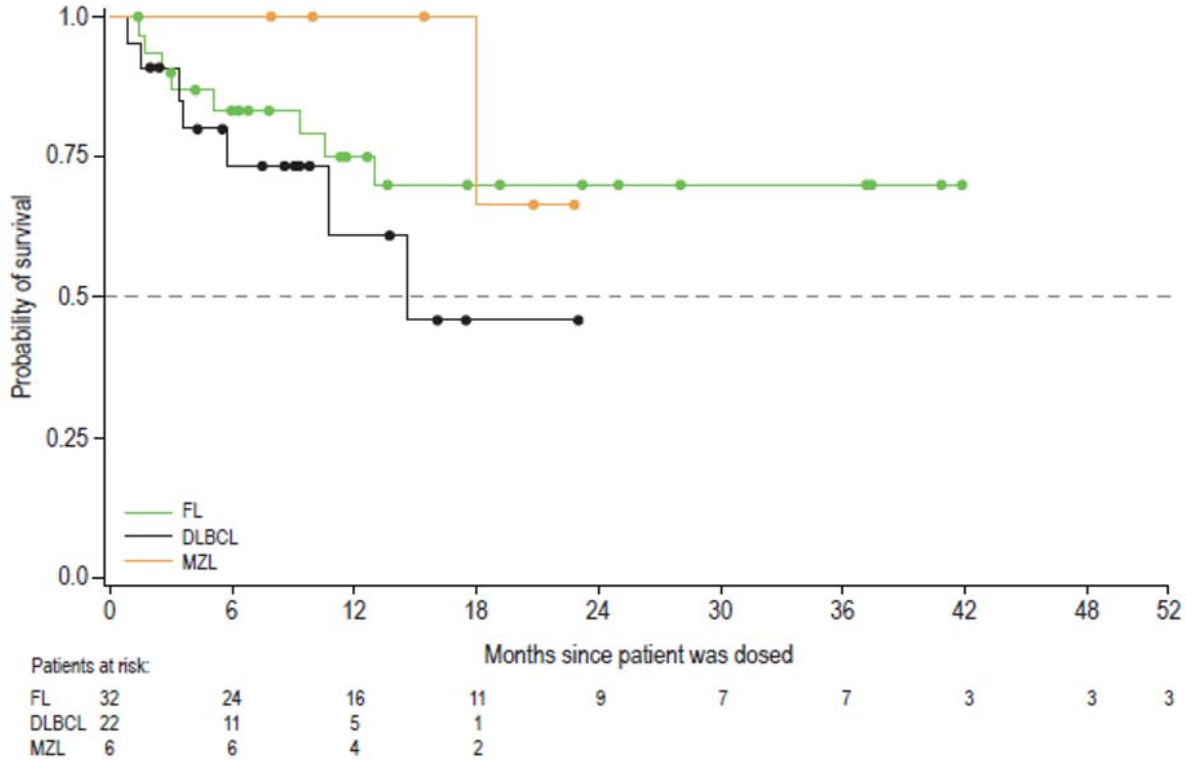
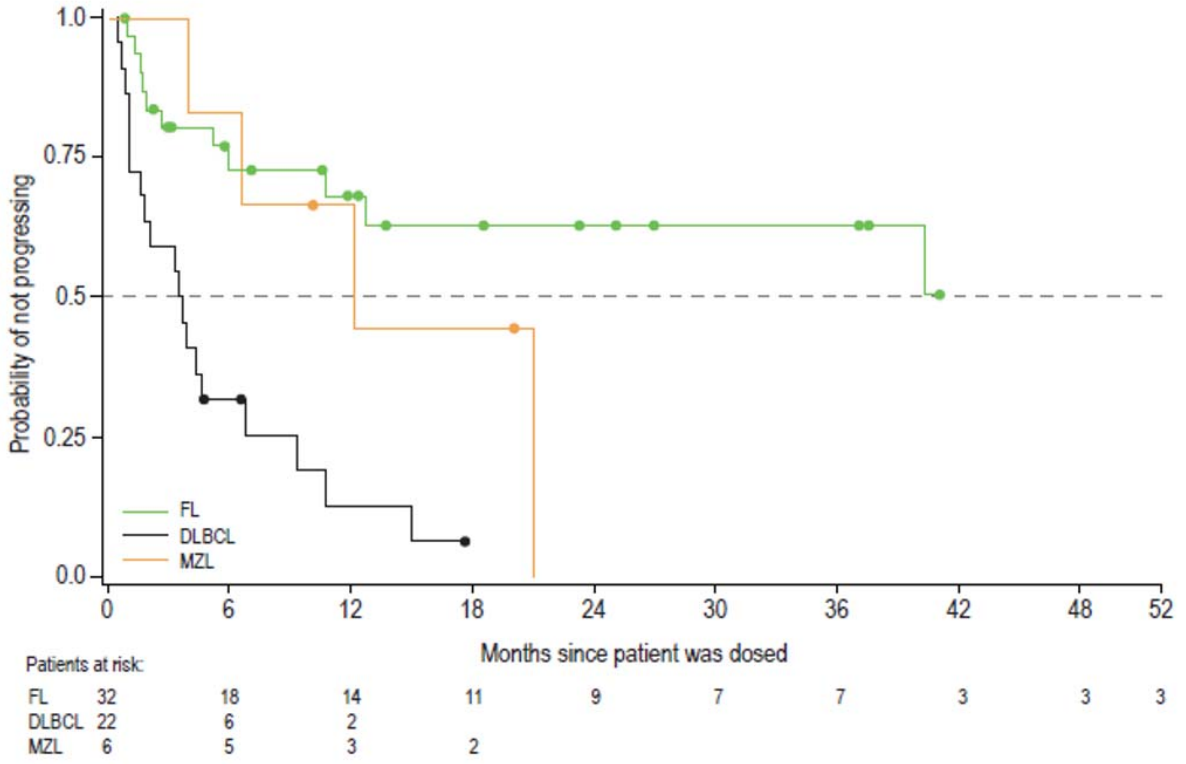


Figure 4. Progression-free survival for patients with relapsed/refractory indolent NHL, by histology in the de Vos (2018) study



Advisory Board Reports

A6. Please provide the full report for reference 31 (Document A): "Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0069) March 2019 2019. (Updated: June 2019) Data on file." Please also provide full reports for references 2 and 133 from Document B:

[redacted]

133. Celgene. UK Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Draft meeting report) 13 March 2019. (Updated: 13 May 2019) Data on file.

Please also provide full text reports for references 60, 74, 76, 77, 79, 80 and 84 in document B (and any other references for which PDFs are missing).

The advisory board report was incorrectly referenced as a draft in Document B (reference 133). We have provided the final advisory board report, which should replace all references to the draft advisory board report.

The Clinical expert clinical validation was wrongly referenced, with the intention to reference input from multiple clinical experts post the advisory board. There is no official document for these validations though details have been provided in related responses below.

Full text reports and references in Document B have been provided. Please note the MAGNIFY CSR is based on an older data cut (1 May 2017). More recent data (cut-off: 10 August 2018) is available and therefore efficacy data reported in the CSR, from the older data cut, was not used for this submission.

Comparators

A7. The decision problem as described by the company is different from the NICE scope in that it uses different comparators for R-refractory patients and non- R-refractory patients; it does not include R-monotherapy as a comparator, but it does include 'obinutuzumab in combination with bendamustine' as a comparator for the R-refractory population. This is despite an explicit statement by NICE that "obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4) [14]).

a Please clarify why the population is divided into a R-refractory population and a non-R-refractory population. Please also clarify why this would affect lenalidomide plus rituximab differently to other combinations with rituximab.

Patients who develop resistance to rituximab-containing regimens early in treatment, defined as 'best response of progressive disease (PD) or stable disease (SD) to treatment with a rituximab-containing regimen (single agent or combination) or response lasting less than 6 months following last rituximab dose' [reference per submission] and termed rituximab-refractory patients, are well documented to have poorer prognosis and limited treatment options.[15, 16]

Patients determined to be R-refractory are treated differently to the non-R-refractory population in the UK due to the availability of an alternative treatment, obinutuzumab-bendamustine, approved in the EU and recommended by NICE.[2] Accordingly, we have divided the submission into R-refractory and non-R-refractory populations to reflect the current approach to patient management in the UK.

The co-administration of lenalidomide and rituximab has demonstrated encouraging activity in rituximab-refractory patients. Chong conducted a study combining lenalidomide with rituximab in 50 patients with indolent B-cell lymphoma (60% FL; 8% small lymphocytic lymphoma [SLL] and 4% MZL) and mantle cell lymphoma (MCL; 28%) who were rituximab-refractory.[17] Patients received 2 cycles (8 weeks) of lenalidomide (10 mg) followed by 4 weekly doses of rituximab (375 mg/m^2) in combination with lenalidomide (10 mg), followed by lenalidomide alone until Week 21.

Chong demonstrated that this therapeutic combination significantly improves progression-free survival (PFS) compared with patients' prior rituximab resistance–defining regimens, including R-chemotherapy combinations, providing clinical evidence of additive activity for the combination of rituximab and lenalidomide in the rituximab-refractory population, with data also suggesting that the immunologic effects of lenalidomide may potentiate the action of rituximab by reducing regulatory T cells.

Furthermore, in primary FL samples, the combination of lenalidomide and rituximab reactivates dysfunctional NK and T cells, leading to increased cytokine production and immune synapse signalling, with enhancement of NK-mediated antibody-dependent cell cytotoxicity (ADCC) and CD8+ T cell anti-FL activity. R^2 treatment of FL patients (RELEVANCE data[18]) led to increased circulating T- and NK-cells compared to R-chemotherapy, which was associated with a decline in immune cell numbers.[19] It is anticipated that the immunologic effects of R^2 are largely driven by lenalidomide given that rituximab activity may be adversely affected by the underlying immune dysfunction observed in FL patients (e.g., low absolute lymphocyte counts and reduced numbers of circulating NK cells before treatment are predictive of a poor response). Combining rituximab with traditional chemotherapy (e.g., CHOP), does not mitigate the sustained immunosuppression seen in patients with B-cell lymphoma,[19] whereas combining rituximab with lenalidomide results in an immunotherapy regimen with complementary mechanisms including direct tumour apoptosis in FL and MZL, and immune-mediated activities, such as activation of NK cells and immune synapse formation, resulting in increased ADCC in vitro.[20] While single-agent lenalidomide and rituximab has been shown to increase formation of lytic NK cell immunological synapses with primary FL tumour cells, the R^2 combination was superior and correlated with enhanced cytotoxicity.[19]

The data discussed in the submission derived from the induction phase of the MAGNIFY study provides further evidence of the favourable activity of lenalidomide-rituximab in the rituximab-refractory population.

b. Please include all results for the comparison of lenalidomide with rituximab versus rituximab monotherapy based on the AUGMENT and MAGNIFY trials in the non-R-refractory population only (including AEs); and please provide a full economic evaluation using this comparison. NICE have identified this comparison as relevant for the current appraisal and it is based on a head-to-head comparison in the AUGMENT trial, therefore, this comparison provides the most reliable data for this appraisal.

All patients in the induction phase of MAGNIFY (as presented in the submission) received R^2 only, therefore there are no results for a comparison between R^2 and R mono from that study.

In AUGMENT, all patients were non-R-refractory, so the data that have already been presented from the AUGMENT study in the original submission dossier only relate to a non-R-refractory population. Efficacy results including IRC PFS by EMA guidance, OS, ORR, CR rate, TTNLT, EFS, TTNCT, RTNLT and HRQL for the comparison between R^2 and R mono can be found in Section B.2.6 of document B (pages 50–60), Section B.2.7 (subgroup analysis of PFS; pages 63–64), Appendix N (IRC PFS by FDA guidance; pages 226–229) and Appendix O1 (DOR, DOCR and DCRR; pages 230–232). Adverse events for both the R^2 and R mono treatment arms can be found in Appendix F1 (pages 62–64). The AUGMENT CSR and additional data tables have also been provided.

As requested by the ERG, the company has provided the results of the AUGMENT trial for R^2 in comparison to R-mono and subsequently added this into the CE model (see response to clarification question B.3). However, there is strong clinical opinion that this treatment is not the most relevant for the treatment of relapsed/refractory disease in UK clinical practice[21], due to the availability of more effective treatments (R-chemotherapy), and therefore not an appropriate comparator to aid decision making.

c. NICE have explicitly stated that obinutuzumab in combination with bendamustine is not considered a relevant comparator for disease that is refractory to rituximab. The submission currently does not present any relevant evidence for R-refractory patients. Please review the evidence and include a relevant comparator for this population.

Although we accept that O-Benda is currently only available via the CDF, and NICE's position on comparing to CDF drugs, we maintain that a comparison to O-Benda is the most relevant for the rituximab-refractory population. Advice from clinical experts from an advisory board suggests that O-Benda is the most likely comparator in patients who are refractory to rituximab.[21] Clinical experts estimate that approximately less than 20% of the relapsed/refractory population treated at second-line are R-refractory and hence candidates for treatment with O-Benda. In addition, O-Benda is the only NICE approved treatment for the rituximab-refractory population, and GADOLIN provides the only source of randomised phase III evidence for comparison in this population. A comparison of R^2 with O-Benda in the R-refractory population can be found in Section B2.9 of Document B.

Benda monotherapy, which was the relevant comparator for O-Benda in TA472, is not considered a relevant comparator in this population given that clinical experts believe that O-Benda has largely replaced use of Benda monotherapy in R-refractory patients.[21] Therefore, we do not believe providing a comparison to Benda monotherapy is relevant to decision making.

d. Please clarify why interventions containing rituximab, such as rituximab+bendamustine and R-CHOP have not been considered in the R-refractory population. If the intervention contains rituximab, please provide justification for why it should not also be possible for the comparator to contain rituximab.

As described above (question A7a), the relapsed/refractory population was split into R-refractory and non-R-refractory, based on NICE guidance and clinical expert opinion, to reflect current clinical management. As such, rituximab-containing regimens are considered in the non-rituximab refractory patients and O-Benda considered in the rituximab refractory population.

As described above (question A7a), there is rationale to believe that lenalidomide may provide a platform for potentiation of rituximab activity through immune enhancement. R-chemotherapy regimens, by contrast, are associated with immune suppression. Obinutuzumab-bendamustine is through to be a more pertinent comparator than R-chemotherapy regimens in the R-refractory population.

Included Studies

A8. Please provide the full Clinical Study Report (including all appendices, tables and figures) for the AUGMENT trial as specified in Ref 66 in document B (Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with

relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007) 2018.).

The AUGMENT CSR, along with the related data tables and figures has been provided.

A9. Please provide the number of patients that are Rituximab experienced at baseline (first line or later) in each study: AUGMENT, MAGNIFY, Van Oers 2006, and the HMRN dataset. Please also provide the number of patients in each of these studies that is R-refractory.

Of the [redacted] R-CHOP/R-CVP patients at 2L+ in the HMRN dataset used in the comparison of R^2 to R-CHOP/R-CVP, [redacted] had been treated with prior rituximab (see Appendix D3). Prior rituximab was not permitted in the Van Oers 2006 study, and 100% of the patients were rituximab-naïve. In the AUGMENT study the proportion of patients that received prior rituximab was 84.4% overall, 85.4% in the R^2 arm, and 83.3% in the R mono arm. For MAGNIFY, note that the submission incorrectly states '41% of patients were R-refractory' (old data cut) on pages 26 and 40 of Document B. In the MAGNIFY study, 96.8% of patients in the IEE population were rituximab experienced at baseline, and [redacted] of patients were R-refractory.

Table 4 summarises the proportion of patients that received rituximab and that were R-refractory in the AUGMENT, MAGNIFY and van Oers studies, as well as the HMRN dataset.

Table 4: Summary of prior rituximab treatment and R-refractory status in indirect treatment comparison datasets

Study	Treatment	% Rituximab experienced	% Rituximab-refractory
AUGMENT	R^2	85.4%	0%
	R mono	83.3%	0%
MAGNIFY (IEE pop)	R^2	[redacted]	[redacted]
Van Oers	RCHOP	0%	0%
HMRN	RCVP/RCHOP	[redacted]	[redacted]
Source: Celgene, 2018[22]; Celgene, 2018[23]; van Oers et al 2006[24]; HMRN (draft report), 2019.[25]			

A10. Please provide separate results for FL and MZL for OS, response rate, health related quality of life and adverse events from the AUGMENT trial.

Table 5 presents a summary of secondary endpoints (OS, ORR, CR rate, TTNLT, EFS, DOR, DOCR and DCRR) by FL and MZL in AUGMENT.

A summary of treatment-emergent adverse events (TEAEs) and the most common TEAEs (reported in ≥10% of patients in either treatment arm) by disease histology in AUGMENT is presented in Table 6 and Table 7, respectively.

Table 5: Summary of secondary endpoints by histology in AUGMENT: ITT population

Endpoint	FL		MZL		Overall	
	R² (N=147)	R mono (N=148)	R² (N=31)	R mono (N=32)	R² (N=178)	R mono (N=180)
Median OS, months(95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	[redacted]		[redacted]		0.61 (0.33, 1.13)[b]	
Best response, n (%)						
ORR (CR+PR)	[redacted]	[redacted]	[redacted]	[redacted]	138 (77.5)	96 (53.3)
95% CI[d]	[redacted]	[redacted]	[redacted]	[redacted]	70.7, 83.4	45.8, 60.8
p-value	[redacted]		[redacted]		<0.0001[e]	
CR rate	[redacted]	[redacted]	[redacted]	[redacted]	60 (33.7)	33 (18.3)
95% CI[d]	[redacted]	[redacted]	[redacted]	[redacted]	26.8, 41.2	13.0, 24.8
p-value	[redacted]		[redacted]		0.001[e]	
PR	[redacted]	[redacted]	[redacted]	[redacted]	78 (43.8)	63 (35.0)
SD	[redacted]	[redacted]	[redacted]	[redacted]	20 (11.2)	55 (30.6)
PD/ death	[redacted]	[redacted]	[redacted]	[redacted]	7 (3.9)	23 (12.8)
No evidence of disease	[redacted]	[redacted]	[redacted]	[redacted]	3 (1.7)	4 (2.2)
Unknown/ND/Missing	[redacted]	[redacted]	[redacted]	[redacted]	10 (5.6)	2 (1.1)
Median TTNLT, months (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	NE (NE, NE)	32.2 (23.2, NE)
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	73.6 (65.6, 80.1)	57.3 (49.3, 64.5)
Hazard ratio (95% CI)	[redacted]		[redacted]		0.54 (0.38, 0.78)[b]	
p-value	[redacted]		[redacted]		0.0007[g]	
Median EFS, months (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]	27.6 (22.1, NE)	13.9 (11.4, 16.7)
Hazard ratio (95% CI)	[redacted]		[redacted]		0.51 (0.38 to 0.67)[b]	
p-value	[redacted]		[redacted]		<0.0001[g]	

DCRR, n (%)	[redacted]					
95% CI[i]	[redacted]					
p-value	[redacted]					
	R^2 (N=118)	**R mono** (N=82)	**R^2** (N=20)	**R mono** (N=14)	**R^2** (N=138)	**R mono** (N=96)
Median DOR, months (95% CI)[a]	[redacted]				36.6 (22.9, NE)	21.7 (12.8, 27.6)
Hazard ratio (95% CI)[c]	[redacted]				0.53 (0.36 to 0.79)	
p-value[h]	[redacted]				0.0015	
	R^2 (N=51)	**R mono** (N=29)	**R^2** (N=9)	**R mono** (N=4)	**R^2** (N=60)	**R mono** (N=33)
Median DOCR, months (95% CI)[a]	[redacted]					
Hazard ratio (95% CI)[h]	[redacted]					
p-value[c]	[redacted]					

Key: CI, confidence interval; CR, complete response; DCRR, durable complete response rate, DOCR, duration of complete response; DOR, duration of response; EFS, event-free survival; FL, follicular lymphoma; IRC, Independent Review Committee; ITT, intent-to-treat; MZL, marginal zone lymphoma; ND, not done; NE, not estimable; ORR, overall response rate; OS, overall survival; PD, progressive disease; PR, partial response; R^2, lenalidomide + rituximab; R-placebo, rituximab + placebo; SD, stable disease; TTNLT, time to next antilymphoma treatment.
Notes: [a], median estimate is from Kaplan-Meier analysis; [b], from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last antilymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c], from Cox proportional hazard model; [d], exact confidence interval for binomial distribution; [e], from CMH test adjusting for the three stratification factors; [f], from Fisher-Exact test; [g], from log-rank test adjusting for the three stratification factors; [h], from log-rank test; [i], exact confidence interval for binomial distribution.
Source: Leonard, et al. 2019; Celgene, 2018.[22, 26]

Table 6: Summary of treatment-emergent adverse events in AUGMENT: Safety population

	FL		MZL		Total	
	R² (N=146)	R mono (N=148)	R² (N=30)	R mono (N=32)	R² (N=176)	R mono (N=180)
Number of patients (%)						
Any TEAE	[redacted]	[redacted]	[redacted]	[redacted]	174 (98.9)	173 (96.1)
Len/Pbo related	[redacted]	[redacted]	[redacted]	[redacted]	159 (90.3)	118 (65.6)
R related	[redacted]	[redacted]	[redacted]	[redacted]	132 (75.0)	105 (58.3)
Grade 3–4 TEAE	[redacted]	[redacted]	[redacted]	[redacted]	121 (68.8)	58 (32.2)
Len/Pbo related	[redacted]	[redacted]	[redacted]	[redacted]	101 (57.4)	38 (21.1)
R related	[redacted]	[redacted]	[redacted]	[redacted]	57 (32.4)	19 (10.6)
Grade 5 TEAE	[redacted]	[redacted]	[redacted]	[redacted]	2 (1.1)	2 (1.1)
Any SAE	[redacted]	[redacted]	[redacted]	[redacted]	45 (25.6)	25 (13.9)
Len/Pbo related	[redacted]	[redacted]	[redacted]	[redacted]	23 (13.1)	8 (4.4)
R related	[redacted]	[redacted]	[redacted]	[redacted]	13 (7.4)	3 (1.7)
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]	[redacted]	[redacted]	46 (26.1)	6 (3.3)
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]	[redacted]	[redacted]	112 (63.6)	47 (26.1)
Any TEAE leading to dose interruption of R	[redacted]	[redacted]	[redacted]	[redacted]	60 (34.1)	37 (20.6)
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]	[redacted]	[redacted]	15 (8.5)	9 (5.0)
Any TEAE leading to discontinuation of R	[redacted]	[redacted]	[redacted]	[redacted]	6 (3.4)	2 (1.1)

Key: FL, follicular lymphoma; Len, lenalidomide; MZL, marginal zone lymphoma; Pbo, placebo; R, rituximab; R², lenalidomide + rituximab; R-placebo, rituximab + placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[22]

Table 7: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT: Safety population

	FL		MZL		Total	
	R^2 (N=146)	R mono (N=148)	R^2 (N=30)	R mono (N=32)	R^2 (N=176)	R mono (N=180)
Number of patients (%)						
Blood and lymphatic system disorders	[redacted]	[redacted]	[redacted]	[redacted]	118 (67.0)	58 (32.2)
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]	102 (58.0)	40 (22.2)
Leukopenia	[redacted]	[redacted]	[redacted]	[redacted]	36 (20.5)	17 (9.4)
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]	28 (15.9)	8 (4.4)
Thrombocytopenia	[redacted]	[redacted]	[redacted]	[redacted]	26 (14.8)	8 (4.4)
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	115 (65.3)	88 (48.9)
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	55 (31.3)	41 (22.8)
Constipation	[redacted]	[redacted]	[redacted]	[redacted]	46 (26.1)	25 (13.9)
Abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]	22 (12.5)	16 (8.9)
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	20 (11.4)	23 (12.8)
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]	110 (62.5)	88 (48.9)
URTI	[redacted]	[redacted]	[redacted]	[redacted]	32 (18.2)	23 (12.8)
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	13 (7.4)	18 (10.0)
General disorders and administration site conditions	[redacted]	[redacted]	[redacted]	[redacted]	98 (55.7)	89 (49.4)
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	38 (21.6)	33 (18.3)
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]	37 (21.0)	27 (15.0)
Asthenia	[redacted]	[redacted]	[redacted]	[redacted]	24 (13.6)	19 (10.6)
Oedema peripheral	[redacted]	[redacted]	[redacted]	[redacted]	23 (13.1)	16 (8.9)
Skin and subcutaneous tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	89 (50.6)	43 (23.9)
Pruritus	[redacted]	[redacted]	[redacted]	[redacted]	21 (11.9)	7 (3.9)
Rash	[redacted]	[redacted]	[redacted]	[redacted]	19 (10.8)	7 (3.9)
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	73 (41.5)	58 (32.2)
Muscle spasms	[redacted]	[redacted]	[redacted]	[redacted]	23 (13.1)	9 (5.0)
Back pain	[redacted]	[redacted]	[redacted]	[redacted]	14 (8.0)	18 (10.0)
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	73 (41.5)	65 (36.1)
Cough	[redacted]	[redacted]	[redacted]	[redacted]	40 (22.7)	31 (17.2)
Dyspnoea	[redacted]	[redacted]	[redacted]	[redacted]	19 (10.8)	8 (4.4)
Investigations	[redacted]	[redacted]	[redacted]	[redacted]	60 (34.1)	50 (27.8)
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]	18 (10.2)	15 (8.3)

	FL		MZL		Total	
	R² (N=146)	R mono (N=148)	R² (N=30)	R mono (N=32)	R² (N=176)	R mono (N=180)
Metabolism and nutrition disorders	[redacted]	[redacted]	[redacted]	[redacted]	58 (33.0)	40 (22.2)
Decreased appetite	[redacted]	[redacted]	[redacted]	[redacted]	23 (13.1)	11 (6.1)
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]	58 (33.0)	39 (21.7)
Headache	[redacted]	[redacted]	[redacted]	[redacted]	26 (14.8)	17 (9.4)
Injury, poisoning and procedural complications	[redacted]	[redacted]	[redacted]	[redacted]	42 (23.9)	40 (22.2)
Infusion related reaction	[redacted]	[redacted]	[redacted]	[redacted]	26 (14.8)	24 (13.3)
Eye disorders	[redacted]	[redacted]	[redacted]	[redacted]	28 (15.9)	14 (7.8)
Neoplasms benign, malignant and unspecified (inc. cysts and polyps)	[redacted]	[redacted]	[redacted]	[redacted]	26 (14.8)	9 (5.0)
Tumour flare	[redacted]	[redacted]	[redacted]	[redacted]	19 (10.8)	1 (0.6)
Psychiatric disorders	[redacted]	[redacted]	[redacted]	[redacted]	24 (13.6)	20 (11.1)
Cardiac disorders	[redacted]	[redacted]	[redacted]	[redacted]	21 (11.9)	17 (9.4)
Vascular disorders	[redacted]	[redacted]	[redacted]	[redacted]	21 (11.9)	22 (12.2)

Key: FL, follicular lymphoma; MZL, marginal zone lymphoma; R², lenalidomide + rituximab; R-placebo, rituximab + placebo; URTI, upper respiratory tract infection.
Source: Celgene, 2018.[22]

A11. Please explain why the following studies were not included in the indirect comparisons: CALGB-50401 for rituximab+lenalidomide and Rummel 2016 for rituximab+bendamustine.

CALGB-50401 is a Phase II study, and due to the availability of more relevant Phase III studies (AUGMENT and MAGNIFY), with readily available IPD, it was not thought to be relevant to the MAIC. In addition, the ITC stated that for each treatment, the studies with the largest number of patients would be used, therefore, AUGMENT/MAGNIFY would be used in favour of CALGB-50401 (n=66) in the ITC.

The Rummel 2016 study was not included in the indirect comparison as rituximab+bendamustine was not thought to be a relevant comparator in the relapsed/refractory setting due to feedback from clinical experts of its predominant use in first-line in UK clinical practice.[21]

A12. Please explain why there are differences in inclusion criteria for the MAGNIFY study in the company submission (CS, Table 6, page 38) and clinicaltrials.gov (https://clinicaltrials.gov/ct2/show/NCT01996865). For instance, FL Grade 3b is

included according to the CS, but not according to clinicaltrials.gov, and ECOG-PS < 2 in the CS versus ECOG-PS ≤ 2 on clinicaltrials.gov.

It would appear that the clinicaltrials.gov entry for the MAGNIFY study has not been updated to include the addition of Grade 3b and transformed follicular lymphoma provided for in a protocol amendment. The company submission reference to ECOG < 2 is incorrect; per clinicaltrials.gov and the CSR, the inclusion criteria required ECOG ≤ 2.

Indirect comparisons

A13. Priority Question: For all MAIC analyses, please provide full details of the statistical methods, including the type of statistical model, the rationale for variable selection, the weighting applied and the statistical software packages used, in sufficient detail to enable replication by an independent statistician. Please also provide all the relevant datasets (both IPD and summary statistics) and analysis code to enable the ERG to replicate and check the analyses.

For all MAIC analyses, the matching variables were identified as follows:

- A list of potential effect modifiers and/or prognostic factors were identified and validated by external clinical experts.
- Availability of the identified potential effect modifiers and/or prognostic factors was summarised for each comparator evidence source.
- All available potential effect modifiers and/or prognostic factors were used in the matching unless the adjustment resulted in an effective sample size (ESS) that was too small for analysis.
    - In the case where the ESS was too small, the list of variables used for adjustment was reduced before analysis. This was done to maintain the maximum number of the most clinically important variables in the adjustment. Several combinations of variables were explored. However, note that excluding known imbalanced covariates from matching may result in populations with differing levels of effect modifiers/prognostic variables on each treatment, which can bias the analysis results.

Each comparison is presented here with the variables that were used in the matching:

- R^2 (AUGMENT) versus R-CVP/R-CHOP (HMRN) - non-rituximab refractory population
    - Age (≥60, <60)
    - Prior therapies (1, 2, ≥3)

- Prior rituximab (yes, no)
- POD24 (yes, no)
- Ann Arbor Stage (1/2, 3/4)
  - This variable in the HMRN data was only available for those patients who are fully staged and therefore missing information was assumed to be equally distributed across the two categories (1/2 and 3/4)
- Number of nodal sites (≤4, >4)
  - This variable in the HMRN data was only available for those patients who are fully staged and therefore missing information was assumed to be equally distributed across the two categories (≤4 and >4)
- Bone marrow involvement (yes, no/missing)
  - This variable has been added into the matching based on the response provided to question A17.
  - This variable in the HMRN data was only available for those patients who are fully staged and there was a similar percentage of missing data in the AUGMENT study - missing data has therefore been grouped with no bone marrow involvement.

R^2 (AUGMENT) versus R-CHOP (Van Oers, 2006) - non-rituximab refractory population, FL+MZL histology (for OS and PFS)

- Age (mean)
  - The Van Oers publication only reported the median and therefore it has been assumed that age for R-CHOP patients follows a symmetric distribution and the median has been imputed for the mean
- Prior therapies (1, 2, ≥3)
- Ann Arbor Stage (1/2, 3/4)
- Refractory to last prior therapy (yes, no)
- FLIPI (low, medium, high)

R^2 (MAGNIFY) versus O-benda (GADOLIN [Sehn, 2016/Cheson, 2018]) – rituximab refractory population, FL histology (for OS and PFS)

- Age (mean)
    - The Sehn, 2016/Cheson, 2018 publications only reported the median and therefore it has been assumed that age for O-benda patients follows a symmetric distribution and the median has been imputed for the mean
- Prior therapies (1, 2, ≥3)
- Prior rituximab (yes, no)
- Refractory to last prior therapy (yes, no)
- Bone marrow involvement (yes, no)

MAIC weights for the individual R² subjects (using the patient-level data from AUGMENT or MAGNIFY - depending on the comparison) were calculated based on the MAIC code provided in the NICE DSU TSD 18 appendix in R and using the summary baseline characteristics reported in the comparator evidence as listed above.

Individual patient level data (IPD) for the comparators was not available for the identified comparative evidence sources. For the published data, KM graphs were digitised using Web-based application WebPlotDigitizer (https://automeris.io/WebPlotDigitizer/) for the digitisation of the KM curves and to create pseudo-IPD using the algorithm of Guyot 2012.[27] HMRN provided time and event data in an Excel spreadsheet.

IPD from either MAGNIFY or AUGMENT depending on the comparison, including the relevant matching weights, was then combined with the comparator pseudo-IPD – weights of 1 were assigned to the comparator data.

Cox proportional hazards models were fitted in R using the weighted data to generate hazard ratios. To account for the fact that the weights are estimated rather than fixed and known, confidence intervals were calculated using a bootstrap estimator.[28] The use of a bootstrap estimator is intuitively appealing; weights are estimated and subject to sampling uncertainty, and bootstrapping can quantify this. Bootstrapping was performed using the following algorithm:

1. R²-treated patients will be sampled with replacement (a bootstrap dataset)

2. For each bootstrap dataset, a set of weights will be derived

3. For each bootstrap dataset and corresponding set of weights, an estimate (in this case a HR, will be obtained.

This procedure was repeated 10,000 times to obtain a distribution of estimates for which the 2.5th and 97.5th percentile was used to generate the limits of a confidence interval. For the analyses using published data, confidence intervals were also calculated using robust sandwich estimators of the variance as discussed in NICE DSU TSD 18.[29]

Finally, parametric survival models were fitted to the weighted R^2 data and the comparator data as per the NICE DSU TSD14, which includes fitting the six standard models (Exponential, Weibull, log-normal, log-logistic, Gompertz and generalised Gamma).

The ADaM datasets needed to perform the analyses are:

- ADTTE (for overall survival, MAGNIFY progression-free survival, time on treatment and time to next antilymphoma treatment data)
- ADTTTEIRC (for AUGMENT progression-free survival data)
- ADSL_SUP (for information on the number of nodal sites for each subject)
- ADSL (for remaining variables used in matching)

All dataset creation, matching weight calculations, and survival modelling were performed in R.

Approval processes are required in order to share IPD with the ERG. Necessary documentation required to start approval processes has been shared with the ERG, and code and IPD required to replicate the MAIC will be provided upon completion of the approval process. The IPD may only be used to replicate results of the MAIC analyses and all results of the MAIC should be treated as "academic in confidence."

A14. The company assume similar efficacy in a relapsed/refractory setting for R-CHOP and R-CVP (CS, page 84).

a. Please provide clinical effectiveness data from the HMRN dataset for R-CHOP and R-CVP.

Data for R-CHOP and R-CVP have been pooled given clinical feedback that it is not unreasonable to assume similar efficacy between R-CHOP and R-CVP in the relapsed/refractory setting, and HMRN clinical data supporting this.

KM plots of OS, PFS and TTNLT for R-CHOP and R-CVP are presented in Figure 5,

Figure **6** and Figure **7**, respectively. Cox PH Model outputs for OS, PFS and TTNLT with the pooled R-CHOP R-CVP data when including treatment, age, prior lines of therapy, early relapse, stage, nodal sites and prior rituximab as covariates are presented in Table 8, Table 9, and **Table 10**, respectively.

Figure 5: Kaplan-Meier plot of overall survival for R-CHOP and R-CVP

Key: OS, overall survival; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Table 8: Cox PH Model outputs for overall survival with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
Prior lines of therapy	[redacted]	[redacted]	[redacted]
Early relapse from diagnosis	[redacted]	[redacted]	[redacted]
Stage	[redacted]	[redacted]	[redacted]
Nodal sites	[redacted]	[redacted]	[redacted]

Prior rituximab	[redacted]	[redacted]	[redacted]
Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.			

Figure 6: Kaplan-Meier plot of progression-free survival for R-CHOP and R-CVP

Key: PFS, progression-free survival; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Table 9: Cox PH Model outputs for progression-free survival with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
Prior lines of therapy	[redacted]	[redacted]	[redacted]
Early relapse from diagnosis	[redacted]	[redacted]	[redacted]
Stage	[redacted]	[redacted]	[redacted]

Nodal sites	[redacted]	[redacted]	[redacted]
Prior rituximab	[redacted]	[redacted]	[redacted]
Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.			

Figure 7: Kaplan-Meier plot of time to next anti-lymphoma treatment for R-CHOP and R-CVP

Key: R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; TTNLT, time-to-next antilymphoma therapy.

Table 10: Cox PH Model outputs for time to next antilymphoma treatment or death with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
Prior lines of therapy	[redacted]	[redacted]	[redacted]

Early relapse from diagnosis	[redacted]	[redacted]	[redacted]
Stage	[redacted]	[redacted]	[redacted]
Nodal sites	[redacted]	[redacted]	[redacted]
Prior rituximab	[redacted]	[redacted]	[redacted]
Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.			

b. Please provide the “empirical data demonstrating this to be the case” as stated in the CS on page 84.

The empirical data demonstrating that R-CHOP and R-CVP have similar efficacy was included in Appendix D3 (the reference to the Appendix was missing from the CS on page 84).

Appendix D.3 in the original company submission provides details of the data used to justify pooling of R-CHOP and R-CVP. The KM plots of OS, PFS and TTNLT demonstrate how similar the end points are for these two treatments and overlap at various points (see Figures 2 – 4 in Appendix D.3, and Figure 5 and Figure 7 above). The Cox proportional hazards models also show no evidence of a significant treatment effect after adjusting for other covariates (age, prior lines of therapy, early relapse, stage, nodal sites and prior rituximab). The p-values from the Cox proportional hazard models for OS, PFS and TTNLT for treatment effect between R-CHOP and R-CVP are 0.617, 0.535 and 0.511, respectively (see Tables 25–27 in Appendix D.3).

Additionally, there is clinical support to suggest that R-CHOP and R-CVP would have similar efficacy to patients in the relapsed/refractory setting.

A15. Please provide the full HMRN report for the MAIC based on HMRN data. Please provide all effectiveness data (adjusted and unadjusted) for R-CVP and R-CHOP (separately by treatment and pooled) based on the HMRN data.

The HMRN report, containing effectiveness data for R-CVP and R-CHOP (separately by treatment and pooled) has been provided. Adjusted and unadjusted data for R-CVP/CHOP pooled has been provided in Section B.2.9 in the original Document B. MAIC analysis for R-CHOP and R-CVP has not been conducted due to evidence supporting the pooled approach described in response to Clarification Question A14.

To note: The HMRN report is a draft version, specifically produced for use in the NICE submission. A final version is not currently available and will depend on publication plans.

A16. Please clarify whether definitions of covariates used in the HMRN-MAIC were the same in the HMRN database as in the AUGMENT trial. If definitions were not the same, please do not include these variables in the matching.

The definitions of variables that were collected and considered to be adjusted for in the MAIC analyses, from both the AUGMENT study and the HMRN dataset, are summarised in Table **11**. The only variable that was identified as having a disparity in the definition was bulky disease. This variable was not included in the matching for this reason.

Table 11: Definitions of variables in AUGMENT and HMRN

Variable	AUGMENT definition	HMRN definition
Age	NA	NA
Sex	NA	NA
Number of prior antilymphoma regimens	NA	NA
Prior rituximab treatment	NA	NA
POD24	Progression within 24 months of diagnosis	Progression within 24 months of diagnosis
Bone marrow involved	Bone marrow involvement by lymphoma as demonstrated by recent bone marrow biopsy	Bone marrow involvement by lymphoma as demonstrated by recent bone marrow biopsy
Nodal sites	NA	NA
Bulky disease	At least one lesion that is ≥ 7 cm or at least 3 lesions with 3 cm or larger in the longest diameter by investigator review	At least one lesion that is ≥ 10 cm
Ann Arbour Stage	NA	NA

Key: BMI, bone marrow involvement; CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival; PFS, progression-free survival; R^2, rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone TTNLT, time to next antilymphoma treatment.

A17. In Appendix D3, an overview is presented for variables that were common across AUGMENT and the HMRN dataset and the variables included in the

matching. ‘Sex’, ‘Bone marrow involved’, and ‘Bulky disease’ are not included in the matching (see Table 29, Appendix D of the CS). Also, the populations in AUGMENT and the HMRN dataset are quite different with regards to ‘Bone marrow involved’, and ‘Bulky disease’.

a. Please explain why ‘sex’, ‘Bone marrow involved’, and ‘Bulky disease’ are not included in the matching.

Sex was not identified as being a potential prognostic factor and/or treatment effect modifier in the list of variables that was validated by external clinical experts and was therefore not included as a matching variable. The definition of bulky disease differed in the AUGMENT and HMRN data, as specified in question A16, and therefore bulky disease was not included as a matching variable.

It is agreed that “bone marrow involved” should have been considered as a matching variable given that it was identified as being a potential prognostic factor and/or treatment effect modifier.

b. Please provide MAIC results based on different sets of matching variables (especially those including ‘Bone marrow involved’ and ‘Bulky disease’).

The comparison to R-CVP/R-CHOP has been performed with additional adjustment for bone marrow involvement. The addition of this extra variable has had little impact on the results, as can be seen in the updated Table 20 from Section B.2.9 of the CS below.

Updated Table 20 from Section B.2.9 of the CS

Outcome	R^2, adjusted N	R-CVP/R-CHOP N	HR (95% CI)[a] (without BMI adjustment)	HR (95% CI) (with the addition of bone marrow involved adjustment)
OS	[redacted]	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]	[redacted]
TTNLT/death	[redacted]	[redacted]	[redacted]	[redacted]

Outcome	R^2, adjusted N	R-CVP/R-CHOP N	HR (95% CI)[a] (without BMI adjustment)	HR (95% CI) (with the addition of bone marrow involved adjustment)

Key: CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival; PFS, progression-free survival; R^2, rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone TTNLT, time to next antilymphoma treatment.

Notes: [a], bootstrapped CI.

A18. Because R-refractory and non-R-refractory patients are presented as separate analyses in the CS, please present all results (including AEs) from the MAGNIFY study separately for R-refractory and non-R-refractory patients. Please also provide the data for FL and MZL (as requested in question A10) separately for R-refractory and non-R-refractory patients.

A summary of the efficacy results (response rate, PFS based on EMA guidance, DOR, DOCR and TTR) split by R-refractory status for patients in the induction safety population in MAGNIFY are presented in Table 12, Table 13, Table 14, Table 15 and Table 16, respectively.

A summary of the efficacy results (response rate, PFS based on EMA guidance, DOR, DOCR and TTR) split by R-refractory and non-R-refractory for FL patients in the induction safety population in MAGNIFY are presented in Table 17, Table 19, Table 21, Table 23 and Table 25, respectively.

A summary of the efficacy results (response rate, PFS based on EMA guidance, DOR, DOCR and TTR) split by R-refractory and non-R-refractory for MZL patients in the induction safety population in MAGNIFY are presented in Table 18, Table 20, Table 22, Table 24 and Table 26, respectively.

A summary of TEAEs and most common TEAEs (≥10% of patients) in MAGNIFY split by R-refractory status are presented in Table 27 and Table 28, respectively. A summary of TEAEs and most common TEAEs (≥10% of patients) for FL patients in MAGNIFY split by R-refractory status are presented in Table 29 and Table 31, respectively. A summary of TEAEs and most common TEAEs (≥10% of patients) for MZL patients in MAGNIFY split by R-refractory status are presented in Table 30 and Table 32, respectively.

Table 12: Summary of response rate by best response by R-refractory status in MAGNIFY: IEE population

Best Response in Induction Period	Total	
	Rituximab-refractory:Yes (N=113)	Rituximab-refractory:No (N=197)
Number of Subjects	[redacted]	
Overall Response Rate (CR+CRu+PR) – n (%)	71 (62.8)	154 (78.2)
95% CI [a]	[redacted]	
Complete Response Rate (CR+CRu) – n (%)	45 (39.8)	93 (47.2)
95% CI [a]	[redacted]	

Percentage is based on the total number of subjects in the IEE population.
[a] 95% CI based on the Clopper-Pearson exact method.

Table 13: Summary of progression-free survival based on EMA guidance by R-refractory status in MAGNIFY: Induction safety population

Table 14:Summary of duration of response by R-refractory status in MAGNIFY: IITT population

Parameter	Rituximab-refractory:Yes (N=71)	Rituximab-refractory:No (N=155)	Total Responders (N=226)
Duration of Response [a]			
[redacted]			
Median Duration of Response (months) [b]	35.8	NE	36.8
95% CI for Median Duration of Response	(19.2, NE)	(36.8, NE)	(35.8, NE)
[redacted]			

The analysis is only performed for subjects who have achieved PR or better after the first dose date of induction therapy and prior to any Maintenance Period treatment and any subsequent anti-lymphoma therapy in Induction Period.
Percentage is based on the total number of responders.
[a] Duration of response is defined as the time (months) from the initial response (at least PR) to documented disease progression or death, whichever occurs first.
[b] Statistics obtained from Kaplan-Meier method. 95% CI is based on Greenwood formula.

Table 15: Summary of duration of complete response by R-refractory status in MAGNIFY: IITT population

Table 16: Summary of time to response by R-refractory status in MAGNIFY: IITT population

Parameter	Rituximab-refractory:Yes (N=71)	Rituximab-refractory:No (N=155)	Total Responders (N=226)
Time to Response (months) [a]			
[redacted]			
Median	2.8	2.7	2.7
[redacted]			
Min, Max	1.6, 12.0	1.6, 11.6	1.6, 12.0

The analysis is only performed for subjects who have achieved PR or better after the first dose date of induction therapy and prior to any Maintenance Period treatment and any subsequent anti-lymphoma therapy in Induction Period.
[a] Time to response is defined as the time (months) from the first dose date of induction therapy to initial response (at least PR) in the Induction Period.

Table 17: Summary of response rate by best response split by R-refractory status in MAGNIFY: IEE population (FL patients only)

Table 18: Summary of response rate by best response split by R-refractory status in MAGNIFY: IEE population (MZL patients only)

Table 19: Summary of progression-free survival based on EMA guidance by R-refractory status in MAGNIFY: Induction safety population (FL patients only)

Table 20: Summary of progression-free survival based on EMA guidance by R-refractory status in MAGNIFY: Induction safety population (MZL patients only)

Table 21: Summary of duration of response split by R-refractory status in MAGNIFY: IITT population (FL patients only)

Table 22: Summary of duration of response split by R-refractory status in MAGNIFY: IITT population (MZL patients only)

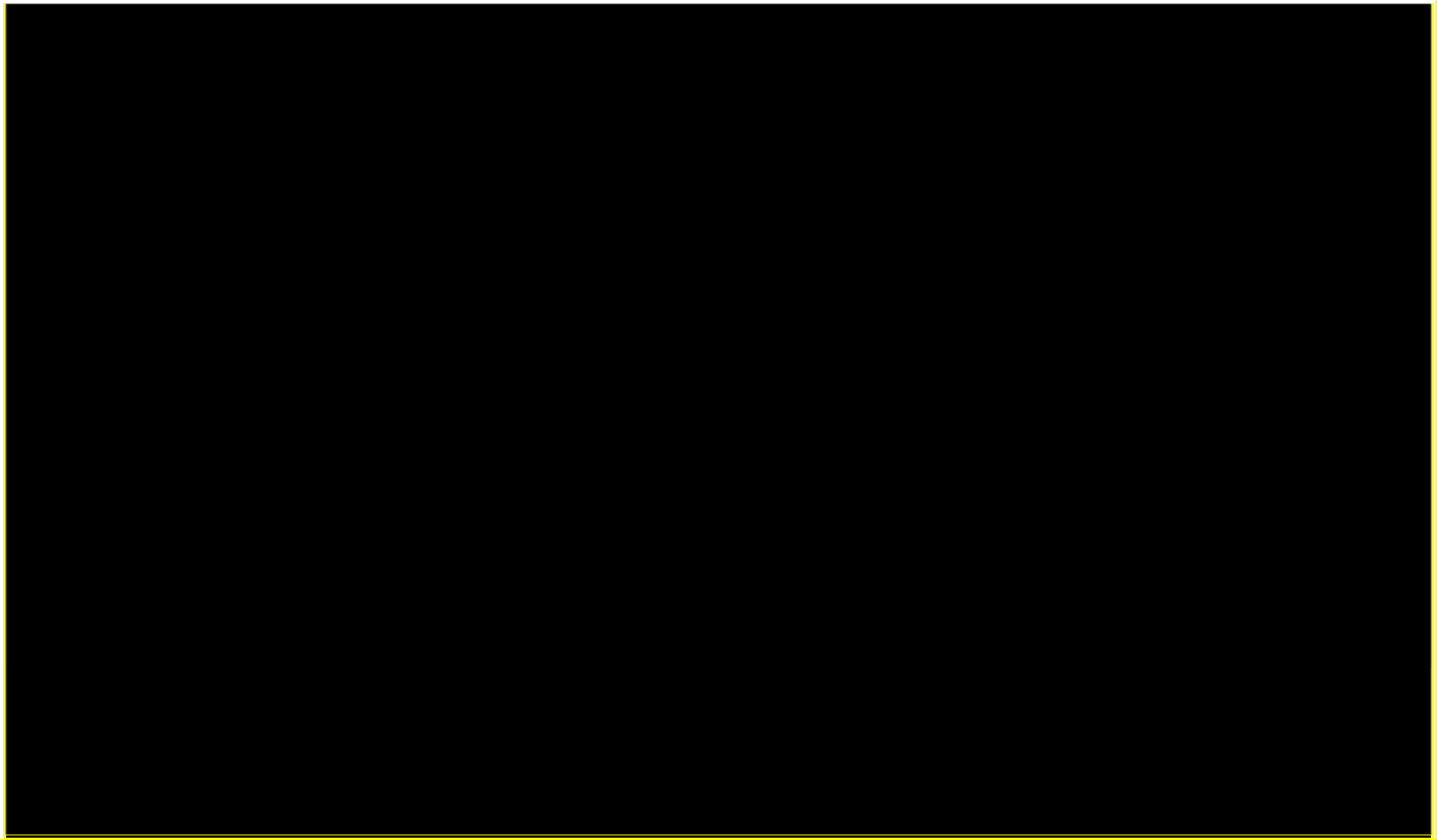


Table 23: Summary of duration of complete response by R-refractory status in MAGNIFY: IITT population (FL patients only)

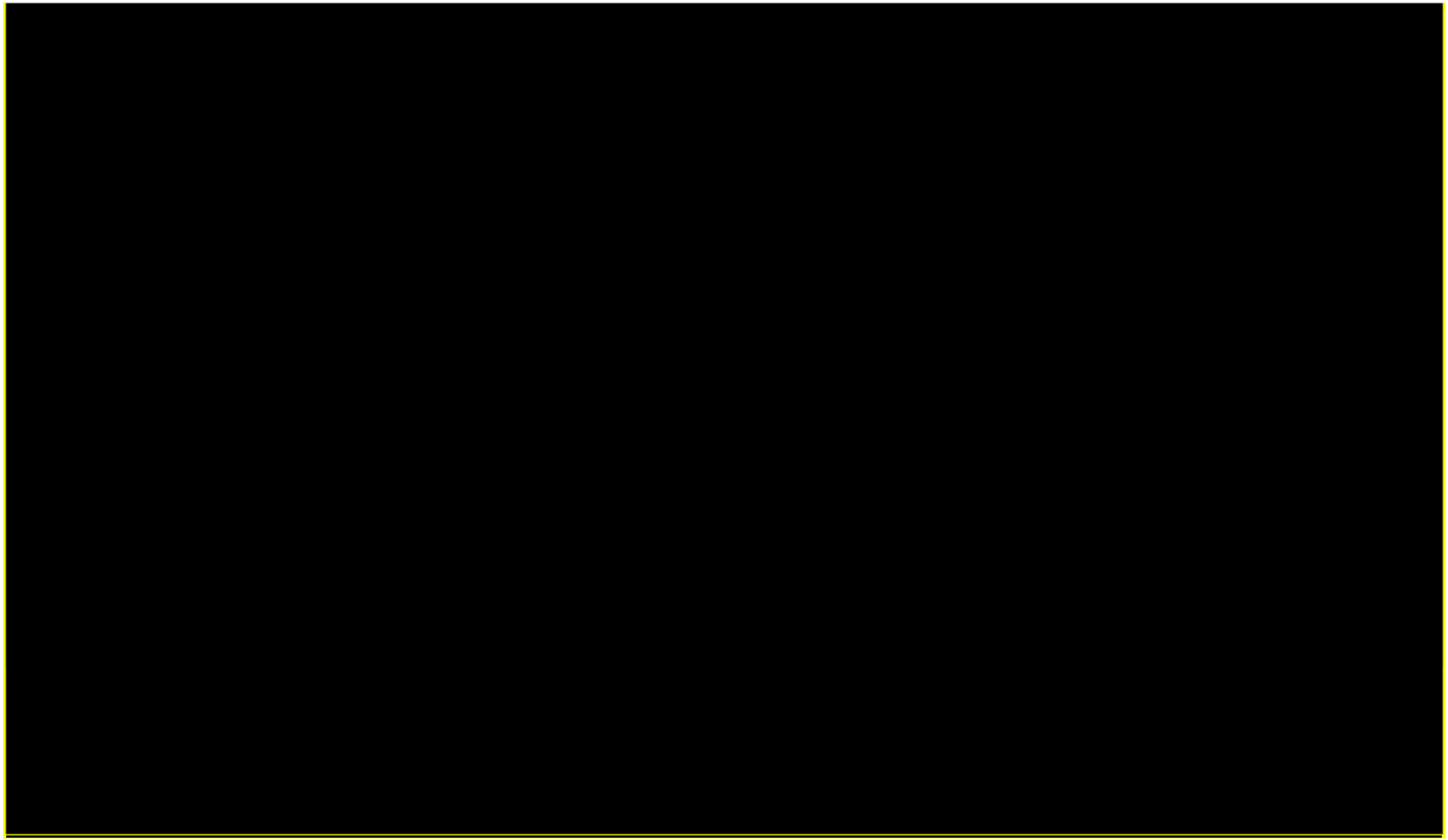


Table 24: Summary of duration of complete response by R-refractory status in MAGNIFY: IITT population (MZL patients only)

Table 25: Summary of time to response by R-refractory status in MAGNIFY: IITT population (FL patients only)

[redacted]

Table 26: Summary of time to response by R-refractory status in MAGNIFY: IITT population (MZL patients only)

[redacted]

Table 27: Summary of treatment-emergent adverse events by R-refractory status during the induction period of MAGNIFY: induction safety population

Table 28: Summary of the most common treatment-emergent adverse events (reported in ≥10% of patients) by R-refractory status in the induction period of MAGNIFY: Induction safety population

Table 29: Summary of treatment-emergent adverse events by R-refractory status during the induction period of MAGNIFY: induction safety population (FL patients only)

Table 30: Summary of treatment-emergent adverse events by R-refractory status during the induction period of MAGNIFY: induction safety population (MZL patients only)

Table 31: Summary of the most common treatment-emergent adverse events (reported in ≥10% of patients) by R-refractory status in the induction period of MAGNIFY: Induction safety population (FL patients only)

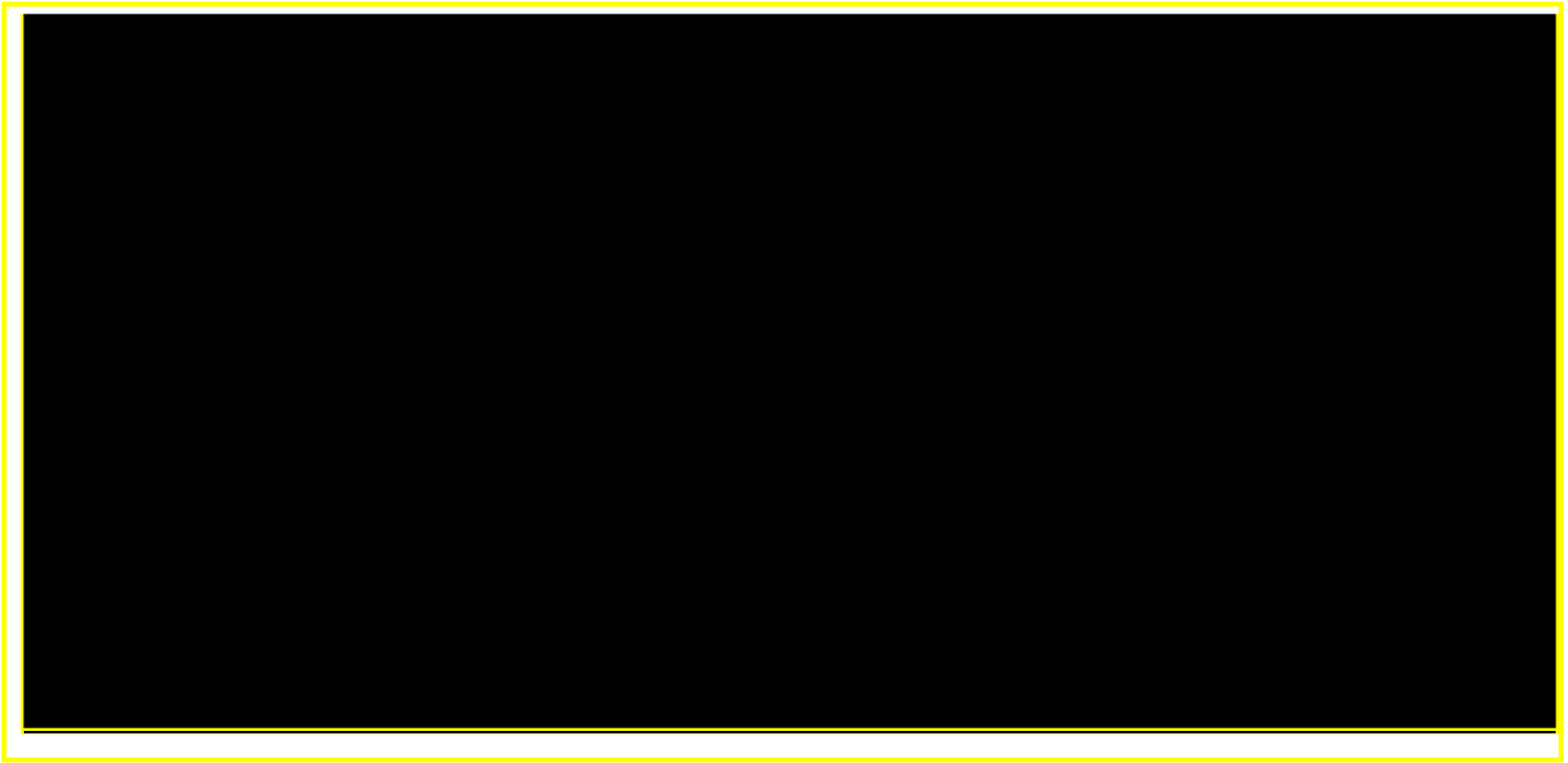
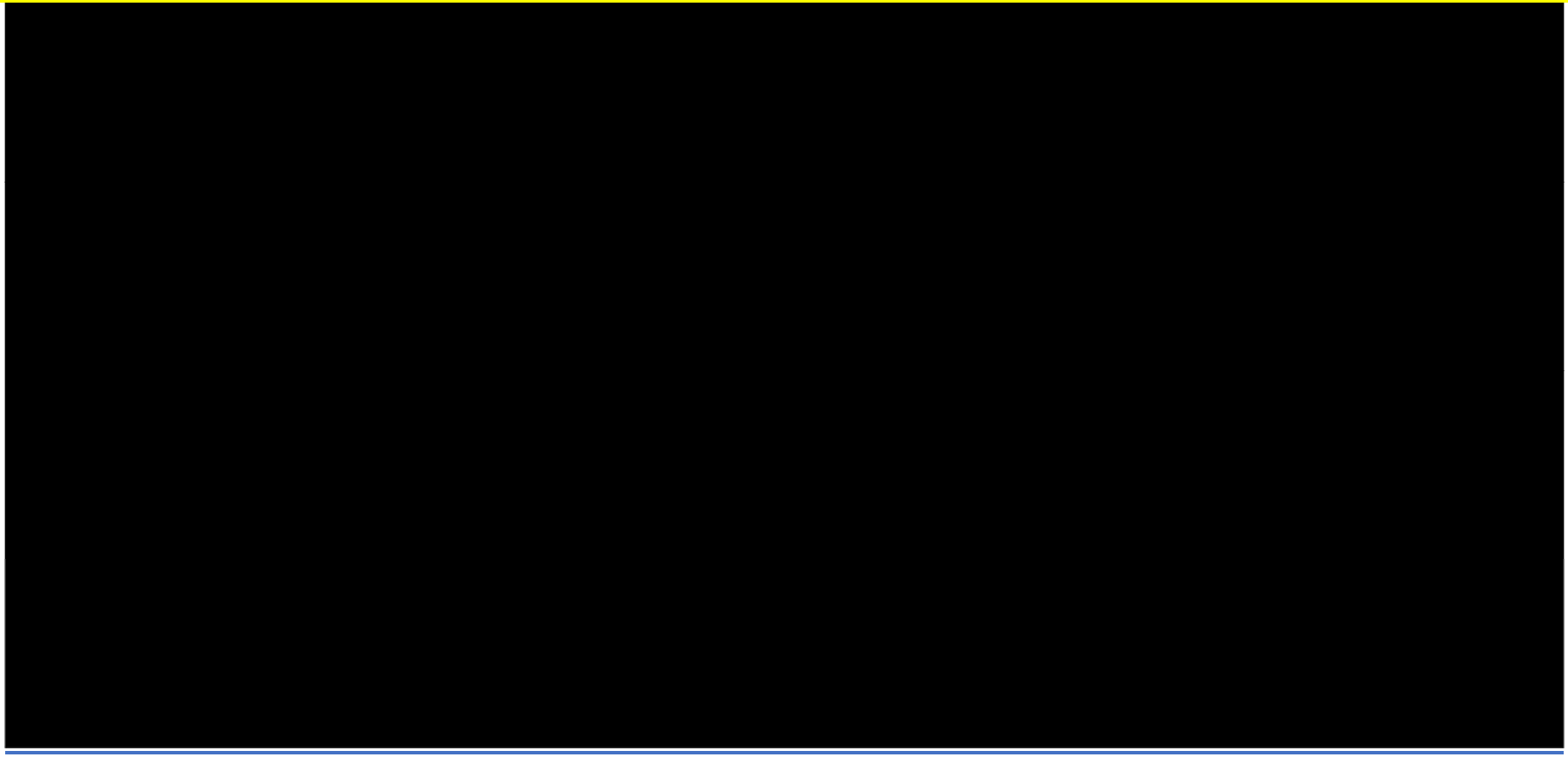
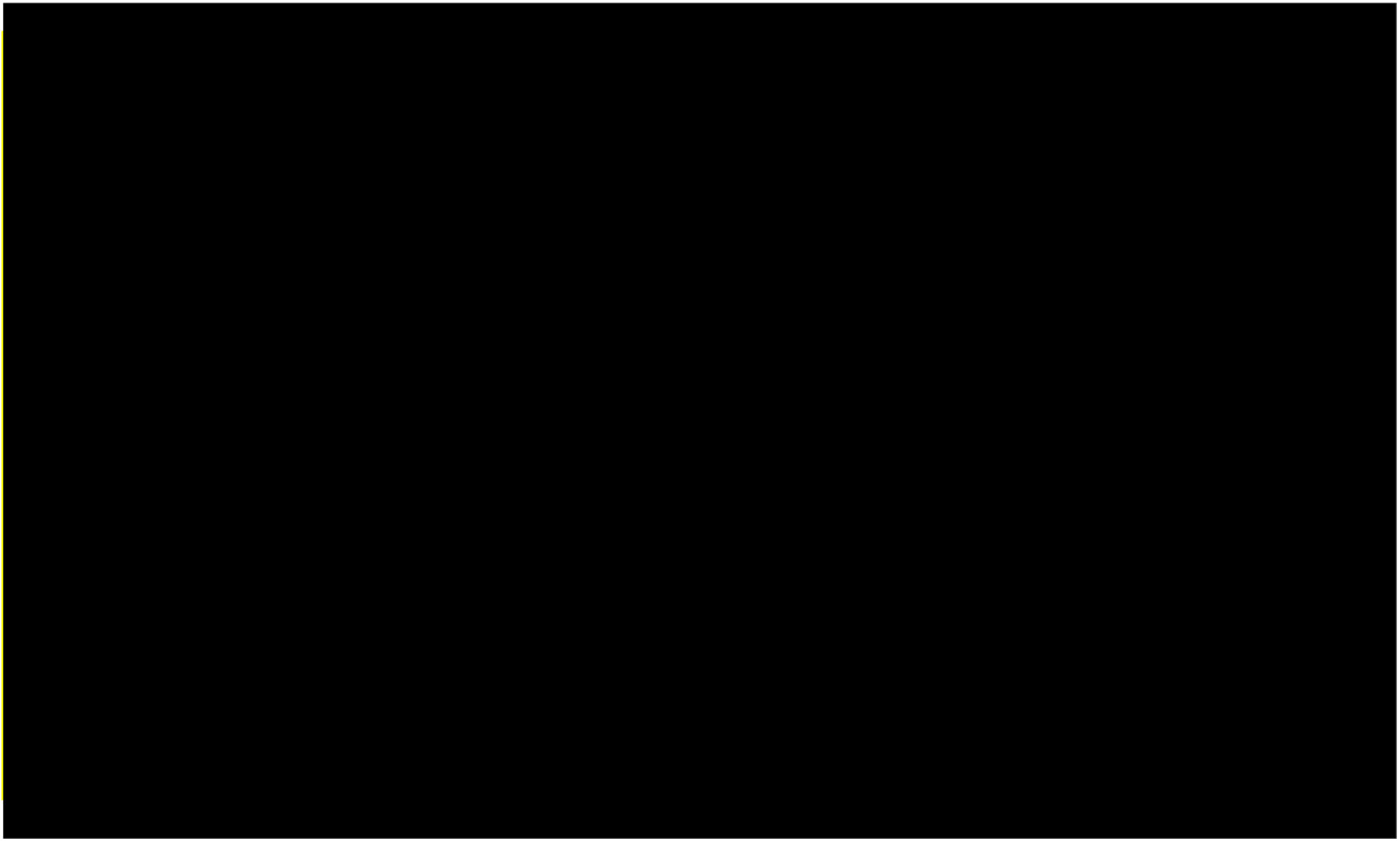


Table 32: Summary of the most common treatment-emergent adverse events (reported in ≥10% of patients) by R-refractory status in the induction period of MAGNIFY: Induction safety population (MZL patients only)





Section B: Clarification on cost-effectiveness data

Population

B1. Priority question: FL and MZL populations were pooled throughout the economic analysis due, as stated by the company, to their similar prognosis and the difficulty in sourcing MZL-specific data. However, the final scope quotes different ranges of survival for the two populations (Cancer Research UK 2004-2011, accessed June 2018). Other literature also reported different 5-year survival rates for FL and MZL.[30] Similarly, the differences between FL and MZL might imply differences in HRQoL, resource use and costs.

a. Please clarify which input parameters for the economic model were based on FL and MZL evidence and which input parameters on FL evidence only.

Table 33 summarises which model input parameters are based on FL and MZL evidence and which input parameters are based on FL evidence only.

Table 33: Population of model input parameters

Input parameters	Population histology	Justification
Efficacy: ToT/OS/PFS/TTNLT	R^2 (non-R refractory)– FL/MZL R-CHOP/CVP – FL R^2 (R-refractory) – FL O-Benda - FL	HMRN was an FL only population. O-Benda is only licensed for FL patients so only FL data was used for R^2. No efficacy data was available for comparators in the MZL setting.
Disease monitoring	FL	FL frequencies derived from ESMO guidance and previous HTA submissions.[2, 31] MZL ESMO guidance does not state frequencies but similar monitoring tests are recommended.[32]
Patient characteristics	FL	Patient characteristics were based on the comparator population which was based on FL populations.
Adverse events	R^2 (non-R refractory)– FL/MZL R-CHOP/CVP – FL R^2 (R-refractory) – FL O-Benda - FL	As per efficacy inputs.
Utilities	FL/MZL	Utility data was based on the AUGMENT population and assumed same for each treatment.
Subsequent treatments	R^2 – FL/MZL Comparators - FL	R^2 subsequent treatment data used AUGMENT and comparator data used subsequent treatment frequencies from HMRN.
Key: FL, follicular lymphoma; MZL, marginal zone lymphoma; OS, overall survival; PFS, progression-free survival; ToT, time on treatment; TTNLT, time to next lymphoma treatment		

b. Please provide evidence (e.g. from the targeted review described in B.3.11) to justify that the prognosis and comparative effectiveness are similar for FL and MZL.

See the response to question A5 above, which fully details this targeted literature review and the results from the studies identified within it.

Further to this targeted literature review, evidence to support the similar comparative effectiveness of R^2 and R-mono in FL and MZL was provided in the AUGMENT CSR.[22] The CSR presents multivariate analyses for progression-free survival based on IRC assessment per FDA censoring for the subgroup of MZL subjects (Table 34).

Table 34. Cox proportional hazard model for progression-free survival based on IRC assessment per FDA censoring rule in subjects with MZL: ITT population

	Univariate Model[a]			Final Multivariate Model[b]		
Variable	Hazard Ratio	95% CI	p-value	Hazard Ratio	95% CI	p-value
Treatment (R^2 Arm vs Control Arm)[c]	1.001	(0.471, 2.125)	0.998	0.509	(0.202, 1.284)	0.153
Age (≥ 70 vs < 70 yrs)	2.083	(0.936, 4.634)	0.072			
Sex (male vs female)	1.990	(0.939, 4.216)	0.072			
Ann Arbor Stage (IV vs I/II/III)	2.746	(1.205, 6.258)	0.016	2.436	(0.998, 5.947)	0.051
FLIPI (high vs low/medium)	1.775	(0.841, 3.745)	0.132			
ECOG (1-2 vs 0)	1.228	(0.562, 2.685)	0.607			
LDH (elevated vs not elevated)	2.348	(1.096, 5.031)	0.028	2.792	(1.131, 6.897)	0.026
B symptom (yes vs no)	0.931	(0.220, 3.941)	0.923			
High tumor burden (yes vs no)	1.465	(0.662, 3.244)	0.346			
Chemoresistant (yes vs no)	0.493	(0.067, 3.644)	0.488			
Unfit for chemotherapy (yes vs no)	2.377	(1.043, 5.416)	0.039	2.086	(0.892, 4.876)	0.090

ECOG = Eastern Cooperative Oncology Group; FDA = Food and Drug Administration; FLIPI = Follicular Lymphoma International Prognostic Index; IRC = Independent Review Committee; ITT = intent-to-treat; LDH = lactate dehydrogenase; MZL = marginal zone lymphoma; yrs = years.
[a] Model includes one risk factor.
[b] Model includes treatment arm and significant risk factors (p-value < 0.05) from univariate analyses.
[c] R^2 Arm = lenalidomide in combination with rituximab. Control Arm = placebo plus rituximab.
Data cutoff: 22 Jun 2018.

Covariates included in the multivariate model were: randomized treatment (R^2/ R-mono) and those variables that were highlighted as being significant risk factors (p-value <0.05) based on the univariate analyses (Ann Arbor Stage, LDH and Unfit for chemotherapy). These multivariate analyses were repeated for PFS based on IRC assessment per EMA censoring (in the MZL subgroup) and showed that after adjusting for the significant risk factors (Ann Arbor Stage and LDH), the HR comparing R^2 and R-mono for PFS by IRC assessment per EMA censoring in the subgroup of subjects with MZL was 0.460 [95% CI: 0.192, 1.101] in favour of R^2. This HR was similar to the HR for PFS by IRC assessment per EMA censoring comparing R^2 and R-mono in the overall population (i.e., FL and MZL; HR 0.45 [95% CI: 0.33, 0.61]). These results indicate that histology does not have a large

impact on the comparison of PFS, and the comparative effectiveness of R^2 is similar between patients with FL and those with MZL.

Exploratory post-hoc analyses of the AUGMENT trial data were also performed, which investigated the impact of the histology (MZL/FL) covariate on the outcome of PFS by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance. These additional analyses look at the impact of histology as a covariate rather than subgrouping by histology, as was performed in the CSR. Cox proportional hazard models were fitted for PFS by IRC assessment per 2007 IWGRC based on EMA guidance for three separate populations (R^2-treated patients only, R-mono treated patients only and the pooled population [i.e., both R^2-treated patients and R-mono treated patients]) with the following covariates included in each model:

- Model 1: Treatment (pooled data only), histology (FL/MZL), interaction between treatment and histology (pooled data only)
- Model 2: Treatment (pooled data only), histology (FL/MZL), interaction between treatment and histology (pooled data only), Ann Arbor Stage, LDH, unfit for chemotherapy
- Model 3: Treatment (pooled data only), histology (FL/MZL), interaction between treatment and histology (pooled data only) and all variables included in Table 34 (with the exception of B symptoms, which was not included because the patient-level data for this variable was not available at the time of analyses)

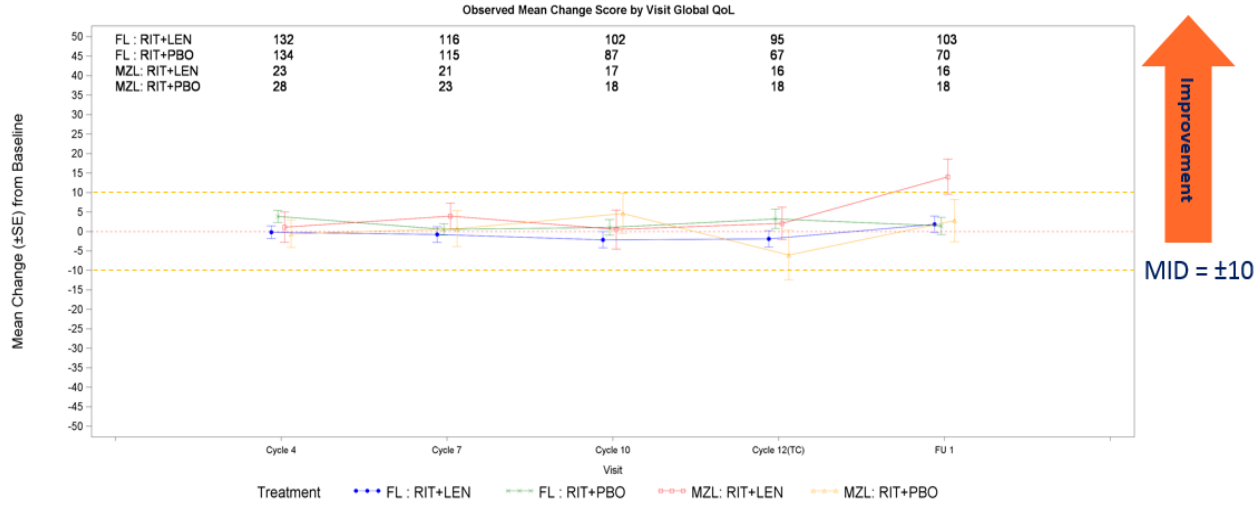
With the three different populations and the three different sets of covariates, nine models were fitted in total. In all nine of the models, histology was not indicated to be significant (p-value >0.05). These analyses therefore support the findings of the CSR that the comparative effectiveness of R^2 and R-mono in FL and MZL is similar. Furthermore, the interaction term between the randomized treatment arm and histology was not significant in any model, where included.

c. Please provide evidence to justify that HRQoL and resource use are similar for FL and MZL.

From the systematic literature reviews conducted as part of the NICE submission process, no evidence was found which reported MZL specific utilities or resource use. The ESMO guidelines for MZL patients do not provide suggested frequencies for routine disease monitoring but do recommend that mandatory follow-up should be a full blood count and CT scan with recommended blood flow cytometry.[32] These are similar to the tests that are recommended for FL routine monitoring in ESMO guidelines.[31] Due to lack of evidence, clinical opinion was sought as to the expected differences between FL and MZL patients. At an advisory board, clinicians confirmed that health-related quality of life for FL and MZL patients would be the same.[21] At following validation meetings, clinicians stated that disease monitoring would be expected to be the same between FL and MZL patients. One minor additional test for approximately 50% of MZL patients would be serum paraprotein which would be on the same schedule as the other blood tests. As this only covered half the MZL population and did not include any additional visits or blood tests this was not costed for separately and resource use was assumed to be the same between FL and MZL patients.

In the QoL analysis conducted on the AUGMENT study, QoL was compared by histology groups (i.e., FL vs. MZL) on the assessed by the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire, Core 30 (EORTC QLQ-C30) and EuroQoL Five Dimension Three Level (EQ-5D-3L) Questionnaire.[33] Mean changes from baseline (HRQoL-evaluable population) between treatment groups by FL/MZL in the EORTC QLQ-C30 global health status/QoL domain are presented in Figure 8. Overall, the results show no clinically meaningful difference between treatment groups within each subgroup of disease histology (FL/MZL) across all assessment visits.[33] If histology was included in the mixed effects regression model used for the CE model utilities, this results in a mean utility difference of 0.03 for MZL patients, however this was not statistically significant (p=0.145).

Figure 8. Mean Change from Baseline in Global Health Status/QoL between Treatments by Disease Histology (FL/MZL)



FL = follicular lymphoma; FU = follow-up; LEN = lenalidomide; MID = minimally important difference; MZL = marginal zone lymphoma; PBO = placebo; QoL = quality of life; RIT = rituximab; SE = standard error; TC = treatment completion
Subjects within each lymphoma group were separated by treatment arm.
The p-value was calculated based on an ANCOVA test, comparing R2 versus control within a lymphoma group, while controlling for baseline scores.
Source Celgene, 2018[33]

Additionally, treatment effects on mean changes from baseline in the exploratory domains of interest did not differ by disease histology in most domains.[33] Although MZL patients in the R^2 arm seemed to have a greater improvement across most of the assessment visits in the QLQ C30 emotional and functional domains, and in the EQ-5D VAS, but a greater deterioration in the constipation domain, compared with MZL patients in the R mono arm, such findings were inconclusive because of the small MZL patient numbers. Additionally, patients in the R^2 arm had a greater deterioration in the diarrhoea domain across most of the visits during the treatment period compared with patients in the R mono group, regardless of disease histology.[33]

No studies have been identified within the literature that directly compare HRQoL or resource use of FL versus MZL patients. Several studies have reported HRQoL detriments associated with FL (described in Section B1.3 of Document B). Another study (not previously detailed in Document B)

demonstrated similar impacts on HRQoL in a population of 97 iNHL patients, of which the majority were MZL patients (67 with MZL, 27 with FL, 3 with other iNHL).[34] Global QoL was reduced in this iNHL population with greatest impacts seen on the emotional and social functioning domains, and the fatigue and insomnia symptom domains of the EORTC-QLQ-C30.[34] Quality of life and symptoms scores as measured on the EORTC QLQ-C30 were similar in this study by Kang and colleagues (2018) for a mostly MZL population to those in the study by Oerlemans and colleagues (2014), which reported QoL of 148 FL patients.[35]

d. In line with question A10, please provide a scenario analysis (and the accompanying model) with separate parameter estimates for FL and MZL for OS, response rate, health related quality of life, adverse events, and resource use.

Due to the lack of comparator data in the MZL population, a scenario analyses (and accompanying model) using separate parametric estimates for MZL is not possible for the relevant comparators. However, it is reasonable to analyse a combined population of relapsed MZL and FL patients, given the similarities in their prognosis (see response to questions B1b and A5) and their similar burden to patients (see response to question B1c), as well as expert clinical opinion from an advisory board that their management and treatment outcomes were broadly similar.[21] Furthermore, there is a need to retain the MZL patients within the analysis given the vast unmet need in this population as highlighted in Section B1.3 of Document B, specifically in that there are currently no licensed treatments specifically indicated for patients with relapsed/refractory MZL.[21] As such we maintain the appropriateness of pooling the FL and MZL histologies as per the company base case.

At the request of the ERG a scenario analysis (and accompanying model) has been conducted for the FL only population for the non-R-refractory population comparisons.

ITCs with data from HMRN

For the comparison with R-CHOP and R-CVP, the same methodology has been used as described in Section B.2.9 of the original submission document. Table 35 provides a summary of the number of patients and events in the MAIC analysis for OS, TTNLT and PFS.

Table 35: Number of patients and events in the MAIC analyses for the R^2 (FL only AUGMENT) versus R-CVP/R-CHOP (HMRN) comparison

Treatment	N	Events	Median survival
OS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
PFS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]

Treatment	N	Events	Median survival
TTNLT			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
Key: HMRN, Haematological Malignancy Research Network; MAIC, matching-adjusted indirect comparison; NA, not applicable; OS, overall survival; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; TTNLT, time-to-next antilymphoma therapy.			

HRs from the Cox Proportional-Hazards models comparing R^2 and R-CHOP/CVP are presented in Table 36. These results demonstrate that R^2 has significant benefit for OS and TTNLT compared to R-CHOP/CVP, with modest PFS improvement which are consistent with the pooled FL and MZL analysis.

Table 36: Results from Cox Proportional Hazard models comparing R^2 (FL only) and R-CVP/R-CHOP

Outcome	R^2, adjusted N	R-CVP/R-CHOP N	HR (95% CI)[a]
OS	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]
TTNLT	[redacted]	[redacted]	[redacted]
Key: CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival; PFS, progression-free survival; R^2, rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone TTNLT, time to next antilymphoma treatment. **Notes**: [a], bootstrapped CI.			

Efficacy

For the economic model, parametric curves were fitted to the weighted R^2 Kaplan-Meier's for OS, PFS, TTNLT and ToT. Table 37 present the AIC and BIC fit statistics for R^2 for all end points in the model and Figure 9 - Figure 11 present the parametric distributions fit to the R^2 Kaplan-Meier data for OS, PFS and TTNLT, respectively.

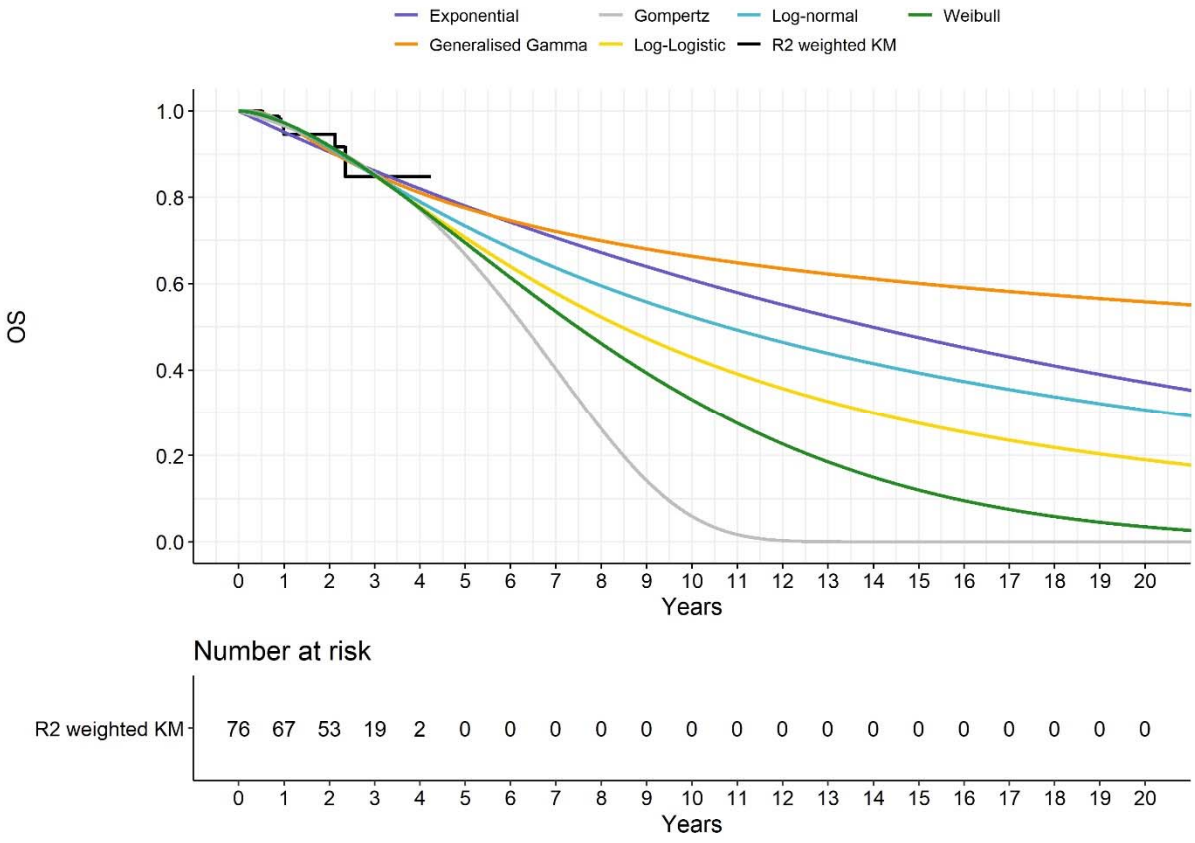
Due to similarities in the FL only data compared to the ITT data set, and for consistency, the same distributions have been selected for the base case as per the pooled FL & MZL analysis. These are; Weibull distribution for OS, R^2 Kaplan-Meier then R-CHOP/CVP hazard for PFS, log-normal for TTNLT and the Kaplan-Meier directly for ToT.

Table 37: AIC and BIC – R² (FL only) versus R-CHOP/R-CVP

Distribution	OS		PFS		TTNLT		ToT	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	177.67	180.66	457.55	460.54	316.51	319.50	1028.70	1031.69
Generalized gamma	178.71	187.68	456.00	464.97	318.29	327.26	828.23	837.18
Gompertz	178.68	184.66	458.82	464.80	318.08	324.06	830.22	836.19
Log-logistic	177.52	183.50	454.41	460.39	316.38	322.36	931.13	937.10
Log-normal	177.03	183.01	454.00	459.99	316.32	322.30	974.55	980.52
Weibull	177.65	183.63	456.14	462.12	316.86	322.84	886.16	892.12

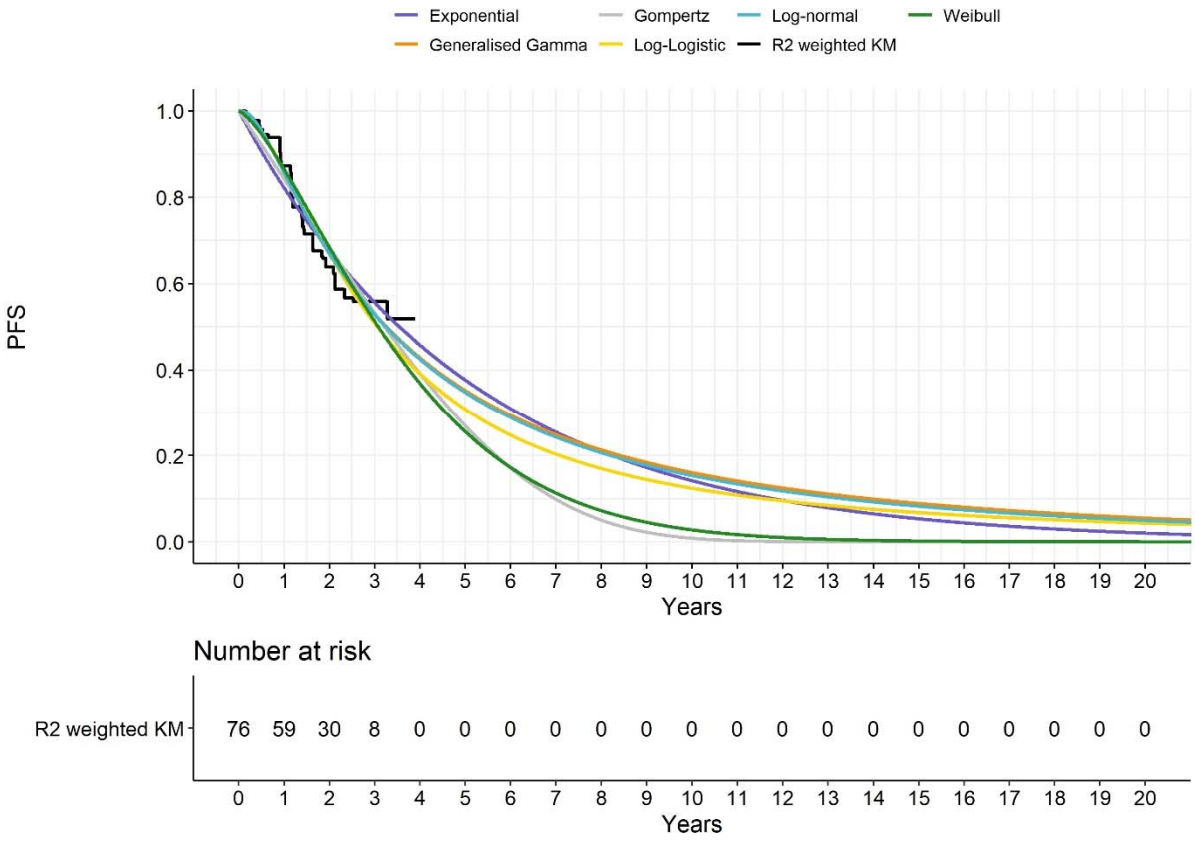
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; OS, overall survival; PFS, progression-free survival; R, rituximab; R², lenalidomide plus rituximab, ToT, time on treatment; TTNLT, time to next anti-lymphoma treatment.

Figure 9: OS parametric curves for R² (FL only) in the non-rituximab refractory population



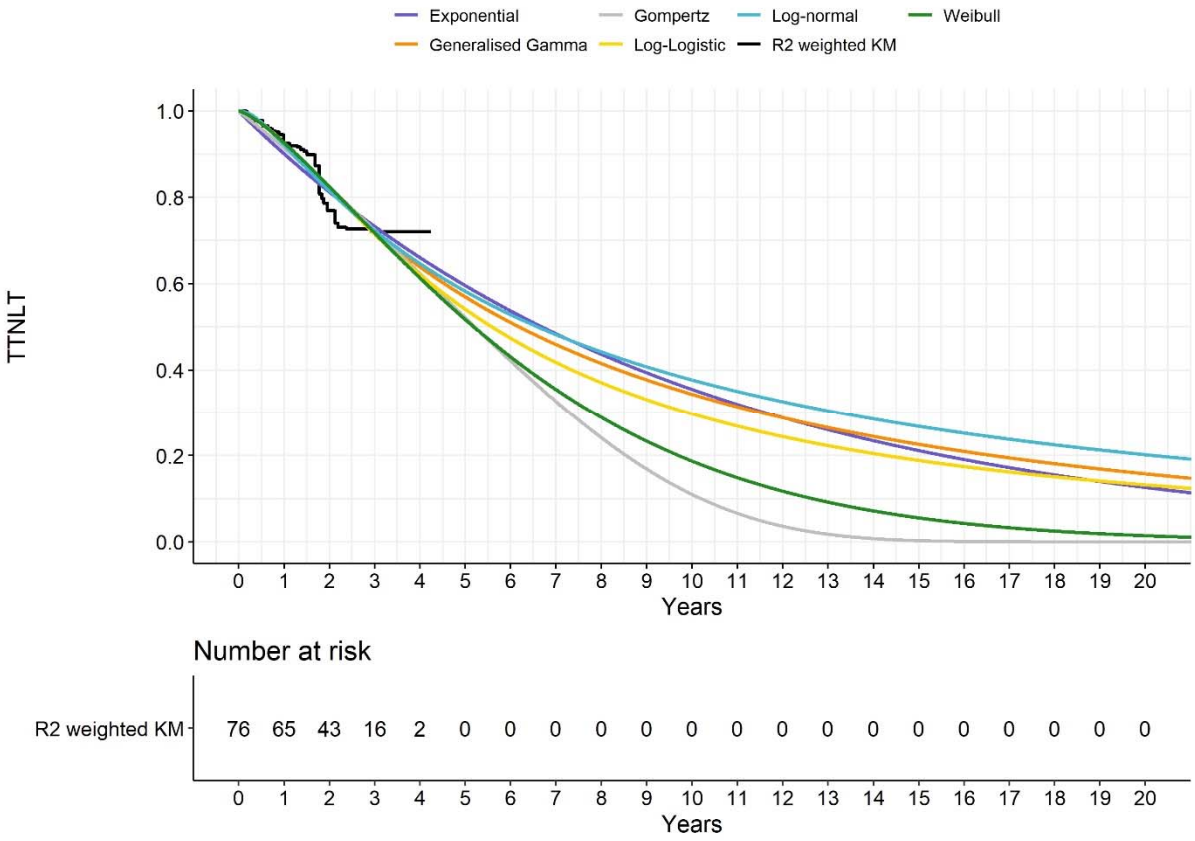
Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R², lenalidomide plus rituximab.

Figure 10: PFS parametric curves for R^2 (FL only) in the non-rituximab refractory population



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Note: For the base case, R^2 is assumed to assumed to follow the KM and then use the hazard of R-CHOP/R-CVP.

Figure 11: TTNLT parametric curves for R² (FL only) in the non-rituximab refractory population



Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R², lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

Adverse events

As described in Section B3.3, grade 3/4 adverse events of greater than 2% in either treatment arm was considered. For this scenario, only the AEs from the FL population in AUGMENT were used to calculate the incidence and resulting costs of adverse events. As per the FL and MZL pooled population, the adverse events from RELEVANCE were adjusted to the AUGMENT data and used to inform the adverse event incidence for the R-CHOP and R-CVP arms. The same costs presented in Section B.3.5 were applied to these adverse event rates to calculate the total adverse event cost for each treatment. Table 38 presents the adverse event incidences and total costs per treatment arm.

Table 38: Grade 3/4 AE incidence: non-rituximab (FL only) refractory population

AE	R² (n=146)		R-CHOP (%)	R-CVP (%)
	n	%		
Neutropenia	74	50.7%	91.5%	86.5%
Leukopenia	9	6.2%	27.8%	15.0%
Anaemia	6	4.1%	4.1%	4.1%
Pneumonia	5	3.4%	NR	NR
Lymphocyte count decreased	5	3.4%	NR	NR
Lymphopenia	5	3.4%	NR	NR

AE	R^2 (n=146)		R-CHOP (%)	R-CVP (%)
	n	%		
Febrile neutropenia	4	2.7%	9.0%	4.9%
White blood cell count decreased	5	3.4%	NR	NR
Diarrhoea	5	3.4%	1.9%	6.7%
Thrombocytopenia	2	1.4%	1.2%	0.0%
Hypokalaemia	4	2.7%	NR	NR
Pulmonary embolism	4	2.7%	NR	NR
Infusion related reaction	1	0.7%	0.1%	0.0%
Nausea and emesis	0	0.0%	1.4%	3.8%
Allergic reaction	1	0.7%	NR	NR
Hypotension	1	0.7%	NR	NR
Fatigue	2	1.4%	3.8%	0.0%
Alopecia	NR	NR	0.8%	0.0%
Abdominal pain	1	0.7%	0.7%	0.0%
Acute kidney injury	2	1.4%	NR	NR
Total cost	**£1,796**		**£3,471**	**£2,714**
Source	AUGMENT		RELEVANCE[18] (adjusted)[a]	

Key: AE, adverse event; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; NR, not reported; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a] Comparator AE incidence = ($AE_{COMPARATOR}$ incidence in RELEVANCE/AE_{R2} incidence in RELEVANCE) x AE_{R2} incidence in AUGMENT.

Health related quality of life

The EQ-5D meta-regression model was re-ran from AUGMENT using only the FL population and was used to inform the utilities for this scenario. The same methodology was applied as described in Section B.3.4 of the original submission.

Table 39: Parsimonious mixed effects FL only model

Parameters	Estimate	SE	p-value
Intercept	0.341	0.037	<0.001
Health state: Pre-progression	0.026	0.010	0.016
Health state: Progressed on treatment	-0.035	0.014	0.015
Randomized treatment arm: R^2	0.002	0.015	0.918
Baseline utility	0.548	0.039	<0.001
Previous rituximab exposure: no	0.037	0.021	0.078
Refractory to last therapy: yes	-0.046	0.021	0.031
Number of prior therapies: 1	0.039	0.015	0.012

Key: R^2, lenalidomide plus rituximab; SE, standard error.

Table 40: Final utility values from the FL only model (R^2 vs R-CHOP/CVP)

Health state	Utility
Progression-free	0.867
Progressed (off treatment)	0.841
Progressed (on treatment)	0.806

Treatment costs

The dose reductions for lenalidomide data has been taken directly from the AUGMENT trial using only the FL patients for the analysis (as described in Section B.3.5 in the original submission document). These are then used to calculate the cost per cycle applied to the proportion of patients on treatment.

Subsequent treatments

Subsequent treatments usage from the FL only population in AUGMENT were used to inform the costs of subsequent treatments for the R^2 arm in this analysis. The durations of the subsequent treatments and cost per treatment remain the same as per Section B.3.5 in the original submission document. Table 41 presents the FL only subsequent treatment usage from AUGMENT and the resulting subsequent treatment costs for the R^2 arm.

Table 41: Subsequent treatments and costs

Subsequent treatment	R^2 (FL only AUGMENT) n (%)[a] n=147
R-mono	[redacted]
R-Benda	[redacted]
R-CHOP	[redacted]
R-CVP	[redacted]
Other R-chemo	[redacted]
O-Benda	[redacted]
Bendamustine	[redacted]
Other chemotherapy	[redacted]
Targeted therapies	[redacted]
Radiotherapy	[redacted]
Other	[redacted]
ASCT	[redacted]
Total weighted treatment cost (R^2)	**£3,128**
Total weighted administration cost (R^2)	**£349**

Key: ASCT, autologous stem cell transplant; Benda, bendamustine; chemo, chemotherapy; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; mono, monotherapy; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a]Percentages include multiple lines therefore total may be over 100%.

Results

Deterministic results are presented in Table 42. As per the request in clarification question B3, the FL only analysis has also been conducted for the R^2 versus R-mono comparison with results also presented in Table 42. The results show that R^2 remains cost-effective in the FL-only population versus R-CHOP and R-CVP as well as being cost-effective versus the AUGMENT study comparator R-mono at the £30,000 willingness-to-pay threshold.

Table 42: Deterministic pairwise results (based on PAS price) – FL only

Technology	Total			Incremental			ICER (£/QALYs)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
R^2 versus R-CHOP							
R-CHOP	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R^2 versus R-CVP							
R-CVP	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R^2 versus R-mono							
R-mono	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£20,310

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; LYG, life years gained; mono, monotherapy; PAS, patient access scheme; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab.

Model structure

B2. The use of a partitioned survival analysis model instead of a state transition model (STM) was justified by the lack of data of relevant comparators in the head-to-head study with R2 to inform a state transition model. Despite the potential limitations of a state transition model, a partitioned survival analysis has several limitations related to the extrapolation (as mentioned in NICE DSU TSD 19[36]). Please provide a scenario analysis (and the accompanying model), enabling STM as a scenario, as recommended in TSD 19.[36]

The company accepts that there are often limitations with any modelling approach and considered both partitioned survival models and state-transition models during the model conceptualisation phase. Based on the reasoning presented in Document B Section B.3.2, the partitioned survival model was considered the most appropriate structure and used for the cost-effectiveness model. Although the NICE DSU TSD 19 recommends presenting state-transition models alongside the partitioned survival model to assist in verifying the plausibility of the partitioned survival models extrapolations [36], the weight of the limitations in the STM approach, combined with the specifics of the

data available mean the company do not believe constructing a state transition model is applicable for this submission.

The main limitation of the state-transition model is the use of unrandomized end points to model transitions such as post-progression survival. This is highly prone to bias due to the selection effects and informative censoring.[36] Certainly, within the FL and MZL setting, the timing of relapse can influence the next treatment option and impact on response to the next treatment affecting overall survival.[21] Given that the data available are either immature or have small patient numbers, the use of such data to inform post-progression or post next therapy outcomes could be misleading. Extrapolating outcomes from a group of patients who no longer have comparable characteristics and based on patients who are progressing early would be bias against the R^2 arm. This is not an issue when using OS directly from the point of randomisation as all patients contribute to the function used to fit the curve. The OS hazard currently observed for the R^2 arm, reflects those patients who die quickest, hence the distributions selected for the base case OS (Weibull versus R-CHOP/CVP) is likely to be conservative. Therefore, providing a state-transition model will not resolve uncertainty in the OS estimates and would likely add further uncertainty, given the likely biased estimates these analyses will produce. OS extrapolation uncertainty has already been thoroughly tested in the model by testing different parametric distributions and probabilistic sensitivity analysis. As such the company do not believe it worthwhile providing these analyses for review.

In addition to the above, further limitations to providing a state-transition model involve underlying data availability and complexity of the approach to allow for all possible transitions within the CE model itself.

For the comparison with O-Benda, OS, PFS and TTNLT was reported and used for the partitioned survival model. However, data required for the state-transition model such as post TTNLT survival and deaths from progression-free off treatment are not specifically reported. In addition, data from MAGNIFY which was used to inform the R^2 R-refractory population was limited to the induction phase as discussed in Section B.2.3 of the company submission. Therefore, data to inform later transitions such as post-progression survival and post- TTNLT survival are not available. A state-transition model for this comparison would therefore not be possible.

For the comparison with R-CHOP/R-CVP the total number of patients for this cohort is [redacted], thus the later model transitions (i.e. post-progression (off treatment) to death) are based on even smaller patient numbers and subsequently small numbers of events creating additional uncertainty in the extrapolated outcomes for later model transitions. Moreover, as this is based on real-world evidence, patients are only assessed for response after a pre-defined number of cycles of treatment, instead of as routinely as in the AUGMENT trial. Therefore, independent transitions between the different health states may not be comparable versus R^2.

For the comparison with R-mono, the later transitions which will be required for the state-transition model would still be based on immature data. Another limitation of the state-transition model is that it

does not negate the need to extrapolate data, therefore extrapolating more immature data (such as post-progression (on-treatment) survival) produces more uncertain estimates for those particular transition probabilities and hence creating more uncertain OS projections from the final model outputs. Additionally, as R-mono is not considered a relevant comparator, the company feels that providing this analysis would not aide the decision makers over the cost-effectiveness of R^2 in relation to appropriate comparators or help validate the extrapolated outcomes from a partitioned survival model.

The current model structure includes three health states which are further split into sub health states for when patients are 'on' or 'off' treatment. For a state-transition model, the development of a five-health state model using time-dependencies in event rates for each possible transition was considered to add significant complexity based on the number of tunnel states that would be required to accurately model the transitions. This would create unnecessary computational complexity that would make the model burdensome to run. This is particularly so when we consider the multiplicative nature of any analyses with a complex decision problem involving multiple subgroup and comparators.

Furthermore, the issues of bias from unrandomized endpoints and limited numbers to inform each transition are compounded by the numerous states required to provide an adequate simplification of the condition. In this sense our specific case differs from the DSU case study which primarily considered a three state STM.

Based on the above points, the company feels that providing a state-transition model would not offer the decision makers any additional clarification in relation to the cost-effectiveness of R^2 and therefore have not generated a state-transition model.

Intervention and comparator

B3. Please provide a scenario analysis (and the accompanying model) including the comparators listed in question A7 and evaluated in response to that question.

In response to A7b, an economic evaluation against rituximab monotherapy based on the head-to-head comparison of the AUGMENT trial was conducted. The resulting ICER for R^2 versus rituximab monotherapy is given in Table 43. However, it is worth noting that rituximab monotherapy is not considered a relevant comparator because it is rarely used in the relapsed/refractory setting in the UK, as discussed in response to A7b.

Table 43: R^2 versus R-monotherapy pairwise results

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R-monotherapy	[redacted]	[redacted]	[redacted]				

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£22,580
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year.							

As discussed further in the response to A7c., advice from clinical experts from an advisory board suggests that O-Benda is the most likely comparator in patients who are refractory to rituximab.[21] We therefore maintain that a comparison to O-Benda is the most relevant for the rituximab-refractory population and no further scenario analysis is presented for that population.

B4. Based on clinical opinion, in the CS it was stated that "it would not be unreasonable to assume that the efficacy of R-CHOP and R-CVP are similar in the relapsed/refractory setting". Please provide a detailed justification supported with clinical evidence why it was considered appropriate to model R-CHOP and R-CVP as a single comparator with equal effectiveness instead of modelling both comparators separately. If applicable, please also provide a scenario analysis (and the accompanying model) including R-CHOP and R-CVP as separate comparators.

As discussed in response to A14, and presented in the Appendix D.3., data from HMRN showed no evidence to suggest a treatment effect between R-CHOP and R-CVP. This is also backed up by clinical opinion that it is not unreasonable to assume that R-CHOP and R-CVP have similar efficacy in the relapsed/refractory setting from multiple advisors post advisory board. From the HMRN data base, the R-CHOP cohort had a total of [redacted] patients and R-CVP had a total of [redacted] patients. Due to these small patient numbers on the individual data sources and the evidence suggesting equal efficacy is plausible, the data was pooled to form one larger efficacy set for more accurate analyses.

Effectiveness

B5. Please provide a scenario analysis (and the accompanying model) using the analyses mentioned in questions A15 to A17 and performed in response to these questions.

As discussed in response to B4, it was not considered appropriate to model R-CHOP and R-CVP as separate comparators. Therefore, no additional scenario analysis is produced in response to clarification question A15.

The definitions used in the HMRN-MAIC analysis were the same in the HMRN database as in the AUGMENT trial, therefore no additional scenario analysis is required in response to clarification question A16.

The response to clarification question A17 contains MAIC results based on a set of matching variables in which 'Bone marrow involved' was included. The MAIC has been included within the cost-effectiveness model and the results of a scenario analysis using this MAIC are presented in Table 44. Including 'Bone marrow involved' in the matching variables used in the base case HMRN-MAIC analysis had little impact on overall results and shows some small improvement to the ICERs of R^2 versus R-CHOP (£11,018 versus £11,471 at base case) and R^2 versus R-CVP (£16,123 versus £16,814 at base case).

Table 44: Scenario analysis pairwise results: HMRN-MAIC analysis including 'Bone marrow involved' as a matching variable (with PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R^2 versus R-CHOP							
R-CHOP	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£11,018
R^2 versus R-CVP							
R-CVP	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,123

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

B6. The company considered TTNLT as a more relevant endpoint than PFS (given that patients who progress via radiographic scans may not necessarily receive subsequent treatment immediately if not symptomatic) in terms of costs of subsequent treatments, quality of life and the impact on OS.

a. Please justify the appropriateness of comparing PFS from the trial data (proactively on study) and PFS from the HMRN clinical practice database (reactively in real-world setting).

PFS was the primary endpoint for the AUGMENT study and is therefore an appropriate endpoint to include in comparison with the real-world data set derived from HMRN. However, comparing PFS from trial data and clinical practice has its limitations due to the discrepant approaches in determining disease progression (as described in section B2.6 of Document B in the company submission). TTNLT offers a clinical end point with additional value beyond PFS and is more readily comparable

between clinical studies and real-world practice. The importance of this endpoint was also highlighted by the committee for TA513.[37] Therefore, TTNLT was included as an additional endpoint to help address some of the limitations of examining PFS and provide for a more like-for-like comparison.

b. Please provide a scenario analysis (and the accompanying model), removing PFS and using TTNLT in all data sources (trial data and HMRN data), assuming no progression before starting post progression treatment (i.e. assuming PFS is identical to TTNLT).

For the scenario analysis removing PFS, the utility regression analysis was also reconducted removing progression as a covariate (see Table 45). The overall utility value per health state are summarized in Table 46.

Table 45: Parsimonious mixed effects model: progression removed as a covariate

Parameters	Estimate	SE	p-value
Intercept	0.341	0.035	<0.001
Randomized treatment arm: R^2	0.011	0.014	0.451
Health state: TTNLT on treatment	-0.045	0.012	<0.001
Baseline utility	0.566	0.037	<0.001
Previous rituximab exposure: no	0.030	0.019	0.124
Refractory to last therapy: yes	-0.037	0.021	0.071
Number of prior therapies: 1	0.031	0.015	0.033
Key: R^2, lenalidomide plus rituximab; SE, standard error; TTNLT, time to next antilymphoma treatment.			

Table 46: Utility values per treatment comparison: progression removed as a covariate

Health state	R^2 vs R-CHOP/R-CVP	R^2 vs O-Benda
Prior next lymphoma treatment (i.e. progression-free when assuming PFS is identical to TTNLT)	0.856	0.806
Post next lymphoma treatment (i.e. post progression when assuming PFS is identical to TTNLT)	0.812	0.761
Key: PFS, progression-free survival; TTNLT, time to next anti-lymphoma treatment.		

The results of the scenario analysis which assumes PFS is identical to TTNLT are presented in Table 47.

Table 47: Scenario analysis pairwise results: assuming PFS is identical to TTNLT (with PAS)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R² versus R-CHOP							
R-CHOP	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£9,573
R² versus R-CVP							
R-CVP	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,857
R² versus O-Benda							
O-Benda	[redacted]	[redacted]	[redacted]				
R²	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£87,868,670 (SW)

Key: Benda, bendamustine; ICER, incremental cost-effectiveness ratio; LYG, life years gained; O, obinutuzumab; QALY, quality-adjusted life year; R2, lenalidomide plus rituximab; SW, southwest.

B7. Please justify and clarify how deaths were handled (as events or censored) in the analysis of TTNLT. If deaths were not analysed as events in the calculation of TTNLT, please provide a scenario analysis (and the accompanying model) where deaths are handled as events, as in the analysis of PFS, for the estimation of TTNLT.

All TTNLT data from HMRN and AUGMENT included deaths as events. Patient-level data for O-Benda was not available, therefore TTNLT was digitised from data reported in TA472 in response to clarification questions.[2] It is unclear whether deaths were classed as events or censored in this data set, however, as TTNLT is assumed to be the same between R² and O-Benda, we do not anticipate this to have much impact on overall results.

B8. Treatment waning starts after five years, as with previous NICE submissions (TA472 and TA137). Please further justify, given the differences (e.g. related to the evidence to estimate relative effectiveness, treatment, population) between the current submission and TA472 and TA137, why selecting a treatment waning effect of 5 years was considered appropriate.

Both TA472 and TA137 were in the relapsed/refractory FL setting. A 5 year treatment benefit was assumed by the company in TA137 and was considered reasonable by the committee.[3] The manufacturer in TA472 changed the treatment effect assumption to be 5.5 years post-consultation based on the last observed event being at 4 years and a further 18 months extrapolation.[2] Neither appraisals appeared to present evidence to support these assumptions and it was clear from the

discussion in TA472 that treatment effect was a key uncertainty and had large impacts on the ICER. There was no evidence to suggest an appropriate time point at which treatment effect diminishes and attempts were made to gain clinical opinion, however it was difficult for the clinicians to give an approximate time frame. Due to the limited evidence, the base case value of 5-years was chosen to reflect the outcomes of the previous submissions in the same disease area and scenario analysis used to test the impact of this parameter. Five years is considered conservative as the immunomodulatory effect of lenalidomide could promote a longer treatment effect versus R-chemo's. Table 68, Table 69 and Table 70 in Document B demonstrate that this time point does not have a huge impact on the results when tested at 3 years or 10 years (Table 48).

Table 48: Scenario analysis on treatment effect (with PAS)

Comparison	Base case ICER (treatment effect = 5 years)	Scenario: treatment effect = 3 years	Scenario: treatment effect = 10 years
R^2 vs R-CHOP	£11,471	£13,746	£10,183
R^2 vs R-CVP	£16,814	£20,471	£14,656
Key: ICER, incremental cost-effectiveness ratio Notes: Comparison to R^2 vs O-Benda is not shown as the base case assumes equal efficacy therefore treatment effect does not have any impact.			

B9. Priority question: Time-to-event data have been extrapolated using parametric survival distributions, based mainly on the statistical fit (AIC and BIC) to the KM data. However,

a. Please describe the efforts undertaken to validate the long-term estimates (for PFS, TTNLT and OS) based on expert opinion and/or external data. Please include detailed information regarding the questions that have been asked (if expert opinion) and the accompanying responses.

Clinical validation was sought to validate extrapolated outcomes. This was first done at an advisory board, where extrapolated curves from the AUGMENT trial were presented (extrapolated outcomes for the comparators were not available at this time). This advisory board consisted of 6 clinical experts within this disease area. A following clinical validation was conducted as an extension to the advisory board via telephone conference with one of the clinical experts from the advisory board. Extrapolated outcomes of the R-CHOP/CVP data were presented and discussed. Details of the discussions from these meetings are presented in Table 49.

Table 49: Clinical validation of extrapolated outcomes

Clinical validation	Topic	Response
UK advisory board	Extrapolation of OS from AUGMENT - Advisors were first asked to consider the long-term extrapolations of	The advisors felt that the lower of the curve distributions were more plausible as these are more in line with approximately 40–50% survival at 10 years.

Clinical validation	Topic	Response
	OS for patients receiving rituximab monotherapy.	
	Advisors were then asked to consider the long-term differences expected for R^2.	Advisors thought that the extrapolated differences in OS would decrease over time.
Extension to advisory board	The clinical expert was asked to assess the six extrapolated curves used for R-CHOP/CVP OS.	The clinical expert thought the bottom 3 curves (exponential, Weibull and log-logistic) were more plausible than the top 3 (generalised gamma, Gompertz, and log-normal). The bottom 3 curves estimated OS at 20 years between 30% - 21%. They felt that any of these could be plausible and suggested to choose the middle curve (Weibull), as the 20-year OS estimate of 27% seems reasonable and wouldn't expect anyone to challenge this.
Key: OS, overall survival		

There is limited long-term external data in the relapsed/refractory setting which can be used to inform long-term estimates. As OS is a key driver in the model results, it was important to ensure that these estimates were plausible, and the clinical guidance received from the validation meetings were used when choosing the base case OS curves. In addition to clinical guidance, AIC, BIC statistics and visual fit were also considered. For the other end points, PFS and TTNLT, the following criteria was considered when choosing the appropriate curve:

- AIC/BIC and visual fit were taken into account.
- The long-term extrapolations were assessed by the time point in which it crossed the other curves. Given that a patient's mean starting age was between 63-65 in the model, 15 years was determined to be a reasonable time point during which curves crossing would be considered implausible.
- The average time within the Progressed on-treatment health state was considered between treatment arms. Given the uncertainty on the effect of treatment once a patient progresses and moves to another treatment, the average time spent in this health state was assessed and curve selected to ensure that these were similar between treatment arms i.e. the mean life-years of patients in the Progressed (on-treatment) health state were similar between treatment arms.

Based on the above criteria, all curves selected for all treatment arms in the model were considered the most clinically plausible, statistically and visually best fitting and did not cross at inappropriate time points.

b. Please provide additional diagnostic plots to assess the visual fit of the parametric survival distributions using the observed data (including smoothed

hazard versus time; LN(smoothed hazard) versus time and LN(cumulative hazard) versus LN(time)).

Log-cumulative hazard plots (LN(cumulative hazard) versus LN(time))) were presented in Section B.3.3 of the submission. The plots show time on the log-scale. Presenting time on the log scale aids interpretation of the log-cumulative hazard plots as time is then comparable to the KM data.

The plots for the smoothed hazard versus time and log smoothed hazard versus time can be found below. The plots for the comparator data were created using the bshazard function from the bshazard package in R. As the bshazard function does not include a weight statement, the plots for the weighted R^2 data were produced by estimating the hazard from a life table and smoothing the hazard using B-splines with degree of 1 and 31 knots (as is the default in the bshazard function). As stated above, as well as visual fit to the data and statistical fits, other criteria were considered when selecting the base case curves.

R^2 (AUGMENT) versus R-CVP/R-CHOP (HMRN)

Overall survival

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for overall survival are presented in Figure 12 - **Figure *15***.

Figure 12: Smoothed hazard versus time for overall survival (R^2 weighted)

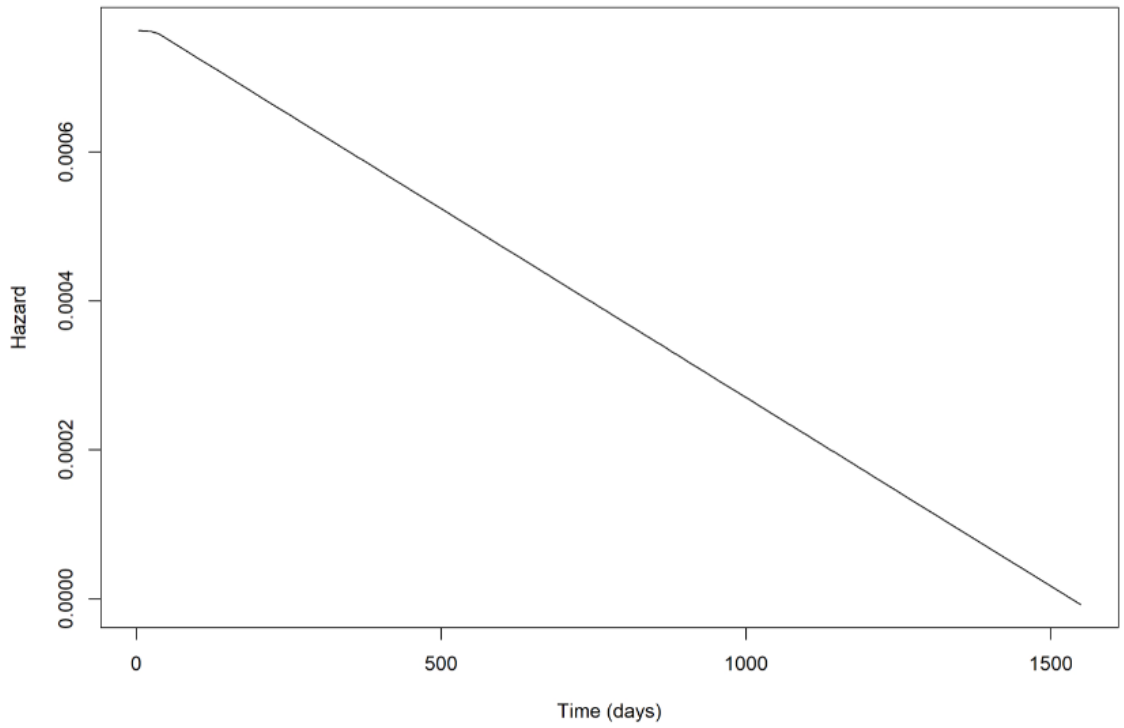


Figure 13: LN(smoothed hazard) versus time for overall survival (R^2 weighted)

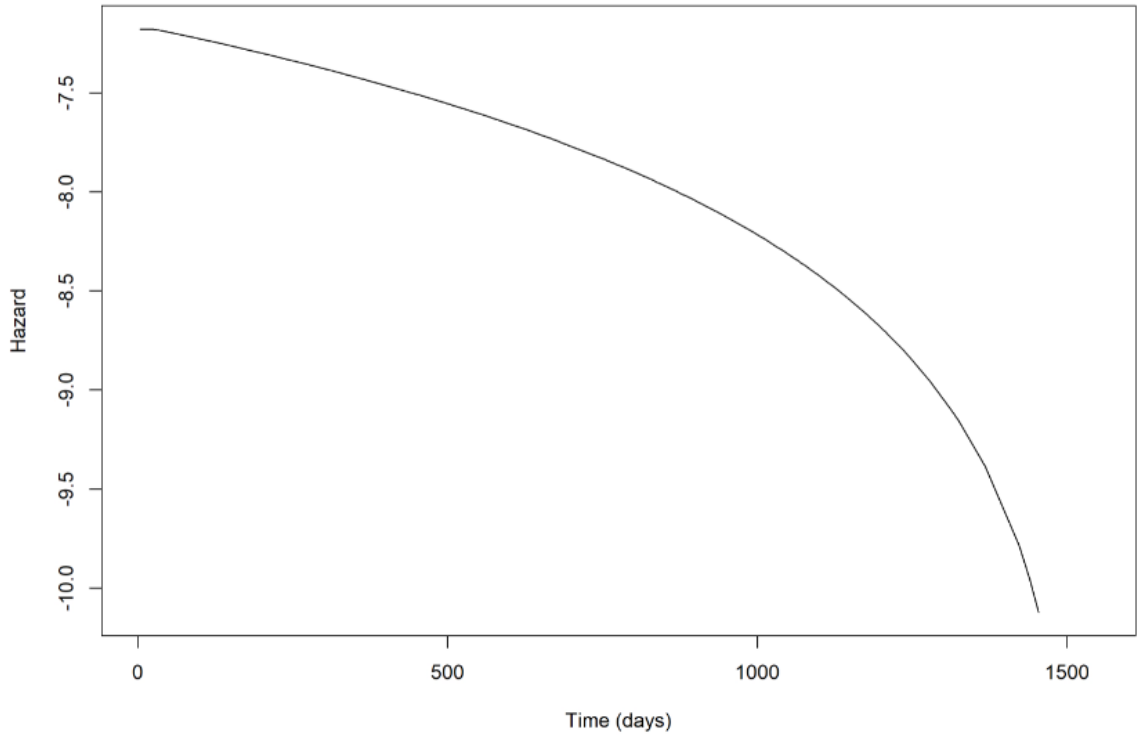


Figure 14: Smoothed hazard versus time for overall survival (R-CVP/RCHOP)

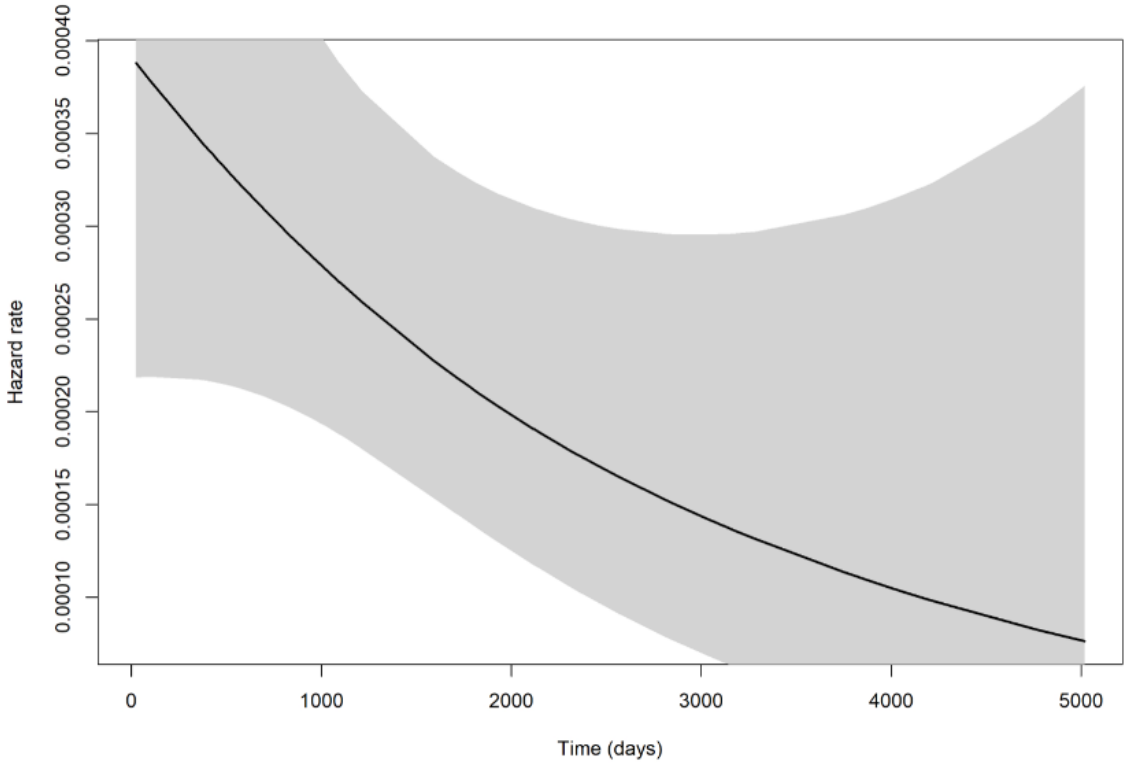
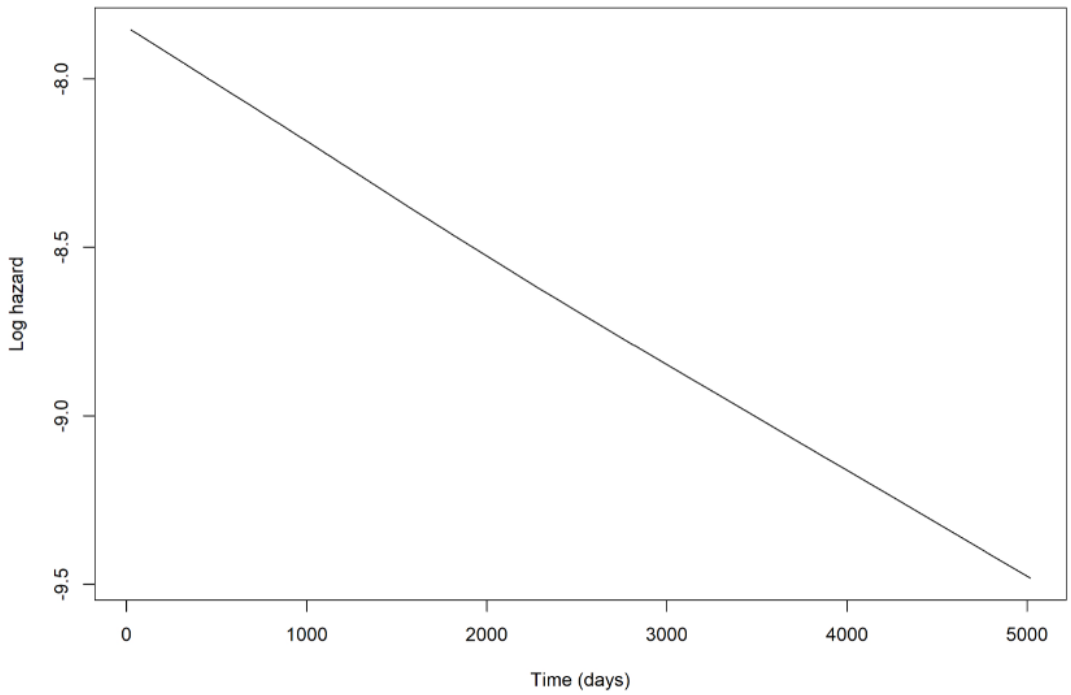


Figure 15: LN(smoothed hazard) versus time for overall survival (R-CVP/R-CHOP)



Progression-free survival

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for progression-free survival are presented in

Figure 16–**Figure *19***.

Figure 16: Smoothed hazard versus time (left) versus time for progression-free survival (R^2 weighted)

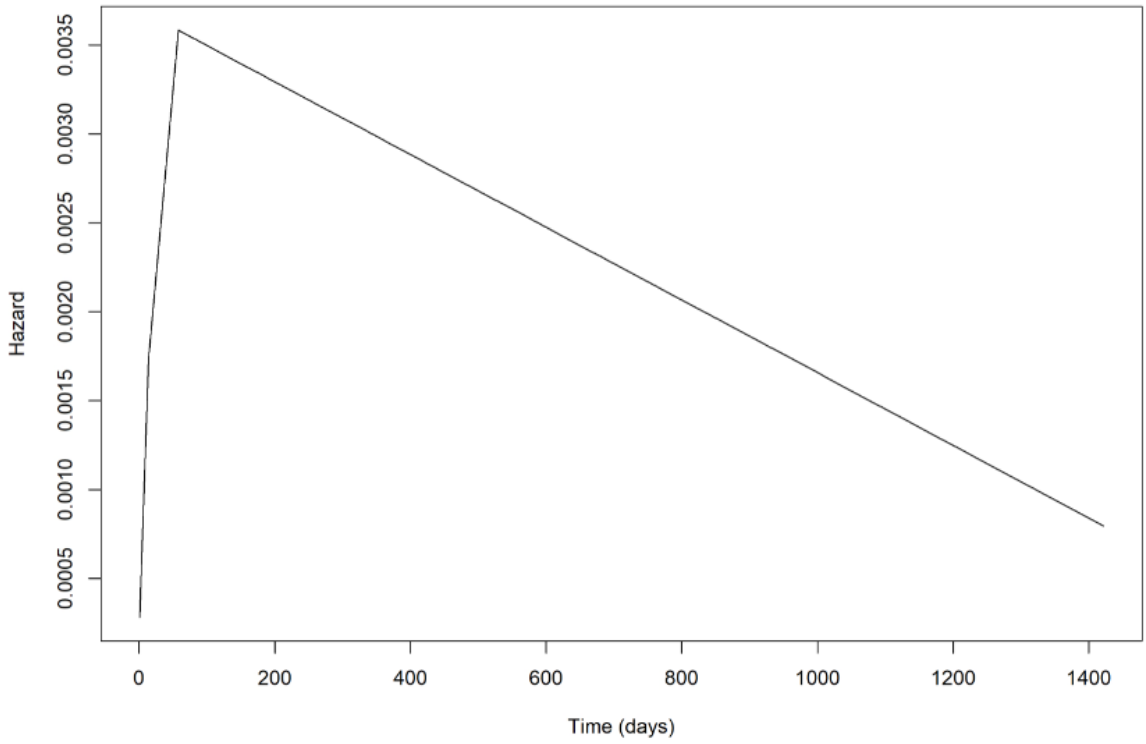


Figure 17: LN(smoothed hazard) versus time for progression-free survival (R^2 weighted)

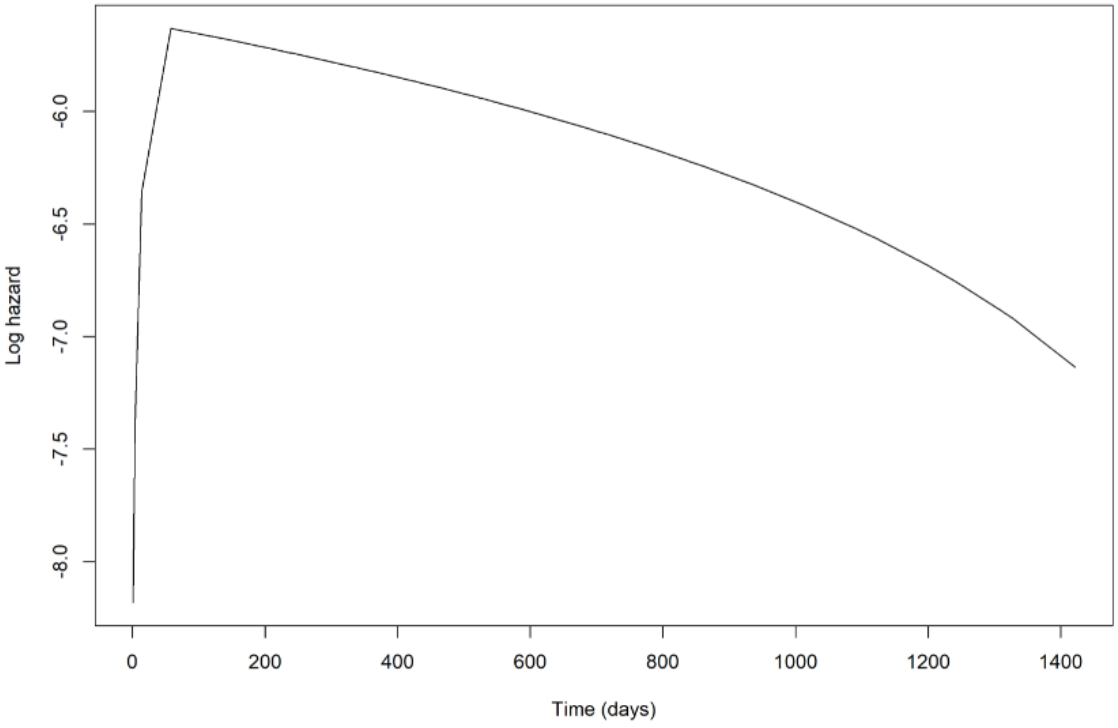


Figure 18: Smoothed hazard versus time for progression-free survival (R-CVP/RCHOP)

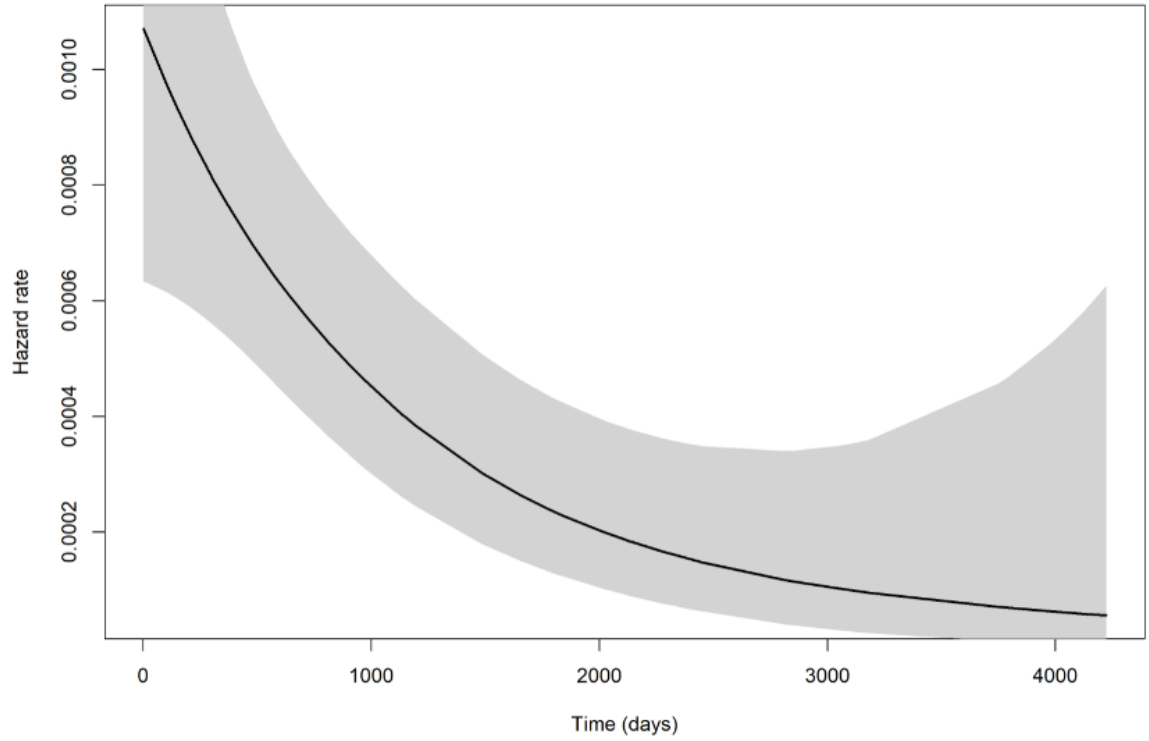
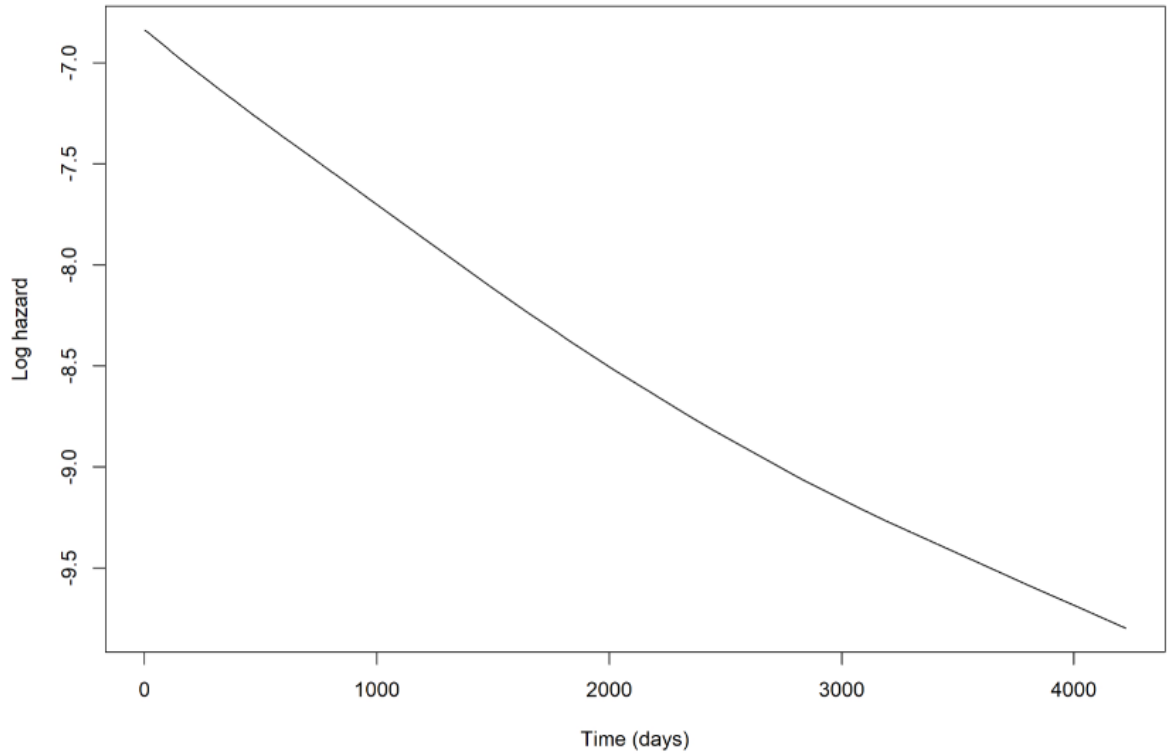


Figure 19: LN(smoothed hazard) versus time for progression-free survival (R-CVP/RCHOP)



Time to next anti-lyphoma treatment/death

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for time to next anti-lymphoma treatment are presented in Figure ***20***–Figure 23.

Figure 20: Smoothed hazard versus time for time to next antilymphoma treatment/death (R^2 weighted)

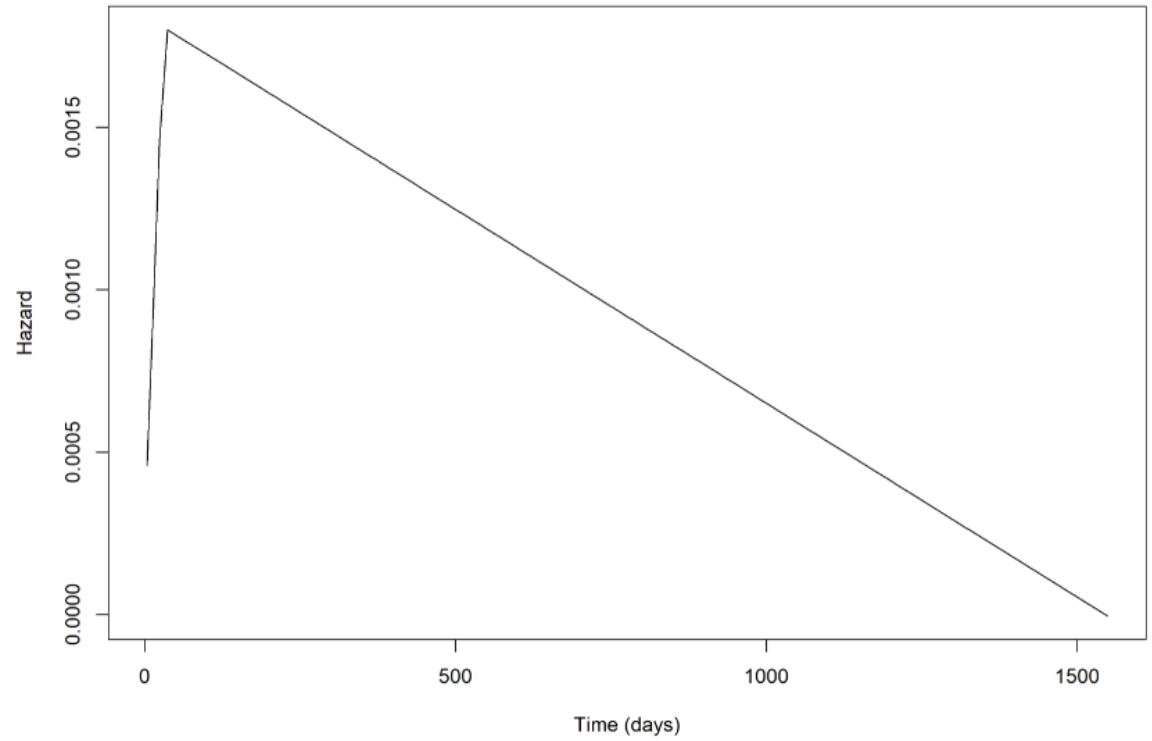


Figure 21: LN(smoothed hazard) versus time for time to next antilymphoma treatment/death (R^2 weighted)

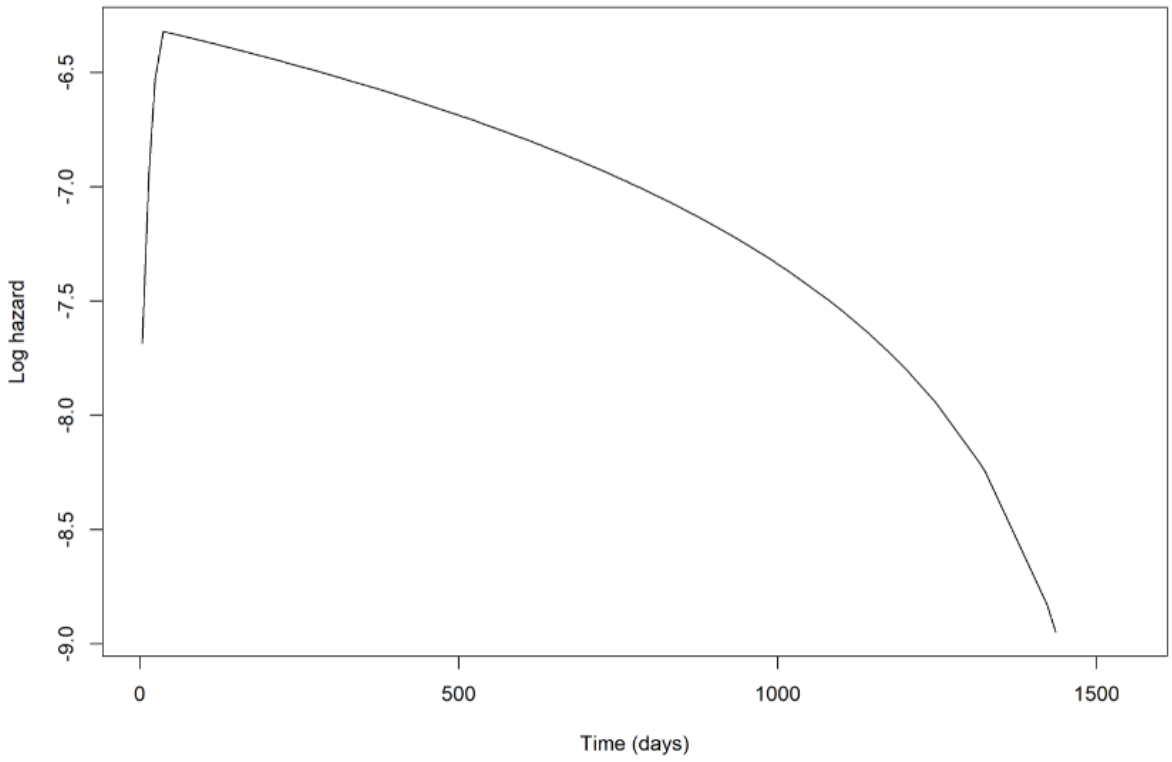


Figure 22: Smoothed hazard versus time for time to next antilymphoma treatment/death (R-CVP/RCHOP)

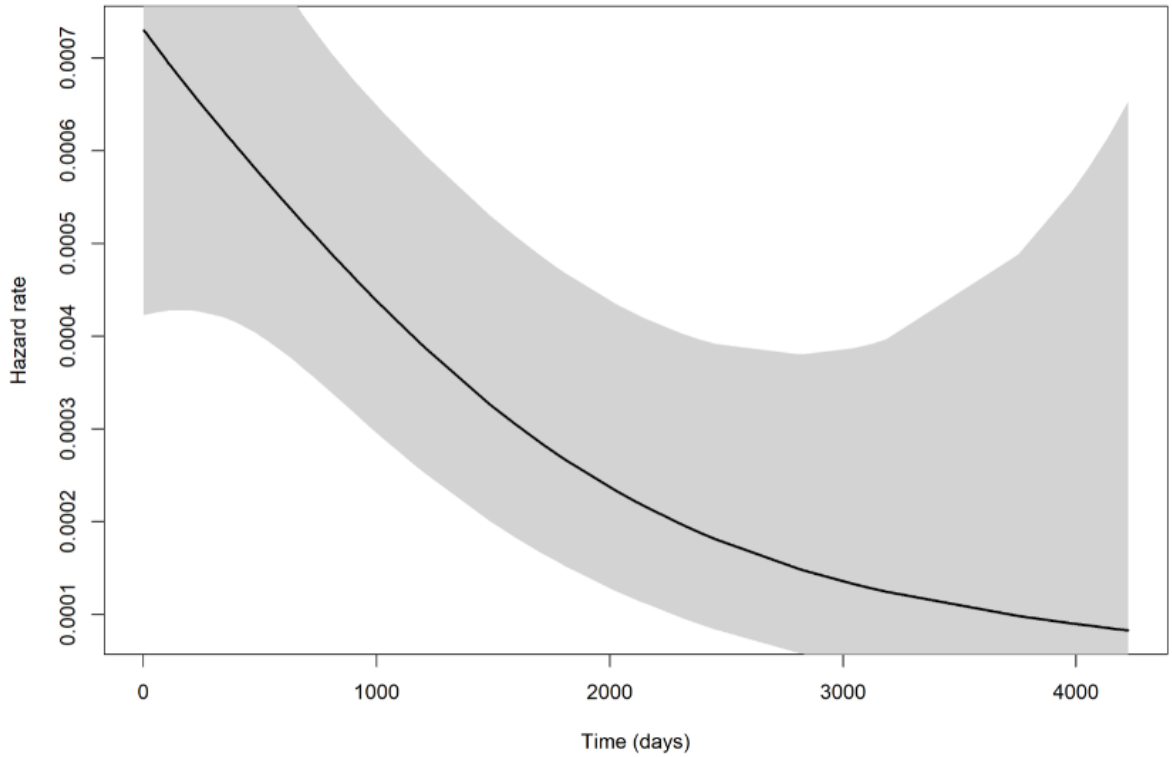
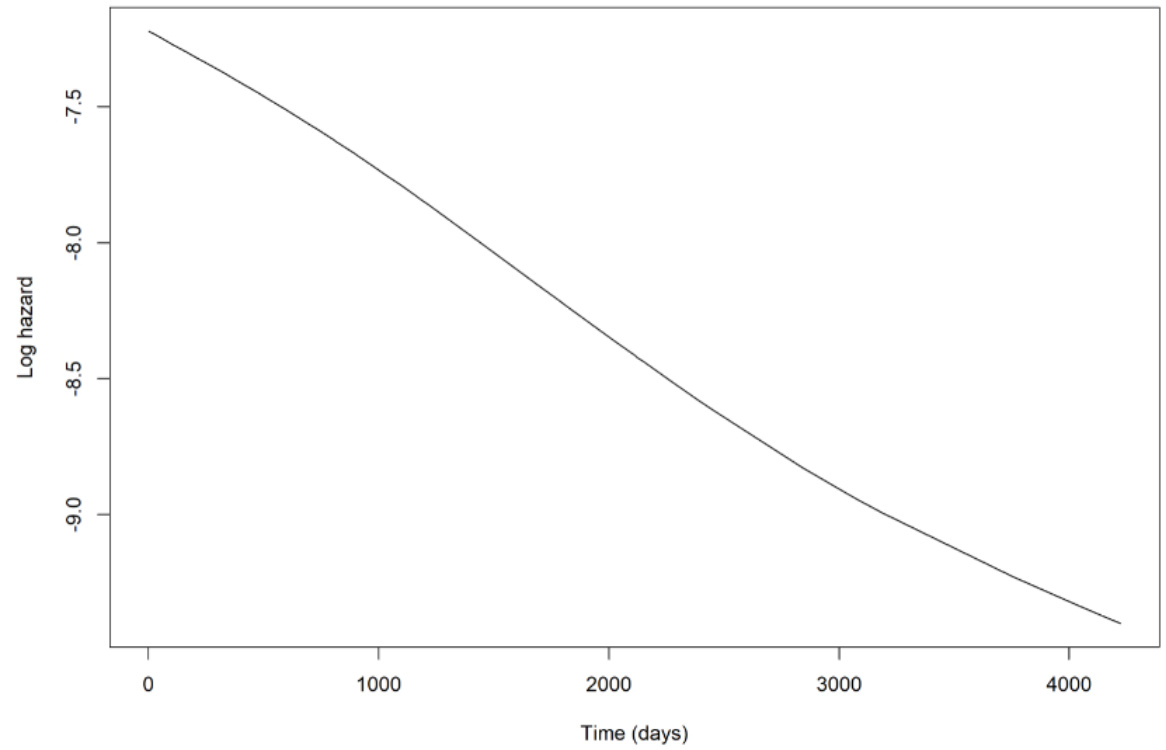


Figure 23: LN(smoothed hazard) versus time for time to next antilymphoma treatment/death (R-CVP/RCHOP)



R^2 (MAGNIFY) versus O-Benda (GADOLIN)

Overall survival

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for overall survival per treatment are presented in Figure 24 - Figure 27.

Figure 24: Smoothed hazard versus time for overall survival (R^2 weighted)

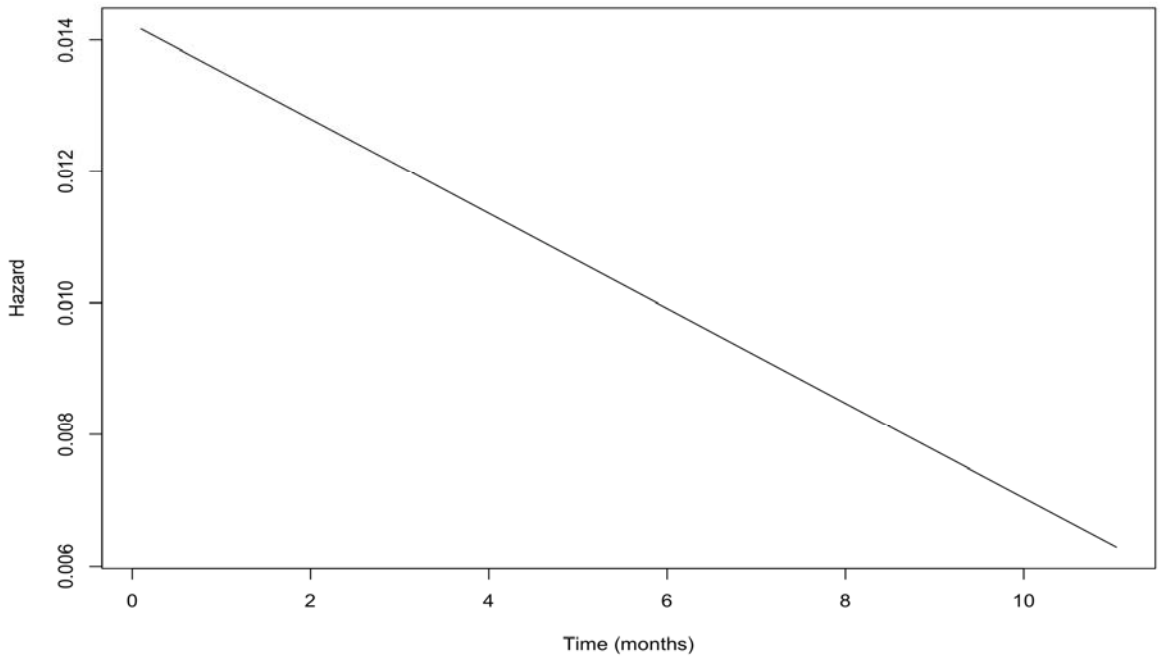


Figure 25: LN(smoothed hazard) versus time for overall survival (R^2 weighted)

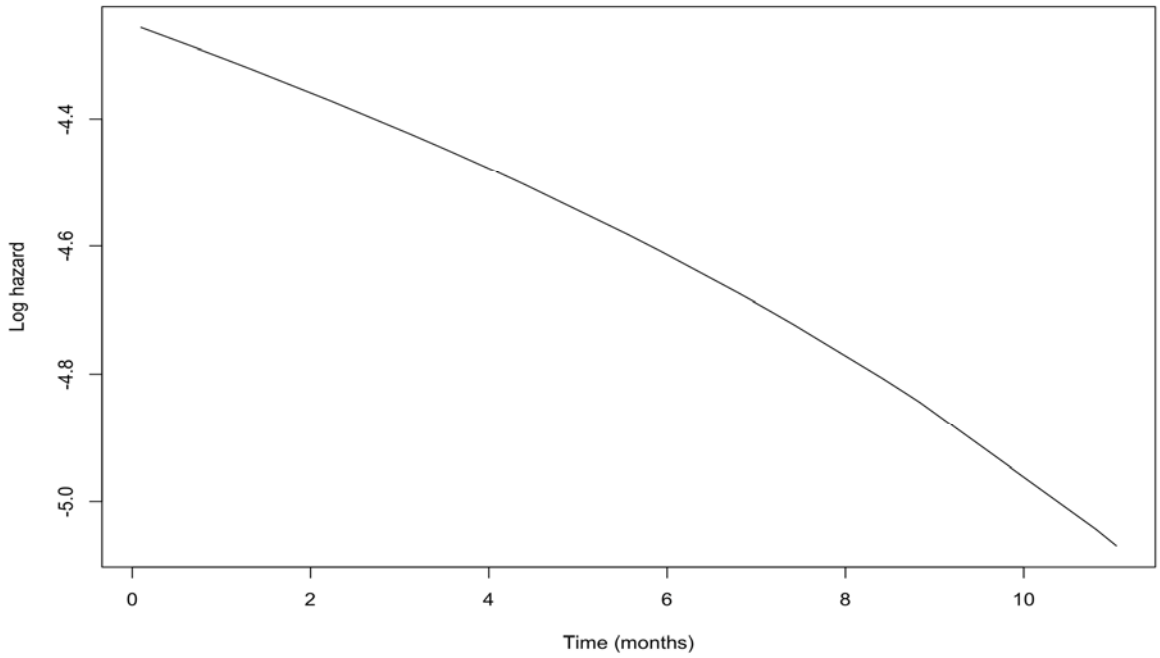


Figure 26: Smoothed hazard versus time for overall survival (O-Benda)

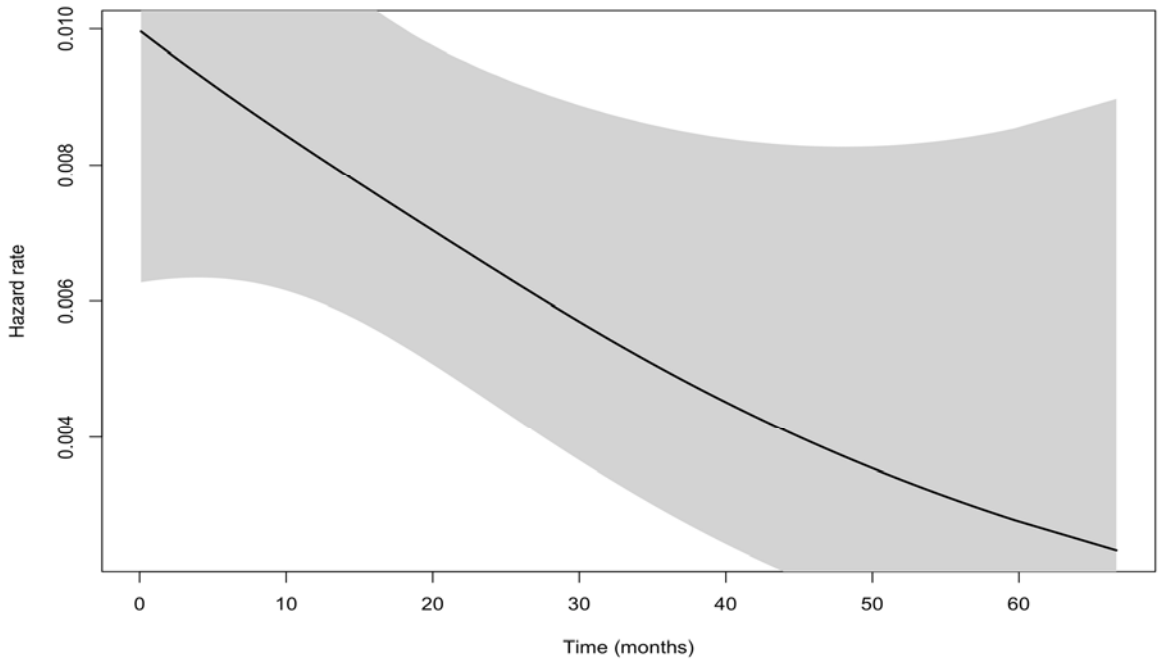
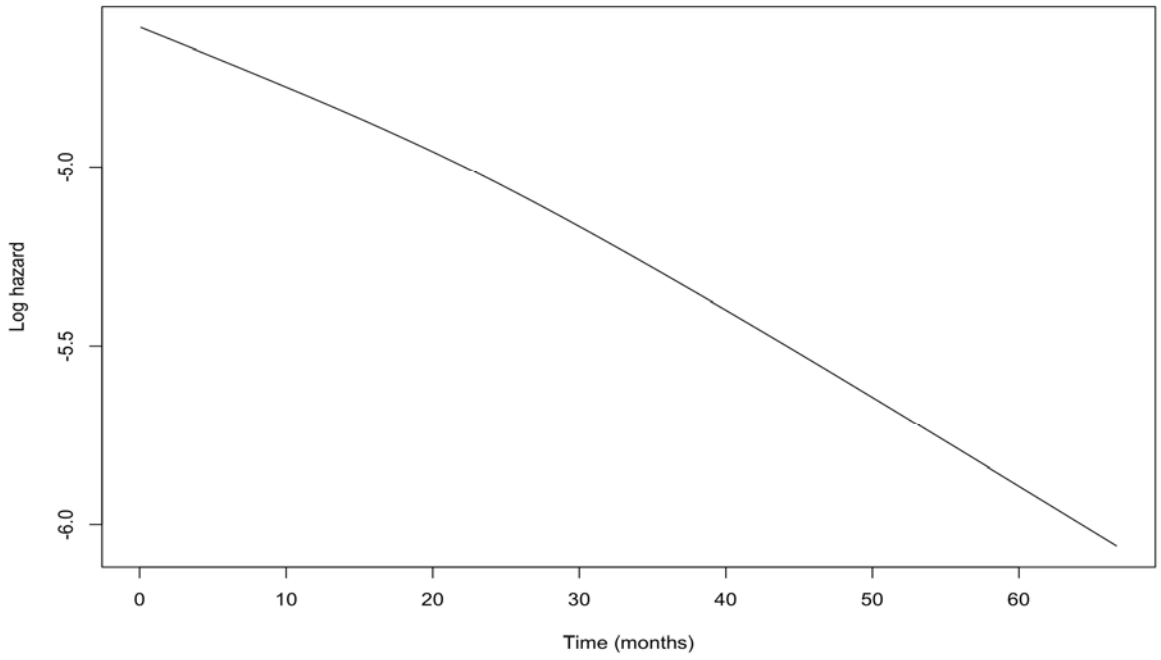


Figure 27: LN(smoothed hazard) versus time for overall survival (O-Benda)



Progression-free survival

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for progression-free survival are presented in Figure 28–Figure 31.

Figure 28: Smoothed hazard versus time for progression-free survival (R^2 weighted)

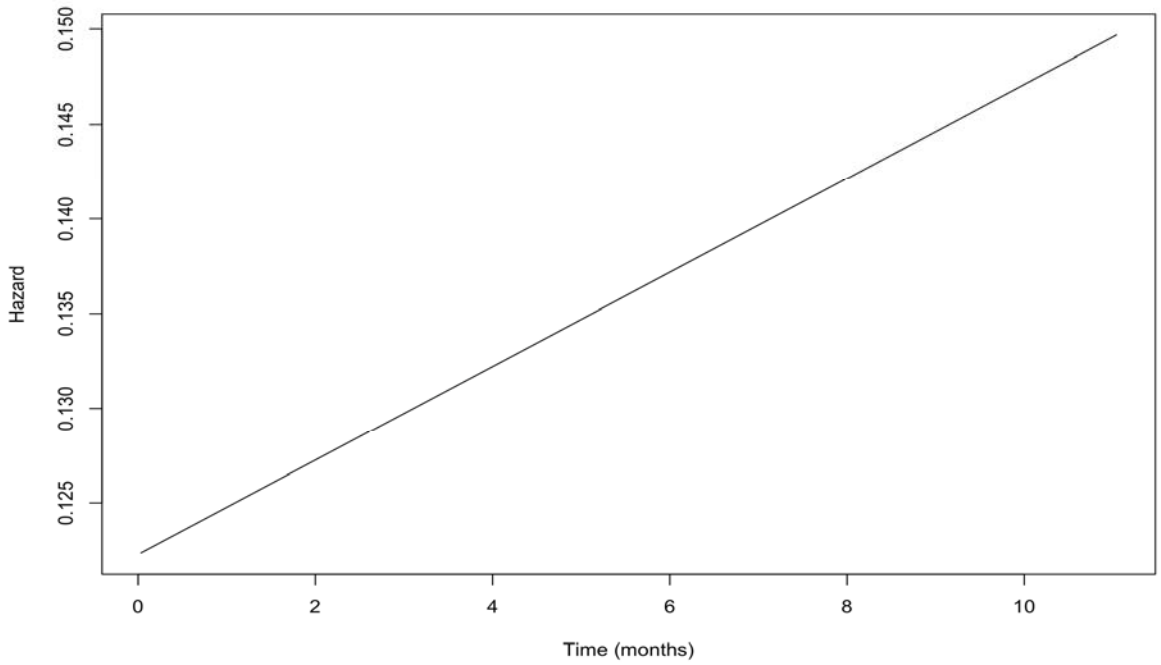


Figure 29: LN(smoothed hazard) versus time for progression-free survival (R^2 weighted)

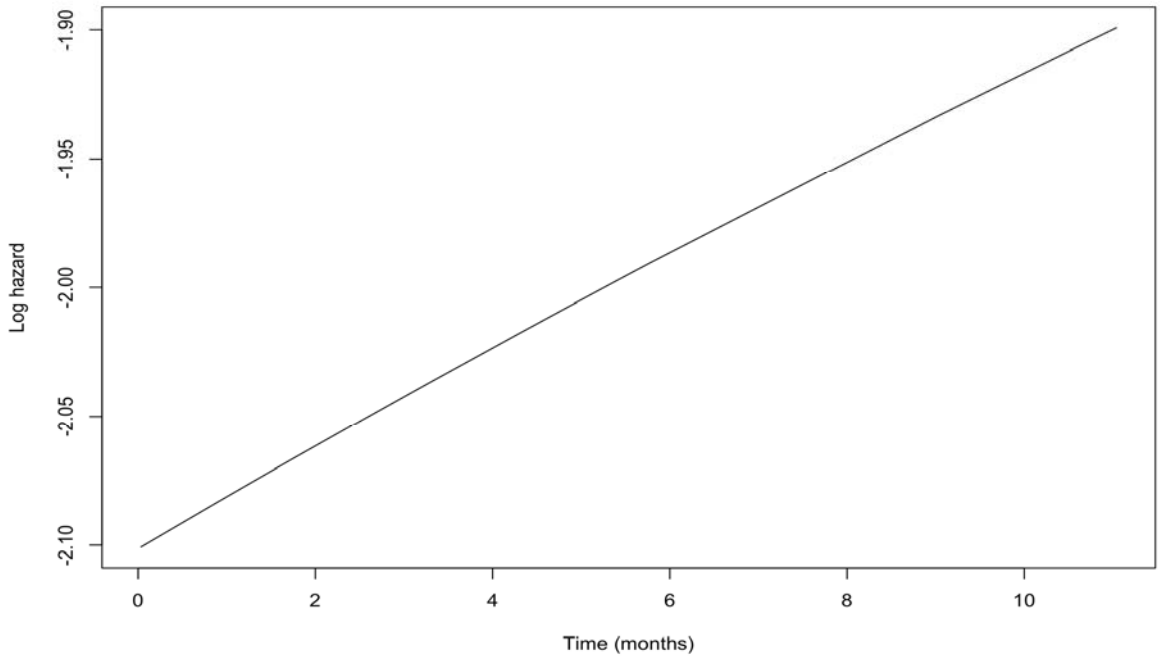


Figure 30: Smoothed hazard versus time for progression-free survival (O-Benda)

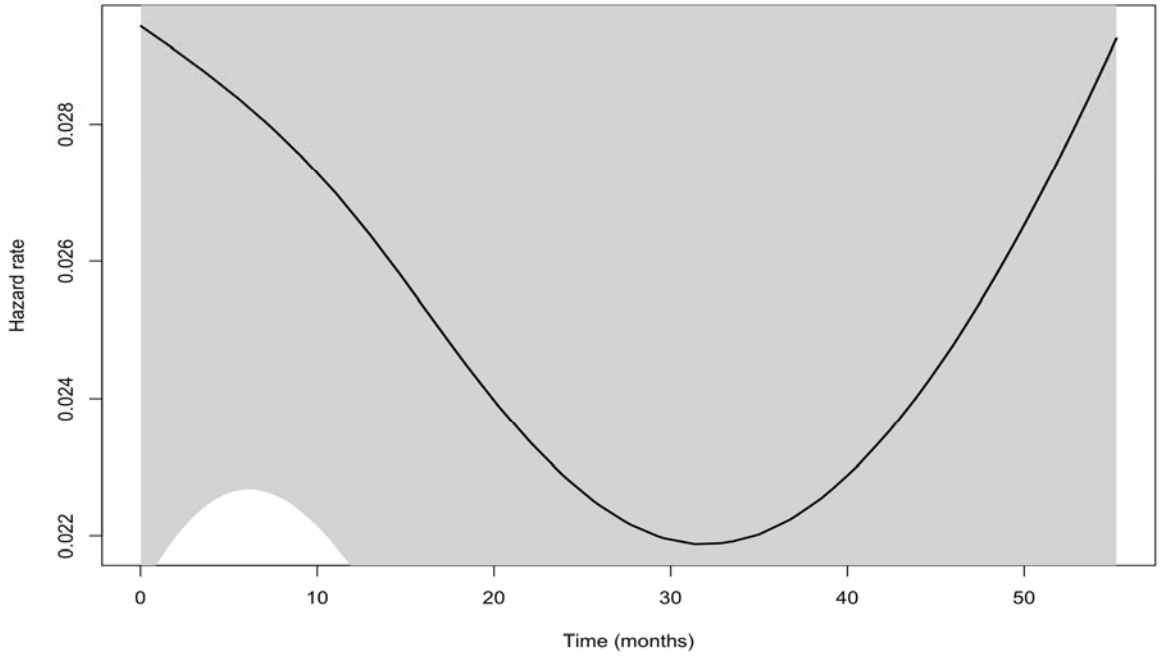
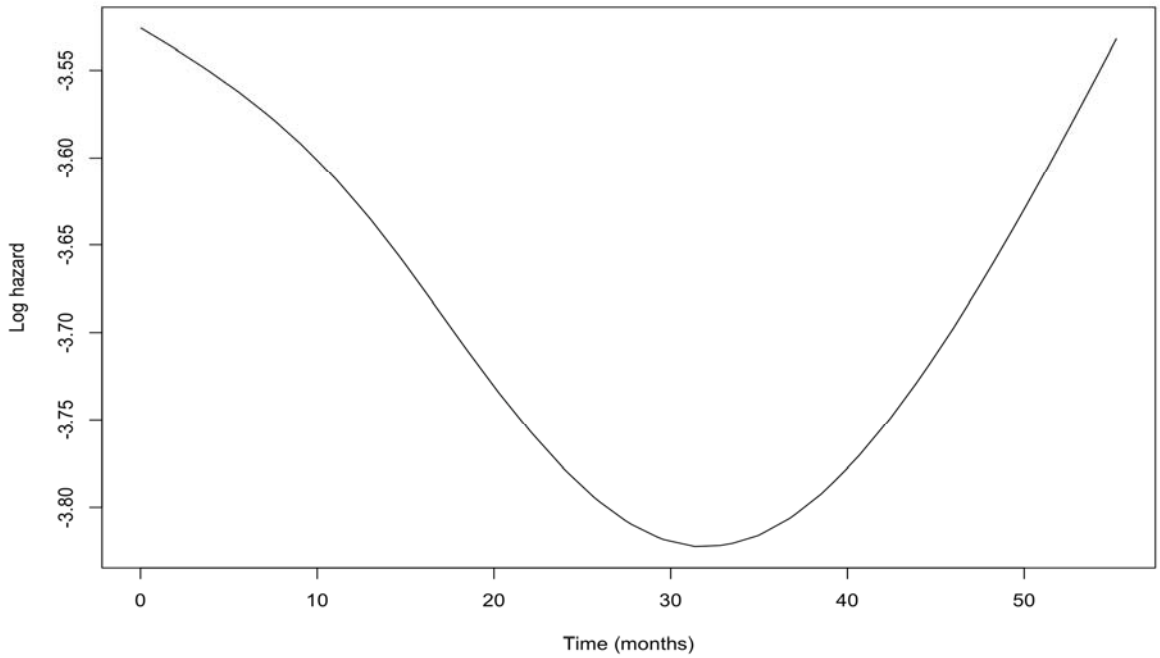


Figure 31: LN(smoothed hazard) versus time for progression-free survival (O-Benda)



Time to next anti-lymphoma treatment

The smoothed hazard versus time plots and LN(smoother hazard) over time plots for time to next anti-lymphoma treatment are presented in Figure 32 and Figure 33.

Figure 32: Smoothed hazard versus time for time to next anti-lymphoma treatment (O-Benda)

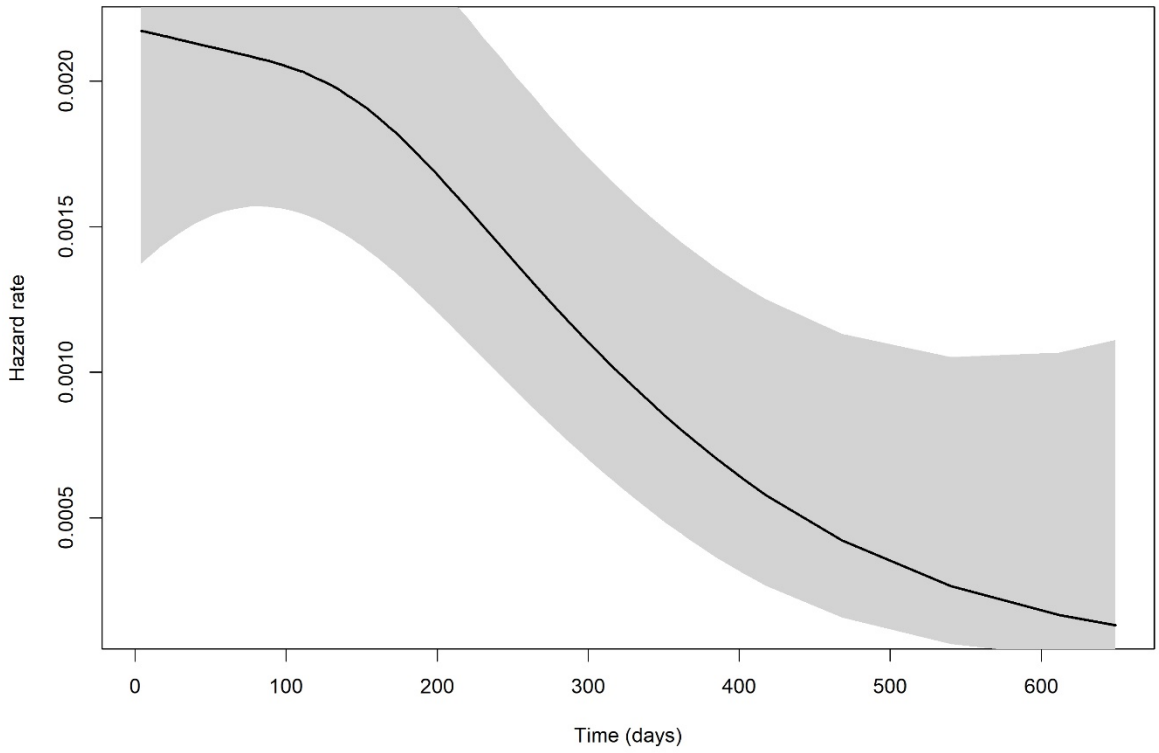
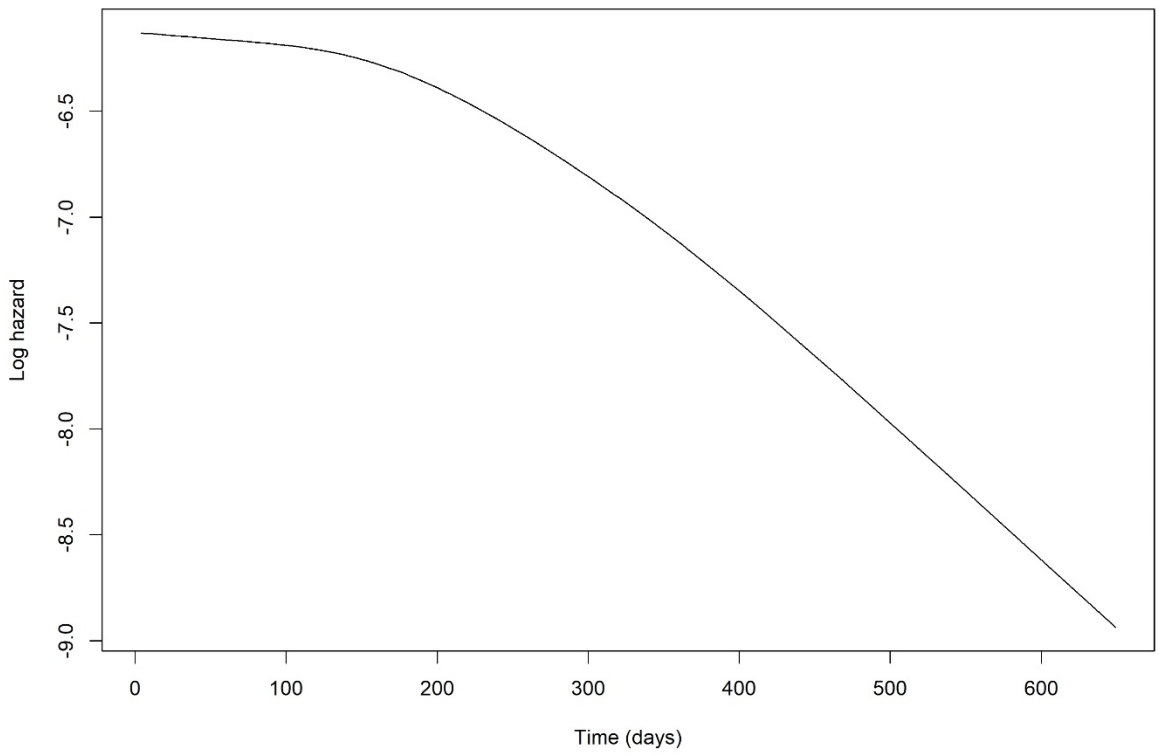


Figure 33: LN(smoothed hazard) versus time for time to next anti-lymphoma treatment (O-Benda)



Quality of life

B10. A mixed effects regression model, controlling for a selection of covariates, was used to estimate utility values for each health state in the cost effectiveness model. This approach resulted in different utility values for the R-refractory and non-R-refractory populations.

a. Please further comment on the plausibility of these utility differences between the R-refractory and non-R-refractory populations.

The utility analysis used the EQ-5D data that was collected in the AUGMENT trial using a mixed effects regression model. The overall utilities were dependant on; health state (progression-free, progressed off-treatment, progressed on-treatment), rituximab naïve status, whether patients were refractory to the last prior regimen, and number of previous treatments. These covariates were found to have significant impact on health-related quality of life. No utility data was available from MAGNIFY to produce a similar model for the R-refractory patients, however as the utility model used different patient characteristics relating to previous rituximab use and refractory status, these utilities could be used to estimate the utilities for the R-refractory population. Table 50 presents the results from the mixed regression model for each of the populations which shows a difference of 0.05 for each health state. These differences seem reasonable given the poorer outcomes expected for the R-refractory population. The R-refractory utilities are also consistent with those reported from GADOLIN which is a R-refractory population suggesting that the estimates for the R-refractory population are plausible.

Table 50: Utility values from the mixed regression model

Health state	Non R-refractory population	R-refractory population (vs O-Benda)	Incremental difference	GADOLIN utilities
Progression-free	0.863	0.814	0.049	0.822 (on treatment) 0.807 (off treatment)
Progressed (off treatment)	0.837	0.787	0.050	0.758
Progressed (on treatment)	0.808	0.758	0.050	

b. The SLR to identify relevant utility studies showed a wide range of utilities, especially the study of Pereira et al.[38] in which PF and PP utility values were much lower than the utility values used in the economic model. Please justify

why the utility values based on the AUGMENT trial would be representative for this submission.

According to NICE's Guide to the methods of technology appraisal 2013 there is a preference for using trial utility data (where available and appropriate), with EQ-5D being sourced from the literature only if not data is not available in the relevant clinical trials.[39] Therefore, in line with NICE's preferred measure for utility values, the utility values based on the AUGMENT trial are representative for this submission because (i) they are directly based on a relevant patient population for the decision problem and (ii) they allow quality of life to be measured according to the health states of the economic model using EQ-5D.

It is worth noting that the reasons the utilities from the study of Pereira et al. were not considered appropriate for this submission are detailed in Section B.3.4. of Document B, including:

- Small sample size (n=21)
- Lack of reporting on methods and patient characteristics
- Inconsistencies in the difference between PF and PP off-treatment utilities between the Pereira et al. study and AUGMENT, Wild et al. and GADOLIN.

Costs and resource use

B11. For costs, it was stated in Table 26 of the CS that "similar frequencies were used as per TA472". Please provide an overview of differences with TA472 and justify these differences.

A comparison of the disease monitoring resource use frequencies used in the company submission and TA472 are given in Table 51. The monthly frequencies are similar across the health state splits, with equivalent monthly frequencies of haematologist visits, diagnostic tests and CT scans given per split progression-free period and during the progressed period. Both the company submission and TA472 base their post-induction treatment monitoring frequencies on ESMO guidelines. [31]

Following the completion of induction therapy, the frequency of the follow-up visits reduce to every 3–4 months based on the ESMO guidelines.[31] TA472 assumed the 3 month based frequencies apply post-induction treatment from 6 to 30 months and the 4 month based frequencies apply from 30 months until progression. The ERG for the TA472 submission noted the assumption that the monitoring at the end of induction treatment with bendamustine monotherapy to be the same as at the start of maintenance treatment with obinutuzumab monotherapy was unclear.

The company submission has utilised the reduction in frequency of follow-up visits from 3–4 months to inform an assumed decrease in monitoring between the maintenance treatment period and post-maintenance follow-up period. The company submission hence assumes the 3 month-based frequencies apply to the maintenance phase and the 4 month-based frequencies apply to the post-maintenance phase.

As R^2 patients received no maintenance treatment following the induction period, no maintenance phase monitoring costs were applied in this treatment arm (Table 52). Furthermore, as detailed in Section B.3.5 of the company submission (summarised in Table 52), differences in disease monitoring resource use can occur where there are differences in the length of the three defined periods for the progression-free resource use outcomes. This is because the duration of induction therapy and whether it is followed by maintenance treatment is dependent on the treatment arm.

Table 51: Disease monitoring resource use frequencies

Item		Company submission				TA472			
		Progression-free monthly frequency			Progressed monthly frequency	Progression-free monthly frequency			Progressed monthly frequency
		Induction	Maintenance	Follow-up		Induction (0-6 months)	Follow-up (6-30 months)	Follow-up (30 months until progression)	
Haematologist led		1	0.33	0.25	1	1	0.33	0.25	1
Diagnostic tests	FBC	1	0.33	0.25	1	1	0.33	0.25	1
	patient history/physical exam								
	Full profile (U&E, LFT, calcium)								
	Serum IgG, IgA, IgM and electrophoresis								
	LDH test								
CT scans		0.17	0.04	0	0	0.17	0.04	0	0

Key: CT, computed tomography; FBC, full blood count; Ig, immunoglobin; LDH, lactate dehydrogenase; LFT, liver function tests; NHS, National Health Service; TA, technology appraisal; U&E, urea and electrolytes.

Table 52: Overview of differences in the durations of monitoring periods for resource use frequencies

	Progression-free monitoring			Progressed monitoring
	Induction	Maintenance	Follow-up	
TA472	0-6 months	6-30 months	30 months until progression	Post progression
Company submission: R²	0-12 months (progression-free on treatment patients)	N/A	0 months – progression (progression-free off treatment patients)	Post progression
Company submission: R-CHOP	0-6 months (progression-free on treatment patients)	6-33* months (progression-free on treatment patients)	0 months – progression (progression-free off treatment patients)	Post progression
Company submission: R-CVP	0-6 months (progression-free on treatment patients)	6-33* months (progression-free on treatment patients)	0 months – progression (progression-free off treatment patients)	Post progression
Company submission: O-benda	0-6months (progression-free on treatment patients)	6-32** months (progression-free on treatment patients)	0 months – progression (progression-free off treatment patients)	Post progression
Notes: *Rituximab maintenance is given to patients 3 months after the induction dose for a maximum of 2 years; **Bendamustine maintenance is given to patients 2 months after the induction dose for a maximum of 2 years.				

B12. Subsequent treatments were included in the model as an average cost per patient which was applied as a one-off cost to those patients entering the PP (on-treatment) health state. Please justify why the duration of subsequent treatments was not considered in the economic model, given the potential impact on costs. Please also justify this for any comparators added in response to question A7.

Subsequent treatments were applied as a one-off cost calculated based on the mean duration using real-world data from HMRN. This was considered appropriate for this model structure given the time patients spend in the PP (on-treatment) health state is similar between both arms based on clinical opinion and selection of base case curves (see response to clarification question B9a), and therefore, accounting for the different durations of subsequent treatments would not impact results. As discussed in response to clarification question B.9a, part of the criteria for selecting the curve choices

was to minimise the difference between mean LYs between receiving subsequent treatment and death. Furthermore, while there is some evidence to suggest R^2 may resensitize patients to subsequent therapy[40] , the clinical experts did not believe patients who received R^2 would be treated any more intensively (and so there is no reason to expect a difference in subsequent treatment costs). The model base case used the relevant subsequent treatment data to inform the costs for subsequent treatments in line with the efficacy data where possible with scenarios exploring the impact of assuming the same subsequent treatment costs. These scenarios had minor impacts on the base case ICER (see Tables 68, 69 and 70 in the company submission, Document B).

B13. Fit and young (less than 70 years) patients who relapsed early but were not refractory to induction therapy were considered for consolidation ASCT. In the R-CHOP comparator, 11.76% of the patients received ASCT, given that R-CHOP may be used as an induction therapy in ASCT candidates. ASCT was not given to the R2 population in the economic model, based on clinical experts and because it was not offered to patients in the AUGMENT protocol. Please clarify why ASCT would not be considered as a subsequent treatment in UK clinical practice in patients receiving R2 who meet the above conditions.

The AUGMENT trial was not designed to assess the impact of R^2 as an induction treatment for patients ahead of undergoing stem-cell transplant, therefore no costs for stem-cell transplant post induction/pre progression were considered for R^2 such that the efficacy was in line with the costs in the model. Two patients in the AUGMENT trial did receive stem-cell transplant subsequently, however this was administered as a 3rd subsequent regimen and thus included within the costs for subsequent treatment assigned to the R^2 arm.

During the clinical validation meeting, when discussions were had regarding the clinical pathway including R^2, clinicians felt that it was unlikely that R-maintenance or stem-cell transplant would be offered after R^2 due to the lack of evidence of R^2 used as an induction to stem-cell transplant. This was in line with the efficacy available for R^2 therefore no further scenarios or assumptions were required for the model.

B14. Subsequent treatment costs were assumed to be identical for O-Benda, R-CHOP and R-CVP (as reported in CS Table 62).

a. Please justify why assuming identical subsequent treatment costs for O-Benda as for R-CHOP/R-CVP is appropriate (rather than for instance assuming identical costs as for R2 from AUGMENT).

There were no subsequent treatment data reported from GADOLIN or for a specific R-refractory population, therefore, for the O-Benda comparison, the subsequent treatments were identical to the treatments used in HMRN using the non-R-refractory population. This assumes that the subsequent

treatments given to a non-R-refractory population would apply to a R-refractory population which was validated by clinicians. The model used the HMRN subsequent treatments as these were based on real-world data from UK specific clinical practice and therefore more reflective of UK costs. Using the same subsequent treatment cost identical to R^2 has little impact on the ICER and is presented as a scenario in Document B, Section B.3.8 (see Table 53).

Table 53: Scenario assuming same subsequent treatment costs between O-Benda and R^2

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
O-Benda	[redacted]	[redacted]	[redacted]				
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,442,218 (SW)
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; R2, lenalidomide plus rituximab.							

b. Please provide a scenario analysis (and the accompanying model) using R-CHOP and R-CVP specific subsequent treatment costs (estimated based on the HMRN database).

The proportion of patients who received each subsequent treatment when using R-CHOP and R-CVP specific data based on the HMRN database are given in

Table ***54***. These proportions of patients were combined with the mean durations, total treatment costs and total administration costs as given in Table 62 of Document B of the company submission (Section B.3.5, p205) to determine total weighted costs specific to R-CHOP and R-CVP. For this scenario analysis, the R-CHOP and R-CVP specific subsequent treatment costs were applied as the comparator subsequent treatment cost for the R-CHOP and R-CVP treatment arms, respectively.

The pairwise results of the scenario analyses which assume R-CHOP and R-CVP specific subsequent treatments (estimated based on the HMRN database) are presented in Table 55. R^2 remains cost-effective versus R-CHOP and R-CVP at the £30,000 willingness-to-pay (WTP) threshold with little impact to the ICERs versus the base case results of the company submission.

It is worth noting that for R-CHOP and R-CVP at base case, data from HMRN using the total subsequent treatment data from the pooled R-chemotherapies has been used in order to draw from a larger sample size.

Table 54: Scenario analysis: R-CHOP and R-CVP specific subsequent treatments

Subsequent treatment	R-CHOP n (%) n=34	R-CVP n (%) n=33
R-mono		
R-Benda		
R-CHOP		
R-CVP		
Other R-chemo		
O-Benda		
Bendamustine		
Other chemotherapy		
Targeted therapies		
Radiotherapy		
Other		
ASCT		
Total weighted treatment cost		
Total weighted administration cost		

Table 55: Scenario analysis pairwise results: R-CHOP and R-CVP specific subsequent treatments

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALYs)
R^2 versus R-CHOP							
R-CHOP							
R^2							£11,829
R^2 versus R-CVP							
R-CVP							
R^2							£17,939
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; R2, lenalidomide plus rituximab.							

Scenario and sensitivity analyses

B15. Based on the model file, all efficacy parameters seem to be included in the probabilistic sensitivity analysis (PSA). However, these parameters and their

distributions were not reported in Table 94 of Appendix S. Please provide further details on how these parameters were included in the PSA.

All efficacy parameters were varied in PSA using the multivariate normal distribution. Each variance covariance matrices used to vary the base case parameters are presented at the end of this document in the Appendix A.

B16. On page 218, section B.3.8 it is stated that "Figure 57 – Figure 59 present the tornado diagrams showing the parameters with the greatest impact on the results with descending sensitivity from one-way sensitivity analysis (OWSA), when their values were set to their upper and lower limits of the confidence intervals reported in Section B.3.6." However, section B.3.6 does not contain information on limits of confidence intervals of model parameters. Please clarify for which parameters OWSA was performed, which values were used, and how these values were determined.

All parameters used in OWSA and PSA, lower and upper limits and distributions are presented in Table 19 of Appendix S. A copy of this table is presented at the end of this document in the Appendix B leaving only the parameters which are used in OWSA. Information of how these values were derived are presented in the relevant sections detailed in Table 56 in Appendix B of this document. If statistical uncertainty was unavailable for certain parameters, then the standard error was assumed to be 20% of the mean.

Validation and transparency

B17. As stated in the CS, the company held an advisory board to validate several key aspects of the model. However, it was unclear which specific validation efforts were undertaken and what the results were.

a. Please provide the full report of reference 133 from Document B: "Celgene. UK Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Draft meeting report) 13 March 2019. (Updated: 13 May 2019) Data on file".

Please see response to question A6 – final advisory board report has been provided.

b. Please provide further details on the validation efforts that were undertaken, using the AdViSHE tool [41].

Please see separate document

c. Please provide detailed results of the validation efforts using the AdViSHE tool[41].

Please see separate document

REFERENCES:

1. National Instituite for Health and Care Excellence (NICE). Single technology appraisal: User guide for company evidence submission template. 2015. Available at: https://www.nice.org.uk/process/pmg24/resources/single-technology-appraisal-user-guide-for-company-evidence-submission-template-pdf-72286715419333. Accessed: August 2019.
2. National Institute for Health and Care Excellence (NICE) TA472 Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab. 2017. Available at: https://www.nice.org.uk/guidance/ta472. Accessed: 05 January 2019.
3. National Institute for Health and Care Excellence (NICE). TA137: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. 2008. Available at: https://www.nice.org.uk/guidance/ta137. Accessed: February 2019.
4. National Institute for Health and Care Excellence (NICE). ID1379: Idelalisib for treating follicular lymphoma refractory to 2 treatments. 2018. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta10290. Accessed: 25 April 2019.
5. Andorsky DJ, Kolibaba KS, Assouline S, et al. An open-label phase 2 trial of entospletinib in indolent non-Hodgkin lymphoma and mantle cell lymphoma. 2019; 184(2):215-22.
6. de Vos S, Swinnen LJ, Wang D, et al. Venetoclax, bendamustine, and rituximab in patients with relapsed or refractory NHL: a phase Ib dose-finding study. *Ann Oncol*. 2018; 29(9):1932-8.
7. Witzig TE, Wiernik PH, Moore T, et al. Lenalidomide oral monotherapy produces durable responses in relapsed or refractory indolent non-Hodgkin's Lymphoma. *J Clin Oncol*. 2009; 27(32):5404-9.
8. 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*. 2017; 28(9):2169-78.
9. Dreyling M SA, Mollica L, Leppä S, Follows GA, Lenz G, Kim W, et al. Copanlisib in patients with relapsed or refractory indolent b-cell lymphoma (CHRONOS-1). *Haematological Oncology*. 2017; 35(S2):119; Abstract 08.
10. Flinn IW, Kahl BS, Leonard JP, et al. Idelalisib, a selective inhibitor of phosphatidylinositol 3-kinase-delta, as therapy for previously treated indolent non-Hodgkin lymphoma. *Blood*. 2014; 123(22):3406-13.
11. Gopal AK, Kahl BS, de Vos S, et al. PI3Kδ Inhibition by Idelalisib in Patients with Relapsed Indolent Lymphoma. *New England Journal of Medicine*. 2014; 370(11):1008-18.
12. Kahl BS, Bartlett NL, Leonard JP, et al. Bendamustine is effective therapy in patients with rituximab-refractory, indolent B-cell non-Hodgkin lymphoma: results from a Multicenter Study. *Cancer*. 2010; 116(1):106-14.

13. Rummel MJ, Al-Batran SE, Kim SZ, 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.
14. GOV.UK. Drugs and pharmaceutical electronic market information tool (eMIT). 2019. Available at: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. Accessed: 25 April 2019.
15. 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*. 2015; 33(23):2516-22.
16. 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*. 2016; 17(8):1081-93.
17. Chong EA, Ahmadi T, Aqui NA, 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.
18. Morschhauser F, Fowler NH, Feugier P, et al. Rituximab plus Lenalidomide in Advanced Untreated Follicular Lymphoma. *New England Journal of Medicine*. 2018; 379(10):934-47.
19. Chiu H, Trisal P, Bjorklund C, et al. Combination lenalidomide-rituximab immunotherapy activates anti-tumour immunity and induces tumour cell death by complementary mechanisms of action in follicular lymphoma. *Br J Haematol*. 2019; 185(2):240-53.
20. Celgene. Draft Summary of Product Characteristics (SPC). Revlimid® (Lenalidomide). 2019. Accessed: 19/02/2019.
21. Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0069) March 2019 2019. (Updated: June 2019)
22. Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007 (updated tables)) 2018.
23. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – Demographic and Baseline Characteristics for IEE population. 04 June 2019 2018.
24. van Oers MH, Klasa R, Marcus RE, 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.
25. HMRN. Haematological Malignancy Research Network: Clinical Management and Outcome of Follicular Lymphoma (FL) with a Focus on Relapsed/Refractory Disease (draft report). 17 July 2019 2019.
26. 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. *J Clin Oncol*. 2019; 37(14):1188-99.
27. Guyot P, Ades AE, Ouwens MJ and Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012; 12:9.
28. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Stat Med*. 2016; 35(30):5642-55.

29. Phillippo D, Ades A, Dias S, et al. NICE DSU Technical Support Document 18: Methods for Population-adjusted Indirect Comparisons in Submissions to NICE. 2016. Available at: https://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf. Accessed: 02 May 2019.
30. Prica A, Chan KKW and Cheung MC. Frontline rituximab monotherapy induction versus a watch and wait approach for asymptomatic advanced stage follicular lymphoma: A cost-effectiveness analysis. *Blood*. 2013; 122(21).
31. 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*. 2016; 27(suppl 5):v83-v90.
32. Dreyling M, Thieblemont C, Gallamini A, et al. ESMO Consensus conferences: guidelines on malignant lymphoma. part 2: marginal zone lymphoma, mantle cell lymphoma, peripheral T-cell lymphoma. *Ann Oncol*. 2013; 24(4):857-77.
33. Celgene. LENALIDOMIDE (CC-5013-NHL-007) – Patient reported outcome report. 27 Nov 2018 2018.
34. Kang D, Cho J, Kim IR, et al. Health-Related Quality of Life in Non-Hodgkin Lymphoma Survivors: A Prospective Cohort Study. *Cancer Res Treat*. 2018; 50(4):1051-63.
35. Oerlemans S, Issa DE, van den Broek EC, et al. Impact of therapy and disease-related symptoms on health-related quality of life in patients with follicular lymphoma: results of the population-based PHAROS-registry. *Eur J Haematol*. 2014; 93(3):229-38.
36. Woods B, Sideris E, Palmer S, et al. NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care: a Critical review. 2017. Available at: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf. Accessed: 28 May 2019.
37. National Institute for Health and Care Excellence (NICE) TA513 Obinutuzumab for untreated advanced follicular lymphoma. 2018. Available at: https://www.nice.org.uk/guidance/ta513. Accessed: 05 January 2019.
38. Pereira C, Negreiro F and Silva C. PSY43 Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. *Value in Health*. 2010; 13(7):A468.
39. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013. Available at: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed: 01 June 2019.
40. Gribben J, Trneny M, Izutsu K, et al. AUGMENT: RELAPSED/REFRACTORY INDOLENT NHL PATIENTS WERE MORE SENSITIVE TO NEXT TREATMENT FOLLOWING LENALIDOMIDE/RITUXIMAB (R2) THAN RITUXIMAB/PLACEBO. *Hematological Oncology*. 2019; 37(S2):227-9.
41. Rafia R, Papaioannou D, Stevenson M, et al. Evaluating the cost-effectiveness of the addition of rituximab to chemotherapy in the first line treatment of follicular lymphoma patients in the UK. *Value Health*. 2011; 14(7):A445.

Appendix A: Variance-covariance matrices for efficacy parameters

R^2 vs R-CHOP/CVP

Efficacy	Distribution	Parameter	Base case	Variance-covariance matrices		
PFS	Generalised gamma			**Mu**	**Sigma**	**Q**
		Mu	8.0454	0.156	-0.071	0.175
		Sigma	0.4190	-0.071	0.099	-0.173
		Q	1.4838	0.175	-0.173	0.384
OS	Weibull (R^2)			**Shape**	**Scale**	
		Shape	0.123	0.074	-0.135	
		Scale	8.665	-0.135	0.312	
	Weibull (R-CHOP/CVP)			**Shape**	**Scale**	
		Shape	-0.139	0.0242	-0.0130	
		Scale	8.386	-0.0130	0.0482	
TTNLT	Log-normal (R^2)			**MeanLog**	**SDLog**	
		Meanlog	7.703	7.703	0.078	
		SDLog	0.457	0.457	0.029	
	Log-normal (R-CHOP/CVP)			**MeanLog**	**SDLog**	
		Meanlog	7.279	0.054	0.008	
		SDLog	0.477	0.008	0.016	

R^2 vs O-Benda

Efficacy	Distribution	Parameter	Base case	Variance-covariance	
PFS	Log-normal			**Intercept**	**Scale**
		Intercept	3.243	0.013	0.003
		Scale	1.248	0.003	0.005
OS	Exponential			**Intercept**	
		Intercept	4.696	0.359	
		Scale	8.665	-0.135	0.312
TTNLT	Log-normal			**MeanLog**	**SDLog**
		Meanlog	3.685	0.035	0.009
		SDLog	0.565	0.009	0.009

Utilities

Parameter	Coefficient	Variance-covariance							
		Intercept	Progression-free	Treatment: R2	Progressed on treatment	Baseline utility	Rituximab naïve	Refractory to last prior regimen	1 prior regimen
Intercept	0.326	0.00123	-0.00008	-0.00009	-0.00006	-0.00115	-0.00004	-0.00014	-0.00012
Progression-free	0.026	-0.00008	0.00009	-0.00001	0.00006	0.00000	0.00000	0.00001	0.00000
Treatment: R2	0.011	-0.00009	-0.00001	0.00020	0.00000	0.00001	0.00000	-0.00001	-0.00001
Progressed on treatment	-0.030	-0.00006	0.00006	0.00000	0.00017	0.00000	0.00000	0.00000	0.00000
Baseline utility	0.557	-0.00115	0.00000	0.00001	0.00000	0.00134	-0.00001	0.00006	0.00001
Rituximab naïve	0.028	-0.00004	0.00000	0.00000	0.00000	-0.00001	0.00037	-0.00009	-0.00001
Refractory to last prior regimen	-0.036	-0.00014	0.00001	-0.00001	0.00000	0.00006	-0.00009	0.00041	0.00006
1 prior regimen	0.034	-0.00012	0.00000	-0.00001	0.00000	0.00001	-0.00001	0.00006	0.00021

Appendix B: Summary of base case analysis inputs

Table 56: Summary of variables applied in the economic model used in OWSA

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Patient characteristics			
Mean baseline patient age (years)	vs R-CHOP: [redacted] vs R-CVP: [redacted] vs O-Benda: 63.0	[redacted] (Normal) [redacted] (Normal) 61.84-64.16 (Normal)	Section B.3.2, page 127
Percentage of female patients	vs R-CHOP: [redacted] vs R-CVP: [redacted] vs O-Benda: 44.5%	[redacted] (Beta) [redacted] (Beta) 37.0-52.1% (Beta)	
Percentage of patients rituximab naïve	vs R-CHOP: [redacted] vs R-CVP: [redacted]	[redacted] (Beta) [redacted] (Beta)	Section B3.4, page 182
Percentage of patients refractory to last prior regimen	vs R-CHOP: [redacted] vs R-CVP: [redacted] vs O-Benda: 93.9%	7.8-25.5% (Beta) 7.8-25.5% (Beta) 89.8-97.0% (Beta)	
Percentage of patients with 1 prior regimen	vs R-CHOP: [redacted] vs R-CVP: [redacted] vs O-Benda: 51.2%	[redacted] (Beta) [redacted] (Beta) 43.6-58.8% (Beta)	
Percentage of patients with baseline CrCl ≥60mL/min	Non-R-refractory: 86.6% R-refractory: 75.8%	82.9-89.9% (Beta) 68.0-82.8% (Beta)	Section B.3.2, page 132
Drug costs			
Pack cost: cyclophosphamide 500mg	£8.31	£8.28 - £8.34 (Normal)	Section B.3.5, page 188
Pack cost: cyclophosphamide 1000mg	£13.47	£13.45 - £13.49 (Normal)	
Pack cost: cyclophosphamide 2000mg	£27.50	£27.19 - £27.81 (Normal)	
Pack cost: doxorubicin 10mg	£4.48	£4.44 - £4.52 (Normal)	
Pack cost: doxorubicin 50mg	£17.78	£17.58 - £17.98 (Normal)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Pack cost: doxorubicin 200mg	£15.59	£15.25 - £15.93 (Normal)	
Pack cost: vincristine 1mg	£11.59	£11.54 - £11.64 (Normal)	
Pack cost: vincristine 2mg	£17.82	£17.50 - £18.14 (Normal)	
Pack cost: vincristine 5mg	£99.00	£95.24 - £102.76 (Normal)	
Pack cost: prednisolone 20mg	£4.17	£4.16 - £4.18 (Normal)	
Pack cost: bendamustine 25mg	£26.32	£22.34 - £30.30 (Normal)	
Pack cost: bendamustine 100mg	£75.13	£73.06 - £77.20 (Normal)	
R-maintenance usage after R-CHOP	47.8%	31.4 - 64.3% (Beta)	Section B3.5, page 193
R-maintenance usage after R-CVP	47.8%	31.2 – 64.6% (Beta)	
O-maintenance usage after O-Benda	77.5%	71.3 - 83.0% (Beta)	
Dose intensity of R^2: lenalidomide (AUGMENT)	[redacted]	[redacted] (Normal)	Section B.3.5, page 189
Dose intensity of R^2: rituximab (AUGMENT)	[redacted]	[redacted] (Normal)	
Dose intensity of R^2: lenalidomide (MAGNIFY)	[redacted]	[redacted] (Normal)	
Dose intensity of R^2: rituximab (MAGNIFY)	[redacted]	[redacted] (Normal)	
Dose intensity of O+B: bendamustine	92.6%	56.3 - 128.9% (Normal)	
Dose intensity of chemotherapy	87.5%	53.2 - 121.8% (Normal)	
Percent of patients needing G-CSF per 48 weeks: R^2	35.8%	28.9 – 43.0% (Normal)	Section B.3.5, page 192
Subsequent treatment costs			
Subsequent treatment total cost- R^2	£3,053.24	£1,856.39 - £4,250.09 (Normal)	Section B3.5, page 206
Subsequent treatment total cost - comparator	£7,712.10	£4,689.01 - £10,735.18 (Normal)	
Administration costs			
More complex chemotherapy, first attendance	£374.52	£227.71 - £521.33 (Normal)	Section B.3.5, page 196

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Chemotherapy, first attendance	£309.22	£188.01 - £430.43 (Normal)	
Chemotherapy, subsequent attendance	£312.34	£189.90 - £434.77 (Normal)	
Maintenance (IV)	£309.22	£188.01 - £430.43 (Normal)	
Maintenance (SC)	£247.74	£150.63 - £344.86 (Normal)	
Pharmacy preparation	£12.00	£7.30 - £16.70 (Normal)	
NHS transportation	£13.43	£8.17 - £18.70 (Normal)	
Adverse events – R^2 (non R-refractory)			
% Neutropenia	50.0%	42.6 - 57.4% (Beta)	Section B.3.3, page 171
% Leukopenia	6.8%	3.6 - 11.0% (Beta)	
% Anaemia	4.5%	2.0 - 8.1% (Beta)	
% Pneumonia	3.4%	1.3 - 6.5% (Beta)	
% Lymphocyte count decreased	3.4%	1.3 - 6.5% (Beta)	
% Lymphopenia	2.8%	0.9 - 5.7% (Beta)	
% Febrile neutropenia	2.8%	0.9 - 5.7% (Beta)	
% White blood cell count decreased	2.8%	0.9 - 5.7% (Beta)	
% Diarrhoea	2.8%	0.9 - 5.7% (Beta)	
% Thrombocytopenia	2.3%	0.6 - 4.9% (Beta)	
% Hypokalaemia	2.3%	0.6 - 4.9% (Beta)	
% Pulmonary embolism	2.3%	0.6 - 4.9% (Beta)	
% Infusion related reaction	2.3%	0.6 - 4.9% (Beta)	
% Nausea and emesis	0.0%	0 - 0% (Beta)	
% Allergic reaction	0.6%	0.0 - 2.1% (Beta)	
% Hypotension	0.6%	0.0 - 2.1% (Beta)	
% Fatigue	1.1%	0.1 - 3.1% (Beta)	
% Alopecia	0.0%	0 - 0% (Beta)	
% Infection	0.0%	0 - 0% (Beta)	
% Sepsis	0.0%	0 - 0% (Beta)	
% Abdominal pain	1.1%	0.1 - 3.1% (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
% Acute kidney injury	1.1%	0.1 - 3.1% (Beta)	
Adverse events – R-CHOP			
% Neutropenia	90.3%	87.1 - 93.1% (Beta)	Section B3.3, page 171
% Leukopenia	30.8%	26.2 - 35.6% (Beta)	
% Anaemia	4.5%	2.7 - 6.9% (Beta)	
% Pneumonia	0.0%	0 - 0% (Beta)	
% Lymphocyte count decreased	0.0%	0 - 0% (Beta)	
% Lymphopenia	0.0%	0 - 0% (Beta)	
% Febrile neutropenia	9.3%	6.6 - 12.5% (Beta)	
% White blood cell count decreased	0.0%	0 - 0% (Beta)	
% Diarrhoea	1.6%	0.6 - 3.1% (Beta)	
% Thrombocytopenia	2.0%	0.8 - 3.7% (Beta)	
% Hypokalaemia	0.0%	0 - 0% (Beta)	
% Pulmonary embolism	0.0%	0 - 0% (Beta)	
% Infusion related reaction	0.5%	0.0 - 1.4% (Beta)	
% Nausea and emesis	1.4%	0.4 - 2.8% (Beta)	
% Allergic reaction	0.0%	0 - 0% (Beta)	
% Hypotension	0.0%	0 - 0% (Beta)	
% Fatigue	3.2%	1.6 - 5.2% (Beta)	
% Alopecia	0.8%	0.2 – 2.0% (Beta)	
% Infection	0.0%	0 - 0% (Beta)	
% Sepsis	0.0%	0 - 0% (Beta)	
% Abdominal pain	1.2%	0.3 - 2.5% (Beta)	
% Acute kidney injury	0.0%	0 - 0% (Beta)	
Adverse events – R-CVP			
% Neutropenia	85.3%	69.7 - 95.8% (Beta)	Section B.3.3, page 171
% Leukopenia	16.6%	5.2 - 32.8% (Beta)	
% Anaemia	4.5%	0.2 - 15.0% (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
% Pneumonia	0.0%	0 - 0% (Beta)	
% Lymphocyte count decreased	0.0%	0 - 0% (Beta)	
% Lymphopenia	0.0%	0 - 0% (Beta)	
% Febrile neutropenia	5.0%	0.3 - 15.9% (Beta)	
% White blood cell count decreased	0.0%	0 - 0% (Beta)	
% Diarrhoea	5.5%	0.4 - 16.8% (Beta)	
% Thrombocytopenia	0.0%	0 - 0% (Beta)	
% Hypokalaemia	0.0%	0 - 0% (Beta)	
% Pulmonary embolism	0.0%	0 - 0% (Beta)	
% Infusion related reaction	0.0%	0 - 0% (Beta)	
% Nausea and emesis	3.8%	0.1 - 13.7% (Beta)	
% Allergic reaction	0.0%	0 - 0% (Beta)	
% Hypotension	0.0%	0 - 0% (Beta)	
% Fatigue	0.0%	0 - 0% (Beta)	
% Alopecia	0.0%	0 - 0% (Beta)	
% Infection	0.0%	0 - 0% (Beta)	
% Sepsis	0.0%	0 - 0% (Beta)	
% Abdominal pain	0.0%	0 - 0% (Beta)	
% Acute kidney injury	0.0%	0 - 0% (Beta)	
Adverse events – R^2 (R-refractory)			
% Neutropenia	42.2%	33.8 - 50.8% (Beta)	Section B.3.3, page 171
% Leukopenia	7.0%	3.3 - 12.0% (Beta)	
% Anaemia	3.1%	0.9 - 6.7% (Beta)	
% Pneumonia	3.1%	0.9 - 6.7% (Beta)	
% Lymphocyte count decreased	3.1%	0.9 - 6.7% (Beta)	
% Lymphopenia	3.1%	0.9 - 6.7% (Beta)	
% Febrile neutropenia	2.3%	0.5 - 5.6% (Beta)	
% White blood cell count decreased	3.9%	1.3 - 7.9% (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
% Diarrhoea	2.3%	0.5 - 5.6% (Beta)	
% Thrombocytopenia	5.5%	2.2 – 10.0% (Beta)	
% Hypokalaemia	2.3%	0.5 - 5.6% (Beta)	
% Pulmonary embolism	0.0%	0 - 0% (Beta)	
% Infusion related reaction	0.0%	0 - 0% (Beta)	
% Nausea and emesis	0.0%	0 - 0% (Beta)	
% Allergic reaction	0.0%	0 - 0% (Beta)	
% Hypotension	2.3%	0.5 - 5.6% (Beta)	
% Fatigue	5.5%	2.2 – 10.0% (Beta)	
% Alopecia	0.0%	0 - 0% (Beta)	
% Infection	0.0%	0 - 0% (Beta)	
% Sepsis	1.6%	0.2 - 4.3% (Beta)	
% Abdominal pain	2.3%	0.5 - 5.6% (Beta)	
% Acute kidney injury	2.3%	0.5 - 5.6% (Beta)	
Adverse events – O-Benda			
% Neutropenia	27.5%	21.6 - 33.8% (Beta)	Section B.3.3, page 171
% Leukopenia	0.0%	0 - 0% (Beta)	
% Anaemia	5.4%	2.7 - 8.9% (Beta)	
% Pneumonia	1.5%	0.3 - 3.5% (Beta)	
% Lymphocyte count decreased	0.0%	0 - 0% (Beta)	
% Lymphopenia	0.0%	0 - 0% (Beta)	
% Febrile neutropenia	3.9%	1.7 – 7.0% (Beta)	
% White blood cell count decreased	0.0%	0 - 0% (Beta)	
% Diarrhoea	1.0%	0.1 - 2.7% (Beta)	
% Thrombocytopenia	10.3%	6.5 - 14.8% (Beta)	
% Hypokalaemia	1.0%	0.1 - 2.7% (Beta)	
% Pulmonary embolism	0.0%	0 - 0% (Beta)	
% Infusion related reaction	8.8%	5.3 - 13.1% (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
% Nausea and emesis	1.0%	0.1 - 2.7% (Beta)	
% Allergic reaction	0.0%	0 - 0% (Beta)	
% Hypotension	0.5%	0.0 - 1.9% (Beta)	
% Fatigue	1.5%	0.3 - 3.5% (Beta)	
% Alopecia	0.0%	0 - 0% (Beta)	
% Infection	0.0%	0 - 0% (Beta)	
% Sepsis	1.0%	0.1 - 2.7% (Beta)	
% Abdominal pain	0.0%	0 - 0% (Beta)	
% Acute kidney injury	0.0%	0 - 0% (Beta)	
Adverse events – R maintenance			
% Neutropenia	11.5%	7.1 – 16.7% (Beta)	Section B.3.3, page 173
% Infection	9.7%	5.7 – 14.6% (Beta)	
Adverse events – O maintenance			
% Neutropenia	10.8%	6.4 – 16.0% (Beta)	Section B.3.3, page 173
% Infection	3.2%	1.0 – 6.4% (Beta)	
Adverse events – ASCT			
% Febrile neutropenia	98.3%	93.9 – 100.0% (Beta)	Section B.3.3, page 173
Adverse event costs			
Neutropenia	£1,893	£1151 - £2634 (Normal)	Section B.3.5, page 202
Leukopenia	£3,415	£2076 - £4754 (Normal)	
Anaemia	£2,996	£1821 - £4170 (Normal)	
Pneumonia	£2,527	£1536 - £3517 (Normal)	
Lymphocyte count decreased	£382	£232 - £532 (Normal)	
Lymphopenia	£382	£232 - £532 (Normal)	
Febrile neutropenia	£6,512	£3959 - £9064 (Normal)	
White blood cell count decreased	£382	£232 - £532 (Normal)	
Diarrhoea	£1,508	£917 - £2099 (Normal)	
Thrombocytopenia	£2,755	£1675 - £3835 (Normal)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Hypokalaemia	£339	£206 - £472 (Normal)	
Pulmonary embolism	£1,330	£809 - £1851 (Normal)	
Infusion related reaction	£618	£376 - £861 (Normal)	
Nausea and emesis	£618	£376 - £861 (Normal)	
Allergic reaction	£395	£240 - £550 (Normal)	
Hypotension	£2,169	£1319 - £3019 (Normal)	
Fatigue	£93	£57 - £130 (Normal)	
Alopecia	£0	£0 - £0 (Normal)	
Infection	£1,570	£955 - £2186 (Normal)	
Sepsis	£2,830	£1720 - £3939 (Normal)	
Abdominal pain	£623	£379 - £868 (Normal)	
Acute kidney injury	£2,674	£1626 - £3722 (Normal)	
Resource use cost			
Haematologist led	£165	£100 - £229 (Normal)	Section B.3.5, page 200
CT scans	£137	£83 - £190 (Normal)	
FBC	£6	£4 - £9 (Normal)	
Patient history/physical exam	£6	£4 - £9 (Normal)	
Full profile (U&E, LFT, calcium)	£17	£10 - £24 (Normal)	
Serum IgG, IgA, IgM and electrophoresis)	£25	£15 - £35 (Normal)	
LDH test	£13	£8 - £18 (Normal)	
EOL cost	£6,083	£3,399 - £8,467 (Normal)	Section B.3.5, page 207
Autologous SCT cost	£35,558	£21,620 - £49,497 (Normal)	Section B.3.5, page 202
Resource use monthly frequency			
Induction - Haematologist led	1.00	0.61 - 1.39 (Normal)	Section B.3.5, page 200
Maintenance -Haematologist led	0.33	0.20 - 0.46 (Normal)	
Follow-up -Haematologist led	0.25	0.15 - 0.35 (Normal)	
Progressed -Haematologist led	1.00	0.61 - 1.39 (Normal)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Induction -FBC, patient history, physical exam, LFT, U&E, immunoglobulin tests, lactate dehydrogenase	1.00	0.61 - 1.39 (Normal)	
Maintenance -FBC, patient history, physical exam, LFT, U&E, immunoglobulin tests, lactate dehydrogenase	0.33	0.20 - 0.46 (Normal)	
Follow-up -FBC, patient history, physical exam, LFT, U&E, immunoglobulin tests, lactate dehydrogenase	0.25	0.15 - 0.35 (Normal)	
Progressed -FBC, patient history, physical exam, LFT, U&E, immunoglobulin tests, lactate dehydrogenase	1.00	0.61 - 1.39 (Normal)	
Induction -CT scans	0.17	0.10 - 0.23 (Normal)	
Maintenance -CT scans	0.04	0.03 - 0.06 (Normal)	
Follow-up -CT scans	0.00	0 - 0 (Normal)	
Progressed -CT scans	0.00	0 - 0 (Normal)	
lenalidomide monitoring - cycle 1	3.00	1.82 - 4.18 (Normal)	Section B.3.5, page 197
lenalidomide monitoring - cycle 2-4	2.00	1.22 - 2.78 (Normal)	
lenalidomide monitoring - cycle 5+	1.00	0.61 - 1.39 (Normal)	
Proportion having ASCT after R-CHOP	11.8%	7.6-16.7% (Beta)	Section B.3.5, page 201
Proportion having ASCT after R-CVP	0.0%	0 - 0% (Beta)	
Proportion having ASCT after O-Benda	0.0%	0 - 0% (Beta)	
Utilities			
Pooled baseline utility	0.837	0.82 – 0.86 (Beta)	Section B.3.4, page 176
Adverse event disutility			
Neutropenia	0.090	0.06 - 0.12 (Beta)	Section B.3.4, page 180
Leukopenia	0.119	0.08 - 0.16 (Beta)	
Anaemia	0.119	0.08 - 0.16 (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Pneumonia	0.200	0.16 - 0.24 (Beta)	
Lymphocyte count decreased	0.100	0.06 - 0.14 (Beta)	
Lymphopenia	0.100	0.06 - 0.14 (Beta)	
Febrile neutropenia	0.150	0.10 - 0.21 (Beta)	
White blood cell count decreased	0.100	0.06 - 0.14 (Beta)	
Diarrhoea	0.048	0.02 - 0.08 (Beta)	
Thrombocytopenia	0.108	0.07 - 0.15 (Beta)	
Hypokalaemia	0.124	0.09 - 0.16 (Beta)	
Pulmonary embolism	0.124	0.09 - 0.16 (Beta)	
Infusion related reaction	0.195	0.12 - 0.28 (Beta)	
Nausea and emesis	0.048	0.02 - 0.08 (Beta)	
Allergic reaction	0.098	0.06 - 0.14 (Beta)	
Hypotension	0.057	0.04 - 0.08 (Beta)	
Fatigue	0.073	0.04 - 0.11 (Beta)	
Alopecia	0.045	0.02 - 0.08 (Beta)	
Infection	0.195	0.12 - 0.28 (Beta)	
Sepsis	0.254	0.16 - 0.36 (Beta)	
Abdominal pain	0.069	0.04 - 0.10 (Beta)	
Acute kidney injury	0.270	0.17 - 0.38 (Beta)	
Adverse event durations			
Neutropenia	15.09	9.17 - 21.01 (Normal)	Section B.3.4, page 180
Leukopenia	13.96	8.49 - 19.43 (Normal)	
Anaemia	16.07	9.77 - 22.37 (Normal)	
Pneumonia	14.00	8.51 - 19.49 (Normal)	
Lymphocyte count decreased	34.00	20.67 - 47.33 (Normal)	
Lymphopenia	34.00	20.67 - 47.33 (Normal)	
Febrile neutropenia	7.14	4.34 - 9.94 (Normal)	
White blood cell count decreased	34.00	20.67 - 47.33 (Normal)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Diarrhoea	34.00	20.67 - 47.33 (Normal)	
Thrombocytopenia	23.23	14.12 - 32.34 (Normal)	
Hypokalaemia	34.00	20.67 - 47.33 (Normal)	
Pulmonary embolism	34.00	20.67 - 47.33 (Normal)	
Infusion related reaction	34.00	20.67 - 47.33 (Normal)	
Nausea and emesis	6.00	3.65 - 8.35 (Normal)	
Allergic reaction	34.00	20.67 - 47.33 (Normal)	
Hypotension	8.00	4.86 - 11.14 (Normal)	
Fatigue	31.50	19.15 - 43.85 (Normal)	
Alopecia	34.00	20.67 - 47.33 (Normal)	
Infection	34.00	20.67 - 47.33 (Normal)	
Sepsis	34.00	20.67 - 47.33 (Normal)	
Abdominal pain	17.00	10.34 - 23.66 (Normal)	
Acute kidney injury	29.75	18.09 - 41.41 (Normal)	

Key: CI, confidence interval; CrCl, creatine clearance; CT, computed tomography; EOL, end of life; FBC, full blood count; G-CSF, Granulocyte-colony stimulating factor; Ig, immunoglobin; LDH, lactate dehydrogenase; LFT, liver function tests; NHS, National Health Service; U&E, Urea and Electrolytes.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[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. In 2018 we raised a total income of £1,432,177 from various fundraising activities. 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. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and	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 nine

carers to include in your submission?	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?	The symptoms of follicular lymphoma and marginal zone lymphoma can be variable, ranging from no symptoms at all to lumps, abdominal symptoms or systemic symptoms such as weight loss, fever, night sweats, fatigue and recurrent infections. Some people with follicular lymphoma or marginal zone lymphoma do not need treatment initially and instead enter a period of active monitoring (watch and wait). Psychologically, this can be very difficult to cope with. The uncertainty of active monitoring can be stressful and many people experience anxiety. Some people find it hard to plan for the future because they don't know if or when they may need to start treatment. Most people with follicular lymphoma and marginal zone lymphoma are treated to keep the lymphoma in remission for as long as possible with as few side effects as possible. However, these types of lymphoma usually relapse and most people need several courses of treatment during their illness. Follicular lymphoma and all subtypes of marginal zone lymphoma can transform to more aggressive forms of lymphoma. Some people with follicular lymphoma and marginal zone lymphoma feel well. However, some experience recurrent symptoms, side effects and late effects of treatment, and psychological issues relating to having an incurable disease. Many patients find it emotionally and psychologically difficult to live with a condition that they know is likely to relapse. Some have reported losing the confidence to carry on with their usual activities, becoming agoraphobic and feeling 'scared of living and terrified of dying.' Many feel constantly mindful of a relapse, which causes stress both for them and their carers. Some report worrying that their lymphoma has relapsed every time they are ill, and feeling anxious at the thought of having to go through more

	chemotherapy. Anxiety about the future is common, as is the impact of a long-term condition on family and friends. Some people report worrying that other people have to care for them. 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. Some patients have enduring symptoms, such as fatigue, that can affect their ability to work and take part in their chosen leisure activities. Some have persistent side effects of treatment, such as peripheral neuropathy, that can cause 'constant pain', disrupt sleep and reduce quality of life. Recurrent infections are also common. One patient reported a serious reaction to treatment and went from walking with a stick to being in a wheelchair, totally reliant on a carer, within 4 weeks. It took a year to regain leg and arm strength and left the patient with permanent nerve damage. Another experienced neutropenia that lasted 3 years, leading to repeated bouts of illness and several hospital admissions with infections. Having regular medical appointments, scans and blood tests can also be disruptive. Having support from friends and family is seen as very important. Howver, for carers, it can be difficult seeing a loved one experiencing symptoms or going through treatment and feeling unable to do anything to help. People have told us it is psychologically and physically challenging supporting their loved ones practically and emotionally whilst trying to take care of their own needs. The psychological impact of an 'incurable' disease affects carers as well as patients. Some carers worry that their loved one's lymphoma has relapsed whenever they are ill.

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	When treatment is needed for follicular lymphoma and marginal zone lymphoma, most people have chemotherapy with antibody therapy. This may be followed by maintenance antibody treatment. Radiotherapy and surgery are used occasionally. People with splenic marginal zone lymphoma may need a splenectomy, which leaves them at higher risk of infection for the rest of their lives. When lymphoma relapses, patients might need more intensive treatment such as high-dose chemotherapy or a stem cell transplant. Patients generally feel that current treatment is effective at putting their lymphoma into complete or partial remission but that the response isn't durable long-term. Many had experienced significant side effects or adverse reactions to treatment. Reliance on chemotherapy was a concern. One patient reported being unlikely to consent to more chemotherapy when the lymphoma next relapses, having experienced significant side effects from a previous course of immuno-chemotherapy. Some patients have entered clinical trials to receive targeted treatments that are not currently available as standard therapy.
8. Is there an unmet need for patients with this condition?	Patients felt there was a need for: • treatments that offer a higher chance of durable remission • treatments with fewer side effects to help maintain a good quality of life • chemotherapy-free treatment options • treatments that are easier to administer • earlier and better diagnosis • consistent standards of care throughout the UK.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	The main advantages patients felt lenalidomide and rituximab could offer over current treatment options are: • longer remission (progression-free survival) leading to improved quality of life • easier route of administration/dosing schedule than chemotherapy • no need to have chemotherapy and its associated side effects • 'hope for the future'.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	As with all treatments, patients were concerned about the potential side effects and felt that the benefits must outweigh the risks. However, some patients felt side effects were 'worth suffering' if the treatment works. Some were concerned about any increased monitoring that might be necessary. As with any newer treatment, any potential late effects of lenalidomide are as yet unknown.

Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients felt the combination of lenalidomide and rituximab should be available to all suitable patients. However, it might be of particular benefit to people who are unable to tolerate existing treatments.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Patients felt that if lenalidomide and rituximab is approved by NICE it should be available in all parts of the UK to people whom consultants consider will benefit from the treatment.

Other issues	
13. Are there any other issues that you would like the committee to consider?	Any cost-effectiveness analysis should take into account the cost of repeated hospital admissions and treatment of side effects caused by existing treatment options.

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission:

- People with relapsed or refractory follicular or marginal zone lymphoma can experience a variety of symptoms that can affect their day-to-day lives.
- There is a significant psychological impact on both patients and carers of having these ‘incurable’ types of lymphoma.
- Side effects and late effects of existing treatment options can be challenging to live with and can have a significant impact on quality of life.
- There is an unmet need for chemotherapy-free treatments that offer a higher chance of durable remission with fewer side effects.
- Lenalidomide and rituximab therapy is seen as a ‘hope for the future’.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Professional organisation submission

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	NCRI/RCP/RCR/ACP

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☒ other (please specify):
5a. Brief description of the organisation (including who funds it).	• The National Cancer Research Institute (NCRI) is a partnership of cancer research funders. • The Royal College of Physians (RCP) is an independent, patient-centred and clinically led organisation that drives improvement in disease diagnosis and management. • The Royal College of Radiologists (RCR) leads, educates and supports doctors who are training and working in the specialties of clinical oncology and clinical radiology. • The Association of Cancer Physicians (ACP) is the specialty association for medical oncologists.
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or	1. To improve quality of life by inducing remission and, in doing so, alleviating disease-related symptoms. 2. To prolong life by extending the natural history of the disease and preventing disease-related complications.

disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	A clinically significant short-term treatment response in indolent NHL would normally be defined as a complete or partial anatomical response. A complete response indicates that there is no detectable disease on imaging and bone marrow biopsy whereas a partial response indicates shrinkage of 2-dimensional tumour size by at least 50%. Given the long natural history of indolent NHL, long-term efficacy endpoints are more clinically meaningful than short-term ones. The most meaningful endpoint is overall survival but progression-free survival is a more practical. Other endpoints such as time to next treatment are arguably more meaningful but also more subject to investigator bias.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. Unlike other mature B-cell malignancies such as chronic lymphocytic leukaemia, mantle-cell lymphoma and lymphoplasmacytic lymphoma, further treatment options for patients with follicular lymphoma or marginal zone lymphoma who fail chemoimmunotherapy are limited owing to a lack of approved novel agents.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Advanced-stage follicular lymphoma (FL) is generally considered to an indolent but incurable disease with a clinical course characterised by recurrent relapses and remissions. Treatment is usually deferred in asymptomatic patients with low tumour burden, although rituximab monotherapy is approved by NICE as an option for this indication. Initial treatment of high tumour burden FL is usually with a CD20 antibody in combination with one of 3 different chemotherapy regimens: bendamustine or cyclophosphamide, doxorubicin and prednisolone with or without vincristine (CHOP and CVP, respectively). Until recently, the only CD20 antibody available for frontline chemoimmunotherapy was rituximab, but obinutuzumab (GA101) was recently NICE approved as an option for patients with high or intermediate FLIPI scores. Patients who respond to chemoimmunotherapy have the option of continuing the CD20 antibody as maintenance therapy for 2 years, although this is increasingly controversial.

	The choice of subsequent therapy depends on a number of factors including the type of first-line therapy given, how well it worked and how much and what type of toxicity it caused. General fitness and comorbidity are also important factors. Treatment options for relapsed or refractory follicular lymphoma include further rituximab-based chemoimmunotherapy (with or without rituximab maintenance), obinutuzumab plus bendamustine (NICE approved for rituximab-refractory patients only), or chemoimmunotherapy followed by autologous stem-cell transplantation (ASCT) for patients who are sufficiently fit. Rituximab monotherapy is available for patients who have exhausted other treatment options but is rarely used as such patients are usually refractory to rituximab-containing chemo-immunotherapy. Occasional patients may undergo allogeneic stem-cell transplantation. This is a potentially curative treatment but requires patients to be in remission and is associated with significant morbidity and mortality. Some targeted agents (e.g. lenalidomide and idelalisib) may be available through patient access schemes. The treatment of marginal-zone lymphoma is even more controversial due to a relative paucity of high-quality clinical studies. However, with the exception of localised extranodal disease, treatment is generally fairly similar to that of FL.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	• NICE guidelines on non-Hodgkin lymphoma (NG52) - includes limited sections on FL and extranodal MZL (2016). • European Society for Medical Oncology (ESMO) guidelines for FL (2016). • British Society for Haematologists (BSH) guidelines for FL (2011, updated in 2017 to include frontline bendamustine plus rituximab).
• 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.)	Frontline therapy for FL is reasonably well defined (chemo-immunotherapy) but nevertheless controversial in terms of the choice of chemotherapy, the choice of anti-CD20 antibody and deployment of post-induction anti-CD20 maintenance. Subsequent treatment pathways are less well defined owing to the large number of variables that impact on treatment decisions in the relapsed/refractory setting. Treatment pathways for MZL are even less well defined owing to the heterogeneity in clinical presentation. These controversies are global.

• What impact would the technology have on the current pathway of care?	NICE approval of lenalidomide plsu rituximab in would have a big impact on treatment pathways for FL and MZL as patients who fail chemo-immunotherapy currently have limited treatment options available to them.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Lenalidomide is already NICE approved for myeloma and MDS. Furthermore, lenalidomide plus rituximab is already being used in selected patients relapsed/refractory FL via the Celgene patient access scheme. Consequently, most if not all UK haemato-oncology units should have direct experience of using lenalidomide.
• How does healthcare resource use differ between the technology and current care?	Compared to rituximab monotherapy, lenalidomide plus rituximab should require minimal additional healthcare resource. Compared to chemo-immunotherapy, lenalidomide plus rituximab should requires less healthcare resource as, unlike most commonly used chemotherapy regimens, lenalidomide is given orally.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care (haemato-oncology outpatient clinics).
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No specific investment is needed to introduce the technology as lenalidamide is already NICE approved for myeloma and MDS, and most haemato-oncologists are therefore already familiar with it.
11. Do you expect the technology to provide clinically	Yes. The AUGMENT trial showed an improvement in median PFS from 14.1 months to 39.4 months as a result of adding lenalidomide to rituximab in patients with relapsed or refractory FL and MZL that was not refractory to rituximab per se. Therefore, lenalidomide plus rituximab is expected to provide clinically

meaningful benefits compared with current care?	meaningful benefit in this population compared to rituximab monotherapy. Since lenalidomide plus rituximab also has established activity in patients with indolent NHL that is refractory to rituximab, the combination is an attractive option for these patients too, especially since chemoimmunotherapy is unlikely to be effective in this setting.
• Do you expect the technology to increase length of life more than current care?	Yes. Population-based studies have shown that survival of patients with indolent NHL has significantly improved over the last few decades. The most likely explanation for this observation is the advent of novel agents which prolong PFS. Although PFS benefit does not always translate into OS benefit in randomised clinical trials, it should be emphasised that many patients who do not receive disease-modifying treatments as part of a trial may do so later in the course of their disease where the benefits may be similar. In other words, many patients end up receiving the same treatments, just in a different order.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes. Active follicular lymphoma often produces symptoms such as pain, sweating and fatigue which reduce quality of life. These symptoms usually increase with tumour burden. Technologies that induce remissions and prevent disease progression should therefore reduce such symptoms and in doing so improve quality of life.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	The AUGMENT trial showed that the benefit of lenalidomide plus rituximab extended over all patients groups examined with the exception of those with marginal zone lymphoma. However, since only 63 patients in the intention-to-treat population had MZL, the comparison was not sufficiently powered in this subset of patients. Furthermore, within the MZL group, clinical features that are generally associated with less favourable in outcome in lymphoma were more prominent in the lenalidomide arm compared with placebo. Although AUGMENT excluded patients who were refractory to rituximab, these patients may benefit the most from lenalidomide plus rituximab owing to a lack of alternative treatment options and the established activity of the drug combination in this setting.
The use of the technology	

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	Lenalidomide administration should not pose any problems for patients or healthcare professionals as it is taken orally and does not require any intravenous infusions. Lenalidomide is associated with some notable toxicities including haematopoietic suppression and a small risk of thromboembolism. Consequently, it is generally recommended that patients receiving lenalidomide who are not thrombocytopenic should receive low-dose aspirin or low-molecular-weight heparin if the risk is considered significant enough. Patients should also be monitored for haematopoietic suppression and supported with G-CSF and/or blood transfusion if required. Any infections should be treated promptly with appropriate antibiotics. Healthcare professionals should have an awareness of these potential toxicities, and patients should be managed appropriately if they occur.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Lenalidomide plus rituximab should be suitable for most patients with relapsed/refractory follicular lymphoma. However, because of its pharmacokinetic and toxicity profile, caution should be used in patients with significant renal impairment, active infection, haematopoietic suppression, recent thrombo-embolism and neuropathy. Treatment should be paused in the event of significant toxicity and stopped altogether if the toxicity is life-threatening.
15. Do you consider that the use of the technology will result in any substantial health-	None I can think of.

related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes. The prolongation in survival observed over the last few decades is likely to reflect the incremental availability of new treatments that can be effective when previous ones have failed. Lenalidomide plus rituximab constitutes such a new treatment option.
• Is the technology a 'step-change' in the management of the condition?	Yes. Lenalidomide would be the first targeted agent (with the exception of CD20 antibodies) to be NICE approved for FL and would therefore represents a significant step-change in treatment algorithms if it is approved.
• Does the use of the technology address any particular unmet need of the patient population?	Yes. Lenalidomide is suitable for older, less fit patients who would not tolerate further chemoimmunotherapy. This is pertinent given that the median age of patients with newly diagnosed FL is about 60.
17. How do any side effects or adverse effects of the	The unwanted effects lenalidomide and rituximab need to be balanced against the beneficial effects, taking into account the paucity of alternative treatment options. For most patients, the benefits in terms of disease control and improvement in disease-related symptoms are likely to far outweigh any negative aspects. By

technology affect the management of the condition and the patient's quality of life?	following dose reduction algorithms for toxicity, most patients can be established on a dose of lenalidomide that is reasonably well tolerated.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	In AUGMENT, patients were required to have received at least 2 doses of rituximab. However, many patients nowadays receive frontline chemoimmunotherapy containting obinutuzumab instead of rituximab. Also, most patients with FL and MZL in the UK who exhaust current treatment options are likely to be rituximab refractory, whereas AUGMENT excluded such patients (to include them would have been unethical).
• If not, how could the results be extrapolated to the UK setting?	It would be reasonable to assume that lenalidomide is superior to rituximab monotherapy in patients with relapsed/refractory FL who were previously treated with obinutuzumab-containing chemoimmunotherapy and to extend the indication for lenalidomide plus rituximab to this patient group. It would also be reasonable to assume that lenalidomide plus rituximab is superior to rituximab monotherapy in patients who are refractory to rituximab and to extend the indication for lenalidomide plus rituximab to this important patient group.
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival is the most important outcome in indolent NHL but as already explained there is often a disconnect between PFS and OS in randomised clinical trials as patients in the standard arm may still benefit from the disease-modifying treatment later during the course of their disease. Under these circumstances, it is reasonable to accept PFS as the primary outcome.
• If surrogate outcome measures were used, do they adequately predict long-term clinical	As explained, PFS benefit may not always translate into OS benefit in trials of indolent NHL but this does not necessarily mean that new treatments do not improve OS. In fact, in AUGMENT did show a non-significant trend for improved OS in the lenalidomide arm. The fact that this did not reach statistical significance can be explained by the delayed divergence of OS curves on the Kaplan-Meier plot. This difference in OS may become statistically significant with longer follow-up.

outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. Lenalidomide has been used for other indications (myeloma and MDS) for some time and its toxicity profile is well defined.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA137 and TA226]?	No.
21. How do data on real-world experience compare with the trial data?	I am not aware of any real-world studies investigating lenalidomide plus rituximab in indolent NHL.
Equality	

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	None I can think of.
22b. Consider whether these issues are different from issues with current care and why.	N/A
Topic-specific questions	
23. Are follicular lymphoma and marginal zone lymphoma clinically distinct? a. Does the treatment (2nd or subsequent line) differ by histology? b. Does the population characteristics differ by histology, if so how? c. Would you expect different	Unlike FL, MZL presents in a clinically heterogeneous way and can be divided into 3 main types: i. Extranodal – often localised, may be driven by infection and respond to antimicrobial therapy (e.g. H pylori in gastic MZL) – 8% of patients in AUGMENT ii. Splenic – overlaps with a distinctive form of leukaemia – 5% of patients in AUGMENT iii. Nodal – resembles FL in terms of its clinical behaviour and therapeutic sensitivity – 4% of patients in AUGMENT (a) Rituximab monotherapy is likely to be used more in MZL than in FL. (b) MZL is more likely to present with early-stage disease (c) MZL is likely to have a slightly better outcome than FL

outcomes for follicular compared with marginal zone lymphoma?	
24. Are outcomes likely to differ for people previously treated with rituximab compared with rituximab naïve?	Not necessarily. What matters the most is whether a patient is refractory to rituximab as these patients have a worse outcome.
25. Are there any clinically important prognostic factors for follicular compared with marginal zone lymphoma that predict outcomes? a. Do these differ by histology?	Prognostic factors in FL fall into the following categories: 1. Pre-treatment patient characteristics (e.g. FLIPI) 2. Response to induction therapy (e.g. post-induction PET-CT status) 3. Disease status several years after diagnosis/treatment (e.g. POD24) Prognostic factors are much less well developed and validated in MZL compared to FL. Histology is only relevant if • There is high-grade transformation to large-cell lymphoma irrespective of whether the initial disease was FL or MZL. This is associated with a worse outcome and such patients are usually treated with high-grade NHL protocols. • The malignant germinal centres in FL consist of sheets of centroblasts (large cells). This is referred to as grade 3b FL. Such patients also have a worse prognosis and are usually treated with high-grade NHL protocols.
Key messages	

26. In up to 5 bullet points, please summarise the key messages of your submission.

- Lenalidomide plus rituximab represents a major treatment breakthrough for indolent NHL as there are currently no approved targeted therapies other than CD20 antibodies.
- The superiority of lenalidomide plus rituximab over rituximab alone in the AUGMENT trial is of sufficient magnitude to justify using lenalidomide plus rituximab as an alternative to chemoimmunotherapy in patients with relapsed/refractory indolent NHL.
- Lenalidomide plus rituximab is generally well tolerated with a toxicity profile that seems fully justified by its clinical efficacy in a setting where effective treatment options are limited.
- Lenalidomide is easy to administer and already widely used, not only in indolent NHL where it is accessible via the Celgene patient access scheme, but also in myeloma and MDS for which it is NICE approved.
- Since many patients with FL are now receiving frontline chemoimmunotherapy containing obinutuzumab instead of rituximab, any requirement for “prior rituximab” should be extended to “prior anti-CD20 antibody” when defining the eligible population. Furthermore, if approved, lenalidomide plus rituximab should be available for patients who are refractory to anti-CD20-containing regimens as these patients have the worst outcome and greatest clinical need.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Clinical expert statement

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us 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.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Kim Linton**
2. Name of organisation	**The Christie NHS Foundation Trust**

3. Job title or position	**Clinical Senior Lecturer**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
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 didn't submit one, I don't 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

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	For most people affected, follicular lymphoma is a chronic and incurable condition which usually relapses many times over a period of 15-20 years. Therapy is necessary to control the disease and preserve or improve quality of life.
8. 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.)	Time to next treatment (TTNT) is the most clinically significant endpoint in this disease. Response rates and progression free survival are reasonable surrogates.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. Immunochemotherapy is currently the only available treatment option for patients with relapsed FL. There is a clear unmet need to license novel therapies to enable patients with FL to continue to treatment beyond the point when chemotherapy is no longer an option: 1. When people run out of treatment options – frequent relapses mean most patients quickly exhaust the finite number of immunochemotherapy options 2. When the disease is refractory to chemotherapy and/or rituximab – more chemotherapy, even if available, is therefore ineffective. Any patient who relapses early (i.e. within 1-2 years of therapy) should be considered to be chemo-refractory. 1 in 5 patients has primary refractory disease and

	relapses within 2 years of first line treatment, but ultimately all patients with a relapsing/remitting course become chemo-refractory. 3. When chemotherapy is too toxic – either because of patient age/frailty or when heavily pre-treated patients with multiply relapsed FL become chronically immunosuppressed and therefore unsuitable for more chemotherapy
What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	Relapsed FL/MZL is typically treated with further immunochemotherapy, the choice of which depends on the previous treatment, its duration of response and tolerability; the patient's fitness, wishes, disease characteristics, comorbidity and available treatment options. The most common regimens use an antiCD20 monoclonal antibody in combination with bendamustine, CHOP or CVP, sometimes followed by maintenance antibody. Rituximab monotherapy is used infrequently to treat symptomatic relapsed disease and, like chlorambucil, is usually only offered to very frail patients. Transplant eligible patients may also be offered a platinum-based chemotherapy regimen e.g. GDP or DHAP.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes. Most centres refer to NICE guidelines (NG52 published 2016) and BCSH guidelines (updated 2019 guidelines in press)
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please	The pathway is not clearly defined as there are no standard options at first line or relapse. Practice varies considerably depending on clinician experience and preference. Here is an example of first line of therapy (LOT) decision making: • R-bendamustine is the most common first LOT regimen in England – offered mainly to younger and fitter patients • R-CHOP is preferred for patients with suspected/confirmed high grade transformation

state if your experience is from outside England.)	• R-CVP is preferred for patients who are unsuitable for R-bendamustine (some patients aged >70) or R-CHOP (due to heart disease) • R-chlorambucil or R monotherapy is preferred for very elderly/frail patients • The advent of obinutuzumab (O) has added to the complexity of first line treatment decisions for FL. For example, O-CVP is now increasing used instead of R-benda, especially in FLIPI2+ patients aged >70. There is no standard therapy at relapse. Patients will typically be offered re-treatment with an effective therapy (except RCHOP) if this previously well tolerated and produced a durable remission, or one of the other immuno-chemotherapy options above. Please note that frail patients treated with upfront R-chlorambucil/R-CVP may be unsuitable for R-CHOP/R-benda and will consequently exhaust available options early during the disease course. Transplant-eligible patients may also be offered platinum-based immunochemotherapy.
• What impact would the technology have on the current pathway of care?	The new technology provides an **additional line of therapy** for patients with relapsed FL/MZL. This alone could prolong overall survival for patients. The new technology may be offered to chemotherapy- and/or rituximab-refractory patients **to overcome treatment resistance**, and may be preferred to current options such as O-bendamustine (the new technology avoids the cumulative toxicity of more bendamustine) or R-chemotherapy (the new technology overcomes rituximab refractoriness by its combination with lenalidomide, which is not the case with standard R-chemo) The new technology is well tolerated, providing a particularly attractive **effective and safe treatment option for older/frailer patients** and others for whom chemotherapy is expected to be too toxic.
11. Will the technology be used (or is it already used) in	The technology is not yet available for routine use in NHS clinical practice. Some clinicians have used the technology via a Celgene named patient programme. I expect the technology will be used in the same way as the NNP and clinical trials.

the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	There are negligible differences in healthcare resource use. Patients receiving the new technology must be counselled on contraception and the risks of pregnancy (lenalidomide is teratogenic).
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Specialist oncology/haematology clinics in tertiary referral centres
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	The NHS is already set up to deliver this technology. No specialist equipment is needed as lenalidomide is an oral agent. As with all new treatments, staff training will be required.
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Please refer to my answers above.

• Do you expect the technology to increase length of life more than current care?	This is uncertain. The new technology will *replace standard of care* when the patient is unsuitable for immunochemotherapy, e.g. no remaining standard options; standard treatment deemed too toxic; patient likely to be refractory to standard treatment. The new technology can also *displace standard of care*, i.e. be used instead of standard treatment in a patient who will go on receive standard care at a later relapse. In both scenarios, the new technology is likely to increase length of life as it provides either a more suitable or an extra option for our patients
• Do you expect the technology to increase health-related quality of life more than current care?	Possibly. The new technology is well tolerated with a different side effect profile to standard immunochemotherapy and rituximab monotherapy. The first line RELEVANCE study which compared the new technology with immunochemotherapy in the first line treatment reported differing side effect profiles but setting did not report HR-QOL data.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes. The technology may be more effective for patients with chemotherapy and/or rituximab refractory patients and may be safer for patients for whom conventional chemotherapy may be too toxic.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare	Easier for administration (oral therapy, less day unit time as no infusional chemotherapy) but first cycle dosing is a little more complex than standard treatment. More vigilant pregnancy and contraception counselling and documentation is required for the new technology. Patients will also need to be considered

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.)	for thromboprophylaxis, tumour lysis precautions and antimicrobial prophylaxis which may be more or less than standard care depending on the regimen.
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Yes. Dose adjustments are necessary for patients with renal impairment. Other stopping rules are standard (excessive toxicity, disease progression, patient choice).
16. 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	No

quality-adjusted life year (QALY) calculation?	
17. 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?	This innovative new technology could meet the unmet needs of a) high risk chemo and/or rituximab refractory patients experiencing an early first relapse, b) heavily pre-treated multiply relapsed patients who have either become chemo-refractory or exhausted standard options and c) patients who are not fit for standard chemotherapy due to its potential toxicity. The new technology is has a unique mode of action that is well tolerated and proven to overcome rituximab resistance. If employed, the new technology could make a substantial improvement to duration and quality of life for these patient groups.
• Is the technology a 'step-change' in the management of the condition?	Yes, I believe it is for the reasons outlined above
• Does the use of the technology address any particular unmet need of the patient population?	Please refer to my answers above
18. How do any side effects or adverse effects of the technology affect the	I have treated several relapsed FL patients with the new technology. Most patients tolerate treatment very well and remain on therapy with few dose reductions. This favourable safety profile has been reported in trials to cause less neutropenic sepsis (the leading cause of therapy discontinuation and treatment related

management of the condition and the patient's quality of life?	death) than chemotherapy and only slightly more infections compared to rituximab monotherapy and is well tolerated even in frailer/older patients, thus fulfilling our overall aim of controlling the disease and preserving quality of life.
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	The AUGMENT study (new technology R2 vs R monotherapy) doesn't reflect UK practice where R-chemotherapy is used much more often than R monotherapy. The MAGNIFY trial (R2 induction +/- R2 maintenance) isn't yet relevant to our practice. Unfortunately, there are no trial comparisons with standard of care immunochemotherapy and limited real world immunochemotherapy outcome data.
• If not, how could the results be extrapolated to the UK setting?	Results can be extrapolated to the UK centre by comparisons with real world data. Results of the RELEVANCE first line trial (R2 vs immunochemotherapy) can also be extrapolated to the relapsed setting. This large trial of >1000 patients reported similar response rates and PFS for the two treatment arms. Side effect profiles were different but with no difference in the overall rate of adverse events between arms. Newly diagnosed and relapsed FL/MZL are biologically the same disease and in clinical practice data from first line trials are frequently extrapolated to the relapsed setting. Thus, the immunochemotherapy regimens are used interchangeably for first line and relapsed disease management. It is therefore entirely reasonable for RELEVANCE to inform this technology appraisal, and to assume that the new technology has similar efficacy to R-chemotherapy for treatment of relapsed FL/MZL, to replace or displace standard of care with the choice dependent on the patient's disease characteristics and wishes.

• What, in your view, are the most important outcomes, and were they measured in the trials?	Progression free survival: 39.4 months for R2 in the AUGMENT trial is numerically similar to results published for R-bendamustine (34 months; long term follow-up of STiL NHL 2-2003)
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	AUGMENT reported time to next treatment as a secondary endpoint. The median TTNT was not reached at a median follow-up of 28.3 months. This is an encouraging result in a population with a median of 1 prior LOT, 33% POD24 and 16% refractory to last LOT.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not to my knowledge
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
21. Are you aware of any new evidence for the comparator treatment(s) since the	No

publication of NICE technology appraisal guidance?	
22. How do data on real-world experience compare with the trial data?	Patients enrolled on the MAGNIFY trial are similar to patients treated in the real world in terms of age, performance status, rituximab refractoriness. The AUGMENT population was possibly a slightly more favourable risk group compared to standard care in terms of slightly younger, fewer median lines of therapy and the majority but not all previously treated with rituximab. Nevertheless, efficacy and safety results from these studies are broadly similar to what I expect to see with standard treatment and R2.
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
23b. Consider whether these issues are different from issues with current care and why.	N/A
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

- Rituximab + lenalidomide offers a new and extra treatment option which could extend the treatment period and life expectancy of patients affected with relapsed/refractory FL/MZL.
- The new technology addresses important areas of unmet need including overcoming treatment resistance in patients with early relapse or heavily pre-treated chemotherapy-refractory disease.
- The new technology has a favourable toxicity profile that is suitable for patients for whom standard chemotherapy may be too toxic.
- A first line trial demonstrating similar efficacy to standard care with different but not increased toxicity is informative for management of relapsed disease, and suggests that the new technology could be used interchangeably with immuno-chemotherapy.
- The NHS is set-up to deliver the new technology with no significant resource implications.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

.....

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

Clinical expert statement

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us 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.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	Andrew Pettitt

2. Name of organisation	University of Liverpool (UoL) Clatterbridge Cancer Centre NHS Foundation Trust (CCC) National Cancer Research Institute (NCRI)
3. Job title or position	Professor of Haemato-oncology, UoL Honorary Consultant Haemato-oncologist, CCC Chair, NCRI Lymphoma Group
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☒ other (please specify): Chief investigator for two phase III clinical trials in follicular lymphoma and local PI for numerous clinical trials in follicular and marginal zone lymphoma.

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 didn't submit one, I don't 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
The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	

or prevent progression or disability.)	
8. 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.)	
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	
What is the expected place of the technology in current practice?	
10. How is the condition 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?	
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	
• In what clinical setting should the technology be	

used? (For example, primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	
12. 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?	

13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
The use of the technology	
14. 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.)	

15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	
16. 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?	
17. 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?	
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
Sources of evidence	
19. 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?	
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	
21. Are you aware of any new evidence for the comparator	

treatment(s) since the publication of NICE technology appraisal guidance?	
22. How do data on real-world experience compare with the trial data?	
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	
23b. Consider whether these issues are different from issues with current care and why.	
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

-
-
-
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

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

Patient expert statement

Lenalidomide for previously treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- 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	**Susan Christine Jones**

2. Are you (please tick all that apply):	Yes, I am a patient with the condition
3. Name of your nominating organisation	Lymphoma Action
4. Did your nominating organisation submit a submission?	Yes, they did
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

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. How did you gather the information included in your statement? (please tick all that apply)	I have personal experience of the condition
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	I was diagnosed in 2007 with Follicular Non-Hodgkin's Lymphoma. Starting in October that year I had 8 sessions of R-CVP. I was then in a partial remission for just over two years. My quality of life was good, and I had recovered sufficiently to walk six miles and to climb some hills, and even mountains. I relapsed in 2010, and in August started a course of 8 sessions of GA101 + CHOP. GA101 was later given a name - Obinotuzumab. This was part of a trial and was followed up with 8 top-ups of GA101 at 12 week intervals. For me the outcome of the trial was good. I have been in a complete remission for 9 years. I lead a busy, active and productive life. But, of course, there are side effects of all treatments. I experience peripheral neuropathy, which disturbs me at night, coughs and colds linger, and I'm inclined to get cold sores. Having said that, they are side effects that I can cope with, and I feel I am generally healthy, but I know I am immuno-compromised.

Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	
10. Is there an unmet need for patients with this condition?	
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	In 2007 R-CVP didn't complete the job for me. Perhaps that was because Rituximab top ups weren't routinely available then. I have no way of knowing, but I understand research has shown top ups of the antibody are generally beneficial to the patient. The 2010 treatment was much more effective. Either GA101 with CHOP was more effective for me, or the GA101 top ups made all the difference to me. I have no way of knowing which, or if it was the combination. Receiving the top up infusions of the GA101 without the CHOP drugs at the same time made receiving treatment much easier. I did not experience nausea or any of the other effects of chemotherapy. My body felt tired on the day of receiving the top up, but I could carry on with my normal activities the following day. One of the advantages of the proposed new treatment for me is that it doesn't include the use of chemotherapy drugs as in CHOP or CVP. Another advantage is that Lenalidomide helps the immune system to do its job of attacking the cancerous cells. It is also in tablet form and can be taken at home. I think tablets are generally better than infusions, and presumably it is cheaper to administer if the patient does not need to be in hospital.

	I would like to think that if I relapse, as seems to be expected, that this type of treatment would be available to me.
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	
Patient population	
13. 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	
14. Are there any potential equality issues that should be taken into account when	

considering this condition and the technology?	
Other issues	
15. Are there any other issues that you would like the committee to consider?	
Key messages	

17. In up to 5 bullet points, please summarise the key messages of your statement:

- This treatment does not use the standard chemotherapy drugs.
- Antibodies (in my experience) do not have such a draining effect on the patient.
- Lenalidomide helps the immune system to work.
- Lenalidomide is administered in tablet form.
- The patient does not need to be in hospital to receive Lenalidomide.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

.....

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

Patient expert statement

Lenalidomide for previously treated follicular lymphoma and marginal zone lymphoma [ID1374]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- 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	**Dr Peter Loftus**

2. Are you (please tick all that apply):	x☐ a patient with the condition? ☐ a carer of a patient with the condition? ☐ a patient organisation employee or volunteer? ☐ other (please specify):
3. Name of your nominating organisation	Lymphoma Association
4. Did your nominating organisation submit a submission?	x☐ yes, they did ☐ no, they didn't ☐ I don't know
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)	x☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't 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.)	**x**☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	**x**☐ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	

Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	
10. Is there an unmet need for patients with this condition?	
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	
Patient population	
13. 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	
14. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	
Other issues	
15. Are there any other issues that you would like the committee to consider?	
Key messages	
17. In up to 5 bullet points, please summarise the key messages of your statement: • •	

-
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Rob Riemsma, Reviews Manager, Kleijnen Systematic Reviews Ltd, UK
Thea van Asselt, Health Economist, Maastricht UMC, Netherlands
Willem Witlox, Health Economist, Maastricht UMC
Vanesa Huertas Carrera, Systematic Reviewer, KSR Ltd
Pawel Posadzki, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Steve Ryder, Health Economist, KSR Ltd
Bram Ramaekers, Health Economist, Maastricht UMC
Sabine Grimm, Health Economist, Maastricht UMC
Xavier Pouwels, Health Economist, Maastricht UMC
Gill Worthy, Statistician, KSR Ltd
Steven Duffy, Information Specialist, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 25/09/2019

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 12/82/01.

Declared competing interests of the authors

None.

Acknowledgements

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Riemsma R, Van Asselt ADI, Witlox W, Huertas Carrera V, Posadzki P, Armstrong N, Ryder S, Ramaekers B, Grimm S, Pouwels X, Duffy S, Joore MA, Kleijnen J. Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2019.

Contributions of authors

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Thea van Asselt acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Willem Witlox, Bram Ramaekers, Sabine Grimm, Xavier Pouwels, Steve Ryder and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Vanesa Huertas Carrera and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
ADCP	Antibody-dependent cellular phagocytosis
AE	Adverse events
AIC	Akaike information criterion
ASCO	American Society of Clinical Oncology
ASCT	Autologous stem cell transplantation
AWMSG	All Wales Medicines Strategy Group
bd/b.i.d	Twice daily
Benda	Bendamustine
BI	Budget impact
BIC	Bayesian information criterion
BSA	Body surface area
BSC	Best supportive care
CDF	Cancer Drugs Fund
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CHOP	Cyclophosphamide, doxorubicin, vincristine, prednisolone
CI	Confidence interval
CIC	Commercial in confidence
CR	Complete response
CrCl	Creatinine clearance
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
Cru	Complete response unconfirmed
CS	Company's submission
CSR	Clinical study report
CT	Computerised tomography;
CVP	Cyclophosphamide, vincristine, prednisone
DCRR	Durable complete response rate;
DHAP	Dexamethasone, cytarabine, cisplatin
DMC	Data monitoring committee
DOCR	Duration of complete response
DOR	Duration of response
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EM	Effect modifiers
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EPAR	European public assessment report
EQ-5D	European quality of life-5 dimensions
ERG	Evidence Review Group
ESHAP	Etoposide, methylprednisolone, cytarabine, cisplatin
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
FACT-Lym	Functional assessment of cancer therapy: lymphoma
FBC	Full blood count
FDA	Food and Drug Administration
FL	Follicular lymphoma
FLIPI	Follicular lymphoma international prognostic index

G-CSF	Granulocyte colony-stimulating factor
GELF	Groupe d'Etude des Lymphomes Folliculaires
H	Helicobacter
HMRN	Haematological Malignancy Research Network
HR	Hazard ratio
HRQL	Health-related quality of life
HT	Histological transformation
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
IEE	Induction efficacy population
Ig	Immunoglobin
IITT	Induction intention-to-treat population
iNHL	Indolent non-Hodgkin's lymphoma
IPD	Individual patient data
IRC	Independent review committee
ITC	Indirect treatment comparison
ITT	Intention to treat
IV	Intravenous
IVRS	Interactive voice response system;
IWGRC	International Working Group Response Criteria
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
LDH	Lactate dehydrogenase
LFT	Liver function tests
MAIC	Matching-adjusted indirect comparison
MALT	Mucosa-associated lymphoid tissue lymphoma
MCL	Mantle cell lymphoma
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MIMS	Monthly Index of Medical Specialities
mITT	Modified intention-to-treat
MRI	Magnetic resonance imaging
MZL	Marginal zone lymphoma
NA	Not applicable
NHL	Non-Hodgkin's lymphoma
NHS	National Health Services
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NK	Natural killer
NSAID	Non-steroidal anti-inflammatory drug
NZML	Nodal marginal zone lymphoma
O	Obinutuzumab
O-Benda	Obinutuzumab in combination with bendamustine
ONS	Office for National Statistics
OR	Odds ratio
ORR	Overall response rate
OS	Overall survival
OWSA	One-way sensitivity analysis
PAS	Patient access scheme
PD	Progressive disease;
PFS	Progression-free survival
PPS	Post-progression survival
PR	Partial response
PRESS	Peer review of electronic search strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSA	Probabilistic sensitivity analyses
PSSRU	Personal Social Services Research Unit
PV	Prognostic variables
QALY(s)	Quality-adjusted life year(s)
QoL	Quality of life
R^2	Lenalidomide plus rituximab
R-Benda	Rituximab in combination with bendamustine
R-chemo	Rituximab plus chemotherapy
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, prednisolone
R-CVP	Rituximab, cyclophosphamide, vincristine, prednisone
R-mono	Rituximab monotherapy
RCT	Randomised controlled trial
RS	Relative survival
RTNLT	Response to next anti-lymphoma treatment
RWE	Real world evidence
SAE	Serious adverse events
SC	Subcutaneous
ScHARR	School of Health and Related Research
SCT	Stem cell transplantation
SD	Stable disease/Standard deviation
SE	Standard error
SHTAC	Southampton Health Technology Assessments Centre
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SMZL	Splenic marginal zone lymphoma
STA	Single technology appraisal
TEAE	Treatment-emergent adverse event
TFR	Tumour flare reaction
TLS	Tumour lysis syndrome
ToT	Time on treatment
TSD	Technical support document
TTNLT	Time to next anti-lymphoma treatment
TTP	Time to progression
TTR	Time to response
UK	United Kingdom
WHO	World Health Organisation
WTP	Willingness to pay

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The population defined in the scope is: Adults with previously treated follicular lymphoma or marginal zone lymphoma. The population in the company submission (CS) is in line with the NICE scope.

According to the company lenalidomide plus rituximab (R^2) does not currently have a UK marketing authorisation, although the Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated on [redacted], and marketing authorisation is expected in [redacted]. Therefore, the relevant population for this appraisal is currently unclear. The anticipated license is as follows: Lenalidomide in combination with rituximab (anti-CD20 antibody) is indicated for the treatment of adult patients with previously treated Follicular Lymphoma (FL) or Marginal Zone Lymphoma (MZL).

The intervention (lenalidomide in combination with rituximab) is in line with the scope.

The description of the comparators in the NICE scope is as follows: rituximab monotherapy, rituximab in combination with chemotherapy, and established clinical management without lenalidomide (including but not limited to bendamustine). The NICE scope does not make a distinction in terms of patients being rituximab refractory or not. However, the CS has different comparators for rituximab refractory patients and non-rituximab refractory patients. The company's justification for this approach is because 'patients determined to be R-refractory are treated differently to the non-R-refractory population in the UK due to the availability of an alternative treatment, obinutuzumab-bendamustine, approved in the EU and recommended by NICE'

For non-rituximab refractory patients, the company included two comparators, both different types of rituximab in combination with chemotherapy:

- Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP)
- Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP)

For rituximab refractory patients, the company included one comparator:

- Established clinical management without lenalidomide, i.e. obinutuzumab in combination with bendamustine (O-Benda).

The ERG has several concerns with these comparators. Firstly, the NICE scope does not make a distinction in terms of patients being rituximab refractory or not. Therefore, the CS should have included a comparison of R^2 with rituximab monotherapy for all patients as specified in the scope. Secondly, even if the NICE committee accepts splitting the population in rituximab refractory patients and non-rituximab refractory patients, the CS should still have included a comparison with rituximab monotherapy for both populations as specified in the scope. Thirdly, the company included O-Benda as a comparator for rituximab refractory patients. However, in the response from NICE to comments on the draft scope, NICE clearly stated that "obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4)). Therefore, we believe that the submission currently does not present any relevant evidence for R-refractory patients.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company submission included six studies that were deemed relevant by the company. Four studies evaluated R^2, one of these was a randomised controlled trial (RCT) of R^2 versus R-monotherapy (the AUGMENT trial), the other three did not include relevant comparators according to the NICE scope.

The remaining two studies evaluated R-CHOP versus cyclophosphamide, doxorubicin, vincristine, and prednisolone (CHOP) (Van Oers et al., 2006) and O-Benda versus bendamustine monotherapy (the GADOLIN trial). The trial by Van Oers et al. (2006) was used by the company for an unanchored indirect comparison (using individual arms of different studies) of R^2 versus R-CHOP. However, the study only included rituximab-naïve patients and was therefore not representative for the UK patient population. The GADOLIN study was used by the company for an unanchored indirect comparison of R^2 with O-Benda. However, as explained in Sections 3.3 and 4.4 of this report, O-Benda is not considered by NICE to be a relevant comparator for this appraisal; therefore, this study has been ignored in this report.

In conclusion, the CS included one relevant study, for the comparison of R^2 versus R-monotherapy: the AUGMENT trial. All patients in this trial were non-R-refractory. In addition, the company performed an unanchored indirect comparison of R^2 versus R-CHOP and R-CVP, using data for R^2 from the AUGMENT trial and pooled data for R-CHOP/R-CVP from the Haematological Malignancy Research Network (HMRN) database.

The AUGMENT trial is a randomised, double-blind, multicentre, Phase III study of R^2 versus rituximab plus placebo (R-mono) in non-R-refractory patients with FL Grade 1–3a or MZL. The study was conducted across 96 sites in 17 countries. The trial did not include any patients from the UK. The primary efficacy analyses were conducted on the ITT population, defined as all randomised patients. The primary endpoint of the study was progression-free survival (PFS), as assessed by the Independent Review Committee (IRC).

Results from the AUGMENT trial show favourable results for R^2 when compared to R-mono in terms PFS with a greater median PFS ([redacted] vs. [redacted] months; hazard ratio (HR) of [redacted] (95% confidence interval (CI): [redacted]). However, there was no evidence of a difference in overall survival (OS) with a HR of 0.61 (95% CI: 0.33 to 1.13) for patients treated with R^2 compared to R-mono. At the time of the analysis the overall survival (OS) data were immature with 16 deaths on R^2 and 26 deaths on R-mono. Overall response rate (ORR) was significantly greater for R^2 compared with R-mono (78% vs. 53%; $p<0.0001$). The complete response (CR) rate was also greater for the R^2 arm compared with R-mono (34% vs. 18%; $p=0.001$). Results for R^2 versus R-mono in MZL patients were generally less favourable for R^2 than in FL patients. However, it is important to note that PFS outcomes in the MZL subgroup are difficult to interpret because of the small sample size (63 patients in total) and imbalance in baseline prognostic factors. In terms of health-related quality of life, no clinically meaningful change from baseline in the global health status/quality of life (GHS/QoL) domain of the EORTC Quality of Life Questionnaire, Core 30 (QLQ-C30) was observed across any of the post-baseline assessment visits, regardless of treatment group. Between-group differences in mean changes were small and not clinically meaningful across all assessment visits and did not differ between FL and MZL patients.

R^2 was associated with more grade 3-4 treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs) when compared to R-mono, especially lenalidomide/placebo related adverse events; but rituximab-related grade 3-4 TEAEs and SAEs were also more frequent in the R^2 arm than in the R-mono arm. R^2 was also associated with more TEAEs leading to dose reductions, dose interruptions and discontinuations of lenalidomide/placebo or rituximab when compared to R-mono. Adverse events are generally the same for FL and MZL patients; however, AEs for MZL patients are based on small numbers.

The company performed three unanchored indirect comparisons, two using data from published evidence and one using data from HMRN:

- R^2 versus R-CHOP for non-rituximab refractory patients, based on comparator data from a study by Van Oers et al. 2006 comparing R-CHOP with CHOP (only the R-CHOP arm was used in the analyses).
- R^2 versus established clinical management without lenalidomide, i.e. O-Benda for rituximab refractory patients, based on comparator data from a study by Sehn et al. (2016) comparing O-Benda with bendamustine monotherapy (only the O-Benda arm was used in the analyses).
- R^2 versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN.

As mentioned above, the two unanchored indirect comparisons using published evidence have been ignored in this report. R^2 versus R-CHOP, because the study by Van Oers is not representative for UK patients, and R^2 versus O-Benda because O-Benda is not a relevant comparator for this appraisal according to NICE.

Results from the remaining matching-adjusted indirect comparison (MAIC) (R^2 versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN) show a significant improvement for R^2 in OS (HR = [redacted] (95% CI: [redacted]) and time to next anti-lymphoma treatment (TTNLT, HR = [redacted] (95% CI: [redacted]) compared to R-CHOP/R-CVP, but no evidence of a difference in PFS (HR = [redacted] (95% CI: [redacted]).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches conducted as part of the systematic review to identify clinical effectiveness studies. A good range of databases and resources were searched. The searches did not include study design filters in order to identify both efficacy and safety evidence. Searches conducted in September 2017 were reported, but need not have been as they were subsequently replaced by searches conducted in April 2019.

The results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were excluded from the matching models, small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources, The analysis also used an unanchored MAIC involving two single treatment arms from different studies, as there was no relevant comparative trial data. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

1.4 Summary of cost effectiveness evidence submitted by the company

The company conducted searches for cost effectiveness, health-related quality of life and healthcare resource use evidence.

The company developed a cohort-level partitioned survival model (PSM), programmed in Excel, with three health states: progression-free (PF), post-progression (PP) and death. All patients start 'on treatment' in the PF health state. Subsequently, patients either remain on treatment or come off treatment before progressing or dying per cycle. Within PP, patients can have a treatment-free interval before receiving subsequent therapy. Patients in the PP on treatment health state remain in this health state until they die. The time horizon was lifetime and cycle length 28 days. R^2 does not currently have a UK marketing authorisation, but the patient population considered in the model is in line with the proposed license: adult patients with previously treated FL or MZL. Due to the similar prognosis of FL

and MZL patients, and the difficulty in sourcing MZL-specific data, FL and MZL populations were pooled throughout the economic analysis. The R^2 dosing regimen within the model is lenalidomide 20 mg orally once daily on days 1–21 of repeated 28-day cycles for up to 12 cycles of treatment. Rituximab is given as 375 mg/m2 every week in Cycle 1 (days 1, 8, 15 and 22) and day 1 of every 28-day cycle for Cycles 2–5. This is in line with the recommended dose in the summary of product characteristics (SmPC). Based on expert opinion, the company compared R^2 in the non-R-refractory population with R-CHOP and R-CVP, and in the R-refractory population to O-Benda.

The main source of evidence on treatment effectiveness used for intervention and comparators was the AUGMENT study for R^2 and HMRN data for R-CHOP and R-CVP. The AUGMENT study contained a mixed MZL/FL population, HMRN contained only data on FL patients. The company assumed efficacy of R-CHOP and R-CVP to be similar, therefore HMRN data for R-CHOP and R-CVP were pooled. For the economic model, this implied that the comparisons of R^2 vs. R-CHOP and R-CVP had identical outcomes for effectiveness and only differed with respect to costs.

Parametric survival curves were fitted to the matched patient level data from AUGMENT and HRMN and were then used to extrapolate survival beyond study follow-up. Survival analysis was performed for OS, PFS, TTNLT, and ToT (time on treatment). PFS and ToT data were used to determine the number of patients staying in the PF (on and off treatment) health states. PFS, TTNLT and OS data were used to determine the number of patients transitioning to the PP (on and off treatment) health states. The number of patients transitioning to the death state was derived using OS data. The curves were adjusted for treatment waning, which was assumed to occur at five years. After this time point, the comparator hazard of progressing or dying was applied to the R^2 arm. Any implausible curve crossings (for instance, OS crossing PFS) were corrected for.

For the R^2 versus R-CHOP and R-CVP comparisons, the company selected a Weibull distribution to extrapolate OS, mainly based on a previous single technology appraisal (STA). For the R-mono comparison, which was added upon request of the ERG in the response to clarification, the company chose Weibull for OS as well.

The company decided to model the PFS for R^2 versus R-CHOP/R-CVP using the Kaplan–Meier (KM) data until the maximum follow-up of 46.7 months, and applied the comparator hazard to extrapolate further. In this way, the company stated in the CS, the relative treatment effect of R^2 vs. R-CHOP/R-CVP based on the MAIC was accurately reflected. For the R-CHOP/R-CVP arm, a generalised gamma was chosen. For the R-mono comparison, a simpler approach was taken, using log-logistic distributions for both arms.

Based on the Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), the exponential distribution best fitted the R^2 data, and the log-normal distribution fitted best to R-CHOP/R-CVP. However, as the exponential distribution would result in crossing of PFS and TTNLT around seven years, the company chose the log-normal distribution for the base-case analysis for both arms. For the R-mono comparison, the generalised gamma was used for both arms.

ToT data were used to determine the proportion of patients on treatment to calculate overall drug costs. Parametric survival curves were fitted to the ToT data which, however, produced a poor fit. Therefore, the company chose to use the KM data directly in the model, and maximum treatment durations were used to cap ToT. For the R-mono comparison the same approach was used.

The main sources of evidence on treatment-related adverse events used for intervention and comparators were the AUGMENT and RELEVANCE trials, because of a lack of safety data from

HMRN. RELEVANCE is a Phase III study comparing R^2 with R-chemotherapy for patients with previously untreated FL. AUGMENT was used for R^2, and RELEVANCE was used for R-CHOP and R-CVP, after adjusting for any possible differences in R^2 AEs between AUGMENT and RELEVANCE. In a scenario, AEs for R-CHOP/R-CVP were taken from van Oers et al. (2006) which concerned a relapsed/refractory population. Furthermore, AE incidence for maintenance treatment and autologous stem cell transplant (ASCT) were also considered.

Utility values were estimated for the health states PF, and PP off and on treatment using European quality of life-5 dimensions-3 level (EQ-5D-3L) data collected in AUGMENT. A covariate selection process was used to select the appropriate mixed effects utility model as input for the economic model. The utility values resulting from the mixed effects model were used to inform the health states in the model for all treatments, and utility values from the literature were tested in scenario analyses. Disease characteristics that were used to derive utility values from the mixed effects model were population-dependent, and therefore, the utility values for R^2 versus R-CHOP/R-CVP and R^2 versus R-mono were slightly different. The mean utility values for post-progression based on the AUGMENT trial data were higher than values from the studies identified in the systematic literature review (SLR). Utility decrements for grade 3 and 4 AEs were applied in the model for the expected duration of each AE, based on literature and previous appraisals.

The cost categories included in the model were costs associated with treatment (drug acquisition costs including subsequent therapies, drug administration costs including subsequent therapies, costs associated with treatment-related AEs), disease monitoring costs and costs associated with end of life care. For lenalidomide, dosing data had been taken directly from AUGMENT (non-R-refractory population) to align the drug costs with the efficacy data because according to the company, dose reductions for lenalidomide can occur. In the economic model, the company applied ASCT to [redacted]% of patients in R-CHOP. For R-CVP and R^2, 0% ASCT was applied as it was considered unlikely that these treatments would be used as an induction regimen prior to ASCT. Subsequent treatments were included in the model as an average one-off cost to patients entering the PP (on treatment) health state, derived using TTNLT data. Costs for patients in the R^2 arm were derived from subsequent treatments from AUGMENT. The total subsequent treatment data from the pooled R-chemotherapies in the HMRN database were used for R-CHOP and R-CVP.

Total life years (LYs) and quality adjusted life years (QALYs) gained and total costs were larger for R^2 than for R-CHOP, R-CVP an R-mono. The incremental cost effectiveness ratio (ICERs) amounted to respectively £11,471, £16,814 and £22,580 per QALY gained. Compared with the deterministic results, the PSA with 1,000 iterations showed lower incremental QALYs and costs, which resulted in increased ICERs of £13,443 and £20,896 and £26,116 respectively. The cost effectiveness acceptability curves in the economic model showed that R^2 respectively had a 82%, 72% and 69% probability of being cost effective at a willingness-to-pay (WTP) threshold of £30,000. Deterministic sensitivity analyses (DSAs) were performed by varying key model parameters to the upper and lower limits of their respective confidence intervals, but in none of these analyses the ICER exceeded the £30,000 threshold.

The company performed internal validity checks using AdVISHE and made face validity checks on model structure and other assumptions within an advisory board. External validation with data from AUGMENT showed that PFS, OS and TTNLT at one year for patients treated with R-CHOP/R-CVP were under-estimated in the model compared with the observations.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

Separate sets of searches were conducted to identify cost effectiveness studies, health-related quality of life studies and healthcare resource use evidence. The CS provided clear, transparent and reproducible searches. A good range of databases and additional resources were searched.

The company submission was largely in line with the NICE reference case. The CS did partly deviate from the scope, however, where it concerned the comparators modelled.

The company used a PSM instead of a state transition model (STM), justified by a lack of data for relevant comparators. Although the ERG recognises the potential limitations of a STM, a PSM has several limitations related to the extrapolation, as mentioned in NICE DSU TSD 19. The ERG requested a scenario analysis using a STM as a scenario, as recommended in TSD 19, which the company did not deliver. The company clarified that while FL and MZL populations were pooled, all evidence of the comparators was based on datasets that only contained patients with FL, while the AUGMENT trial contained patients with FL and MZL. In response to questions from the ERG, the company provided additional analyses on AUGMENT trial data that showed the impact of histology on the results were not statistically significant. The company provided a FL-only scenario analysis upon request of the ERG. The ERG also requested an analysis with R-mono as a comparator, as listed in the final scope, which the company provided. O-Benda was not included in the ERG report as NICE has explicitly stated it is not considered a relevant comparator for disease that is R-refractory.

A main concern of the ERG was the questionable trustworthiness of R^2 efficacy resulting from the indirect comparison, which seemed to be inflated relative to the direct comparison data from AUGMENT. Although the ERG did not have the necessary data to quantify this uncertainty, it may have lowered the ICER substantially, favouring R^2.

The ERG had concerns about the way survival curves were selected. Although the company proposed a systematic approach of selecting the appropriate curves, there were many deviations from this systematic approach in the actual selection process. The choice of OS curve was mainly based on a previous STA (TA137: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma). In particular the choice of PFS curves was not sufficiently justified and appeared sub-optimal, with a likely overestimation of PFS in the R^2 arm, and substantial underestimation of PFS in the first year for R-CHOP and R-CVP. This matter was exacerbated by the high utility values for all health states. The ERG considered these to be potentially overestimated, being higher than or comparable to those in the general population. With utilities remaining high throughout the model, any adjustment in survival curves only had little impact on the ICER, as a high utility post-progression implied there was hardly any penalty on progression in terms of cost effectiveness. This was demonstrated in the ERG scenario analyses, where the ERG base-case in combination with a lowered utility score post-progression had the highest impact on the ICER, increasing it to £33,626 and £47,281 for R^2 versus R-CHOP and R-CVP respectively, when using the lowest value from the literature for the post-progression utilities.

The ERG questioned the applicability of AE incidences taken from a previously untreated population for the present STA, and feels it is important to seriously consider the scenario provided by the company with data from a relapsed/refractory population. Therefore, the ERG included this as one of their scenarios. Also, the ERG considered it to be inconsistent that AEs related to subsequent ASCT and R-mono therapy were only taken into account for R-CHOP and R-CVP and so this was corrected for in the ERG base-case.

The ERG also had concerns about the high utility values for the PF and PP health states, and the modest utility decrement for disease progression. Utility values for the PF and PP health states were higher than the utility reported for the general population, which seems quite unlikely in patients with treated FL or MZL. Furthermore, the ERG judges that a larger utility difference between PF and PP health states would be more plausible, and explored this in a scenario analysis using lowered utility values taken from published studies for both PP health states. For R^2 versus R-CHOP and R-CVP, this substantially increased the ICER, while for R^2 versus R-mono the ICER decreased.

The ERG questioned the company's choice to include subsequent treatments as a one-off cost to those patients entering the PP on treatment health state. The company costed for observed incidences of subsequent treatments from the data sources, which for R^2 had a much shorter follow-up than for R-CHOP/R-CVP and therefore may not be reflective of clinical practice. Furthermore, subsequent treatment costs for R-CHOP and R-CVP were, in contrast to the treatment effectiveness, calculated based on the pooled R-chemotherapies data from HMRN instead of the HMRN R-CHOP/R-CVP cohort. The ERG changed this in its base-case but the impact on the ICER was modest. In addition, the company assumed the percentage of post-induction (but pre-progression) ASCTs in R^2 to be zero, because it was not protocolised in AUGMENT and clinicians considered it unlikely that patients would receive ASCT post R^2. The ERG would have liked to see a scenario using observed frequencies, as clinical practice may sometimes contrast with protocols and clinical opinion. A non-zero observed frequency would increase the ICER for R^2 versus R-CHOP.

The ERG had some comments about the PSA, which did not enable a fully incremental analysis for more than two comparators, nor representation of multiple comparators in the cost effectiveness acceptability curve (CEAC). Furthermore, probabilistic QALYs were lower compared to the deterministic QALYs in the company base-case, likely caused by non-linearity of the model. An additional scenario analysis for the FL-only population was provided by the company in response to clarification, resulting in ICERs of £15,909 and £23,746 for the R-CHOP and R-CVP comparisons, respectively, making it the most influential scenario. For the R-mono comparison, using FL-only data lowered the ICER to £20,310.

Internal validation of the model was performed to a good standard. It is not clear whether all assumptions and extrapolations (notably for PFS, OS and TTNLT for patients treated with R-CHOP/R-CVP) were validated by experts.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

A good range of resources were searched and the searches were well documented making them transparent and reproducible. Supplementary searches of conference proceedings and HTA organisation websites were undertaken, along with a search of the ClinicalTrials.gov register in order to identify additional trials.

The company submission was largely in line with the NICE reference case. Utility scores were estimated using a mixed effects model based on observed EQ-5D data in the AUGMENT study.

The model was, in general, well-built and transparent. Apart from the base-case, the model provided ample opportunity for exploratory analyses using alternative assumptions on a range of input parameters.

1.6.2 Weaknesses and areas of uncertainty

The ERG was concerned about the overall quality of the searches conducted, as truncation and proximity operators were used inconsistently, and more synonyms could have been included in the search strategies. The date ranges of searches were not accurately reported. However, the searches were adequate, and given the range of resources searched, it was unlikely that any relevant studies were missed.

The results of the MAIC should be treated with a high degree of caution.

Similarly, the results of the economic evaluation should be treated with a high degree of caution, as the results of the MAIC serve as an important input parameter for the economic model. As the ERG did not have the necessary data to quantify uncertainty around the MAIC, model results do not include this structural uncertainty. Therefore, not only the company base-case, but also the ERG base-case and further exploratory analyses all (except those for the R-mono comparison) are conditional upon the possibly biased effectiveness of R^2 versus R-CHOP/R-CVP resulting from the MAIC. The ERG considers this to be a major source of uncertainty.

A main limitation was the lack of clarity and consistency in the selection of the parametric survival curves for extrapolation of PFS, OS and also TTNLT. The ERG considers particularly PFS to be overestimated for R^2 and underestimated (in the first year) for R-CHOP and R-CVP. Curve selection was often based only on avoiding implausible curve crossings, which may be indicative of a structural issue in the model design. For reference, the ERG would have liked to see the results of a state transition model next to the current partitioned survival model, but the company did not provide this.

Given the large impact of the FL-only scenario on the ICERs of the R-CHOP and R-CVP comparisons, the ERG considers the pooling of MZL and FL populations throughout the analysis to be another substantial source of uncertainty.

Lastly, the utility scores used in the model do not seem representative of the patient population. The ERG considers utilities in both progression free and progressed health states to be an overestimate.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG made various adjustments to the company's base-case, including the fixing of errors, violations, and amending the model according to the company's base-case according to its preferred assumptions (matters of judgement).

Fixing errors

1. Error cells when using 'van Oers' as input for R-CHOP efficacy

Fixing violations

2. Include AEs related to subsequent treatments in R^2 arm
3. Use pooled R-CVP/R-CHOP subsequent treatment rates instead of R-chemo. (Not applicable in the R-mono comparison)
4. Cap utilities at the general population level

Matters of judgment

5. Use exponential distribution to extrapolate OS in both arms
6. Use log-logistic for PFS in R^2 and Weibull for PFS in the comparator (not applied to R-mono comparison)
7. Used log-logistic for TTNLT both arms (not applied to R-mono comparison)

1.7.1 ERG probabilistic base-case results

The probabilistic ERG base-case ICER of R^2 versus R-CHOP was £15,818 per QALY gained (based on 1,000 iterations). This was slightly higher than the deterministic base-case ICER of £15,505. For R^2 versus R-CVP, the probabilistic ICER was £23,367 (deterministic £21,759) and for R^2 versus R-mono it was £29,010 (deterministic £27,372) (See Table 1.1). These rather substantial differences between probabilistic and deterministic ICERs were also observed in the company analyses (to a larger extent even).

Table 1.1: ERG probabilistic base-case results

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case for R^2 versus R-CHOP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,818
R-CHOP	[redacted]	[redacted]			
Probabilistic ERG base-case for R^2 versus R-CVP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,367
R-CVP	[redacted]	[redacted]			
Probabilistic ERG base-case for R^2 versus R-mono					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,010
R-mono	[redacted]	[redacted]			

ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. For the R-CHOP/R-CVP comparisons, using R-CHOP and R-CVP efficacy from van Oers et al. would lower the ICER substantially, £8,251 for R^2 versus R-CHOP and £13,315 for R^2 versus R-CVP. Alternative assumptions regarding lowered utilities in the PP health states had the most significant upward impact, increasing the ICER to £33,626 for R^2 versus R-CHOP and £47,281 for R^2 versus R-CVP. For the R-mono comparison, lowering the PP health state utility had the opposite effect, lowering the ICER to £17,826. Another influential scenario was the change of time-point where treatment waning starts to three years (instead of five years in base-case). This increased the ICER to £40,543.

In conclusion, even though the ERG base-case ICER for R-CHOP was below £20,000, the uncertainty around the cost effectiveness of R^2 is substantial, mainly caused by the possible bias introduced by the indirect treatment comparison, which could not be accounted for in the ERG analyses. The ICER for R-CVP is higher and suffers from the same uncertainty. The R-mono analysis is based on a direct comparison, but is also surrounded by substantial uncertainty, as the ICER is rather sensitive to, for instance, the time-point at which treatment waning starts and utilities in the PP health state.

2. BACKGROUND

In this report, the ERG provides a review of the evidence submitted by Celgene in support of lenalidomide (Revlimid®) in combination with rituximab (MabThera®) (R2), for the treatment of adults with treated follicular lymphoma (FL) or marginal zone lymphoma (MZL).

We will outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken from section B.1.3 of the company's submission (CS) with sections referenced as appropriate. For additional information on the aetiology, epidemiology, health impact, prognosis and management of FL or MZL, please see the CS (pages 13-23).[1]

2.1 Critique of company's description of underlying health problem.

The underlying health problems in this appraisal are follicular lymphoma (FL) and marginal zone lymphoma (MZL), the two most common subtypes of indolent non-Hodgkin lymphoma (iNHL).

As described in the CS, FL is typically characterised by an indolent clinical course, with recurrent remissions and relapses; with each relapse, the disease becomes more resistant and/or refractory to treatment and each remission becomes shorter than the preceding one.[1] The incidence of FL increases with age, with a median presentation between 60 and 65 years, and a slightly higher incidence in females.[2] At diagnosis, most patients have advanced disease (Stage III: 18.4%; Stage IV: 50.5%).[3] The overall five-year relative survival rate for patients with FL in the UK is 89% and specifically for Stages III and IV, is approximately 80%.[4, 5] Since the introduction of rituximab, the median OS of patients with FL has extended to 20 years in some studies,[6] compared with nine years previously reported.[7] Despite the available treatment options, most patients eventually die from this disease.[8]

Patients with MZL represent a generally older (median age at diagnosis is 70–73 years)[2] and more advanced population compared with those with FL.[2, 3, 9] The primary organ of origin is the most significant prognostic factor and dictates organ-specific management strategies.[9] Patients with MZL have a similar prognosis to those with FL. In the UK, the overall five-year survival ranges between 77% and 90% depending on the subtype of MZL.[4] The median OS for UK patients has been reported as between eight and 12.6 years, depending on the subtype of MZL.[10, 11]

For FL, the CS notes that the Office of National Statistics (ONS) estimates 2,168 patients were diagnosed with FL in 2017 in England.[12] Of these, [redacted]% (n=[redacted]) have first-line chemotherapy, while [redacted]% (n=[redacted]) undergo a 'watch and wait' approach,[3] of which [redacted]% (n=[redacted]) go on to receive chemotherapy.[13] Therefore, the total number of FL patients on first-line chemotherapy is [redacted]. Of these, [redacted]% (n=[redacted]) are expected to receive second-line chemotherapy or beyond.[13] For MZL, the CS states that based on the anticipated figures for the different MZL types, the total number of MZL patients in England in 2017 is estimated at 1,411.[14-16] Of these, 34.9% (n=492) have first-line chemotherapy, while 49.9% (n=704) undergo a 'watch and wait' approach[3] of which [redacted]% (n=[redacted]) go on to receive chemotherapy.[13] Therefore, the total number of MZL patients on first-line chemotherapy is [redacted]. Of these, [redacted]% (n=[redacted]) are expected to receive second-line chemotherapy or beyond.[13] The ERG has no reason to doubt these numbers.

2.2 Critique of company's overview of current service provision

In the CS, lenalidomide is described as an agent that binds to cereblon in the Cullin-4 RING E3 ubiquitin ligase that promotes the degradation of the haematopoietic transcription factors Ikaros and Aiolos.[17, 18] As a result lenalidomide inhibits proliferation and enhances apoptosis of certain haematopoietic tumour cells (including FL and MZL tumour cells), enhances T cells and natural killer (NK) cell-mediated

immunity and increases the number of NK, T and NK T cells. Single agent lenalidomide reactivates dysfunctional T and NK cells from FL patients.[17] Rituximab is an anti-CD20 antibody; its mechanisms of action are to augment NK cell-mediated killing of malignant B cells via antibody-dependent cellular cytotoxicity (ADCC), to enhance antibody-dependent cellular phagocytosis (ADCP) and to induce complement-mediated killing.[17] The combination immunotherapy of lenalidomide and rituximab acts by complementary mechanisms including direct tumour apoptosis in FL and MZL and immune-mediated activities, such as activation of NK cells and immune synapse formation, resulting in increased ADCC in vitro.[18]

The CS describes the following sources that were used in the company's interpretation of the positioning of R^2 in the treatment pathway for FL (see Figure 2.1): an advisory board conducted by Celgene in March 2019, involving six UK clinical experts in NHL and two health economics experts,[19] ad-hoc follow up with advisors, technology appraisal 472 (TA472, Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab),[20] TA243 (Rituximab for the first-line treatment of stage III-IV follicular lymphoma),[21] TA226 (Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma),[22] TA513 (Obinutuzumab for untreated advanced follicular lymphoma),[23] NICE treatment pathway for FL[24] and NICE guideline for the diagnosis and management of non-Hodgkin's lymphoma (NG52).[25]

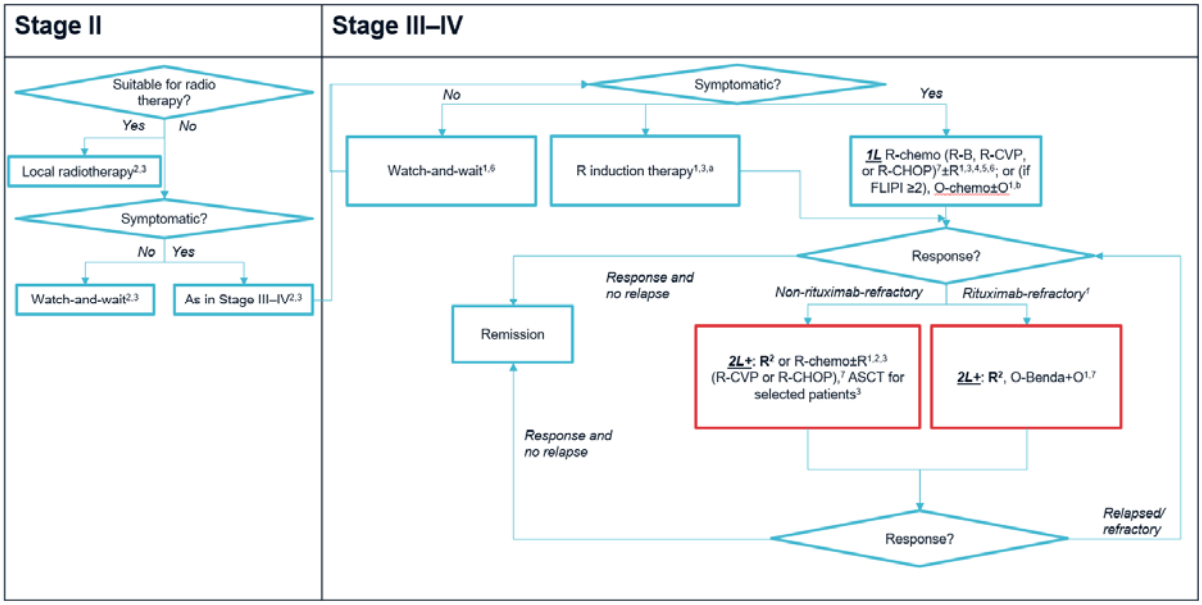
Figure 2.1 shows the treatment pathway proposed by the company for patients with follicular lymphoma. The flowchart distinguishes between Stage II and Stages III and IV. For Stage II FL radiotherapy is advised as a first-line option when suitable, when radiotherapy is unsuitable 'watch and wait' should be the preferred approach for asymptomatic patients, while symptomatic patients should be treated as in Stages III and IV. For asymptomatic patients with advanced Stages III and IV a 'watch and wait' approach or rituximab induction therapy are recommended in first-line. If patients present with symptoms, pharmacological therapy is recommended in first-line (i.e. rituximab-chemotherapy (R-bendamustine, R-CVP or R-CHOP) with or without rituximab maintenance therapy). For patients with a Follicular Lymphoma International Prognostic Index (FLIPI) score of 2 or more obinutuzumab-chemotherapy with or without obinutuzumab maintenance therapy may be given as first-line therapy.

Second-line therapy depends on whether patients are refractory to rituximab or not, according to the company's proposed pathway. The ERG has questioned this in the clarification letter (Clarification letter, Question A.7).[26] To the ERG it seems counter intuitive that rituximab containing treatments are not appropriate for rituximab-refractory patients, but rituximab in combination with lenalidomide is. The company stated that this was done to 'reflect the current approach to patient management in the UK' and that 'patients determined to be R-refractory are treated differently to the non-R-refractory population in the UK due to the availability of an alternative treatment, obinutuzumab-bendamustine, approved in the EU and recommended by NICE' (Response to Clarification Letter).[26] However, obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund; this means 'there is significant remaining clinical uncertainty which needs more investigation, through data collection in the NHS or clinical studies'.[27] This also means that it is not considered a relevant comparator for disease that is refractory to rituximab by NICE.[28]

Depending on the response to first-line therapy, patients who are not refractory to rituximab may be given rituximab-chemotherapy (R-CVP or R-CHOP) with or without rituximab as maintenance. Autologous stem cell transplant (ASCT) may be an option for selected patients at this stage. Patients who were refractory to rituximab are recommended obinutuzumab-bendamustine (O-Benda) with obinutuzumab as maintenance. Rituximab in combination with lenalidomide (R^2) is an option for both, R-refractory patients and non-R-refractory patients in second-line.

As justification for not including rituximab monotherapy as an option for non-R-refractory patients in second-line, the company cited the opinion of clinical experts, elicited during the advisory board meeting conducted by Celgene in March 2019:[19] "*According to clinical experts, R mono is rarely used in the relapsed/refractory setting in UK clinical practice."* Clinical experts also advised that: "*R-Benda is primarily used in a first-line setting and clinicians are reluctant to re-challenge relapsed/refractory patients with bendamustine in subsequent lines of therapy."* Therefore, bendamustine monotherapy was not considered an option in the R-refractory population.

Figure 2.1: Treatment pathway as described by the company for patients with follicular lymphoma with proposed positioning of R²



Source: Section B.1.3 of the CS.[1]

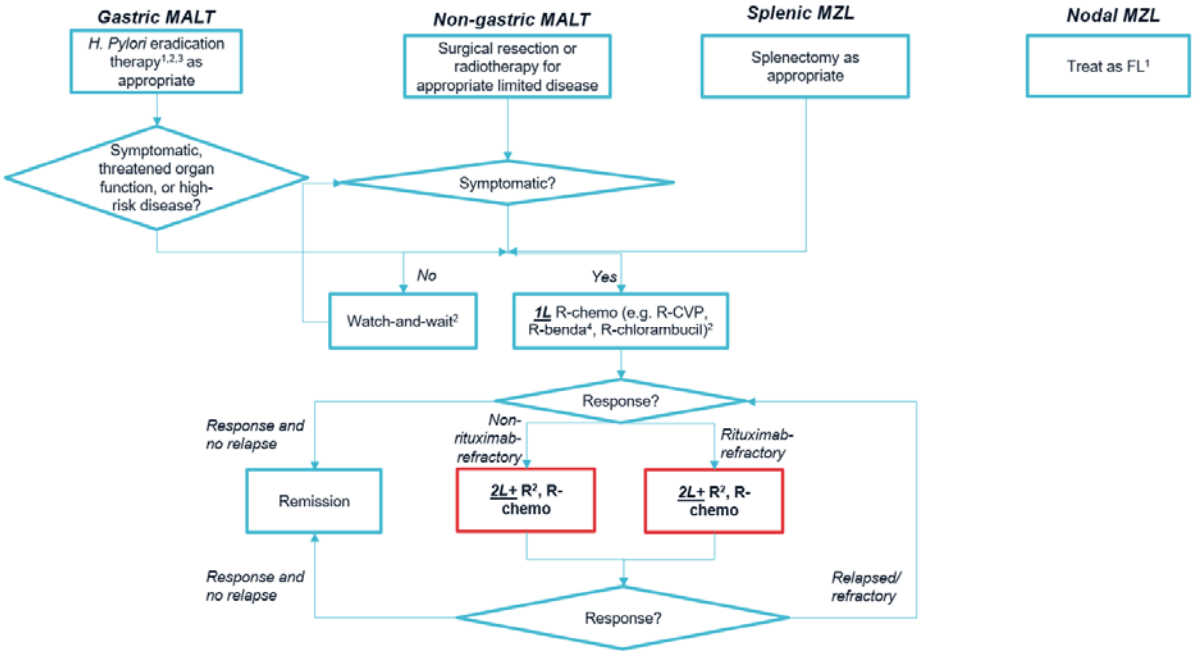
1L = first-line; 2L = second-line; Benda = bendamustine; CHOP = cyclophosphamide, doxorubicin, vincristine, prednisolone; CVP = cyclophosphamide, vincristine, prednisolone; O = obinutuzumab; R = rituximab; R-chemo+R = rituximab and chemotherapy induction followed by rituximab maintenance therapy; ASCT = autologous stem cell transplant.

* Please note that references in the graph are references from the CS.

The CS describes the following sources that were used in the company's interpretation of the positioning of R² in the treatment pathway for MZL (see Figure 2.2): an advisory board conducted by Celgene in March 2019, involving six UK clinical experts in NHL and two health economics experts,[19] NICE guideline for the diagnosis and management of non-Hodgkin's lymphoma (NG52),[25] the ESMO guidelines for marginal zone lymphoma, mantle cell lymphoma and peripheral T-cell lymphoma,[29] and the fact sheet for MZL by the Lymphoma Research Foundation.[30]

Figure 2.2 shows the treatment pathway proposed by the company for patients with marginal zone lymphoma (MZL). Treatment options are dependent on the type of MZL: gastric or non-gastric mucosa associated lymphoid tissue (MALT), splenic marginal zone lymphoma (SMZL) or nodal marginal zone lymphoma (NMZL). First-line treatment options include R-chemo (e.g. R-CVP, R-Benda or R-chlorambucil). Second-line treatment options are R² or R-chemo, both for R-refractory patients and for non-R-refractory patients. It is not clear to the ERG why R-chemo is a second-line treatment option for R-refractory patients with MZL, but not for R-refractory patients with FL.

Figure 2.2: Treatment pathway as described by the company for patients with marginal zone lymphoma with proposed positioning of R²



Source: Section B.1.3 of the CS.[1]

Benda = bendamustine; CHOP = cyclophosphamide, doxorubicin, vincristine, prednisolone; CVP = cyclophosphamide, vincristine, prednisolone; FL = follicular lymphoma; H = Helicobacter; MALT = mucosa-associated lymphoid tissue; MZL = marginal zone lymphoma; R = rituximab.

References: 1. Dreyling 2013;[29] 2. NICE, 2016;[25] 3. Lymphoma Research Foundation, 2018;[30] 4. Celgene, 2019.[19]

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with treated follicular lymphoma or marginal zone lymphoma	Adults with treated follicular lymphoma or marginal zone lymphoma	N/A	The population is in line with the scope. However, R^2 does not currently have a UK marketing authorisation, although CHMP opinion is anticipated on [redacted], and marketing authorisation is expected in [redacted].
Intervention	Lenalidomide with rituximab (R^2)	Lenalidomide with rituximab (R^2)	N/A	The intervention is in line with the scope.
Comparator(s)	• Rituximab monotherapy (R-mono) • Rituximab in combination with chemotherapy • Established clinical management without lenalidomide (including but not limited to bendamustine)	For non-rituximab refractory patients: • Rituximab in combination with chemotherapy • Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP) • Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP) For rituximab refractory patients: • Established clinical management without lenalidomide	For non-rituximab refractory patients: • R-mono is not considered a relevant comparator as clinical expert opinion confirmed it is rarely used in the relapsed/refractory setting in the UK.[19, 31] For rituximab refractory patients: • O-Benda is included as an option for rituximab-refractory patients under the category 'Established clinical management without lenalidomide'. This is the only	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
		o Obinutuzumab in combination with bendamustine (O-Benda)	NICE-recommended option for this patient group (via the CDF) and clinical experts stated this is the likely treatment choice for FL patients refractory to rituximab.[19] • Bendamustine monotherapy (Benda mono) is not considered a comparator in this population given that clinical experts believe O-Benda has largely replaced use of Benda mono in rituximab refractory patients.[19]	
Outcomes	The outcome measures to be considered include: • Overall survival • Progression-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life	The outcome measures to be considered include: • Overall survival • Progression-free survival • Event-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life • Time to next anti-lymphoma treatment • Time to next chemotherapy treatment • Response rate to next anti-lymphoma treatment	Several efficacy outcomes have been presented in addition to those in the scope as several secondary and exploratory outcomes were reported in the AUGMENT and MAGNIFY studies that provide additional insight into the efficacy of R^2	All outcomes are reported in AUGMENT. However, for the indirect comparisons only a limited number of outcomes have been included.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
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 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 PAS for the intervention or comparator technologies will be taken into account.	Adhering to the reference case, the cost effectiveness of treatments is expressed in terms of incremental cost per quality adjusted life year. Adhering to the reference case, a lifetime horizon is used. Adhering to the reference case the economic analyses has been conducted from an NHS and Personal Social Services perspective Adhering to the reference case, the PAS has been applied in all economic analysis for all Celgene products.	Confidential PAS schemes that apply to relevant subsequent comparator therapies are not included in these analyses as Celgene is not privy to such information	
Subgroups to be considered	None listed in scope	No specific subgroups	N/A	

Source: CS, Table 1, pages 7-9.

CDF = Cancer Drugs Fund; FL = follicular lymphoma; MZL = marginal zone lymphoma; NICE = National Institute for Health and Care Excellence.

3.1 Population

The population defined in the scope is: Adults with previously treated follicular lymphoma or marginal zone lymphoma.[32] The population in the CS is in line with the NICE scope.[1]

According to the company R[2] does not currently have a UK marketing authorisation, although CHMP opinion is anticipated on [redacted], and marketing authorisation is expected in [redacted]. Therefore, the relevant population for this appraisal is currently unclear.

The anticipated license is as follows: Lenalidomide in combination with rituximab (anti-CD20 antibody) is indicated for the treatment of adult patients with previously treated FL or MZL.[18] Treatment should not be initiated in patients with hypersensitivity to the active substance or to any of the excipients, in women who are pregnant, in women of childbearing potential unless all of the conditions of the Pregnancy Prevention Programme are met, and in children and adolescents from birth to less than 18 years.

3.2 Intervention

The intervention (lenalidomide in combination with rituximab) is in line with the scope.

Lenalidomide is administered orally and rituximab is administered by intravenous (IV) infusion. Lenalidomide capsules should be taken orally at about the same time on the scheduled days.[18] The recommended starting dose of lenalidomide is 20 mg, orally once daily on days 1 to 21 of repeated 28-day cycles for up to 12 cycles of treatment. The recommended starting dose of rituximab is 375 mg/m^2 IV every week in cycle 1 (days 1, 8, 15, and 22) and day 1 of every 28-day cycle for cycles 2 through 5.[18]

The following tests/investigations are recommended when administering lenalidomide in combination with rituximab:[18]

- Medically supervised pregnancy tests with a minimum sensitivity of 25 mIU/mL must be performed for women of childbearing potential, including those who practice abstinence, before treatment, every four weeks during treatment, and four weeks after the end of treatment (except in the case of confirmed tubal sterilisation)
- Patients with known risk factors for myocardial infarction (including prior thrombosis) should be closely monitored, and action should be taken to try to minimise all modifiable risk factors (e.g. smoking, hypertension, and hyperlipidaemia)
- A complete blood cell count should be performed at baseline and then weekly for the first three weeks of Cycle 1 (28 days), every two weeks during Cycles 2 through 4, and then at the start of each cycle thereafter
- Careful monitoring and evaluation for tumour flare reaction (TFR) is recommended.
- Careful monitoring and evaluation for tumour lysis syndrome (TLS) is recommended. Patients should be well hydrated and receive TLS prophylaxis, in addition to weekly chemistry panels during the first cycle or longer, as clinically indicated.

3.3 Comparators

The description of the comparators in the NICE scope is as follows: Rituximab monotherapy, rituximab in combination with chemotherapy, and established clinical management without lenalidomide (including but not limited to bendamustine).[32]

ERG comment: The NICE scope does not make a distinction in terms of patients being rituximab refractory or not. However, the CS has different comparators for rituximab refractory patients and non-rituximab refractory patients. The company was asked why they made this distinction (Clarification Letter, Question A7), because, according to the ERG, if the intervention includes rituximab, the comparator should also be able to include rituximab. The company stated that this was done to 'reflect the current approach to patient management in the UK' and that 'patients determined to be R-refractory are treated differently to the non-R-refractory population in the UK due to the availability of an alternative treatment, obinutuzumab-bendamustine, approved in the EU and recommended by NICE' (Response to Clarification Letter).[26]

For non-rituximab refractory patients, the company included two comparators, both different types of rituximab in combination with chemotherapy:

- Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP)
- Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP)

For rituximab refractory patients, the company included one comparator:

- Established clinical management without lenalidomide, i.e. obinutuzumab in combination with bendamustine (O-Benda).

ERG comment: The ERG has several concerns with these comparators. Firstly, the NICE scope does not make a distinction in terms of patients being rituximab refractory or not. Therefore, the CS should have included a comparison of R^2 with rituximab monotherapy for all patients as specified in the scope. Secondly, even if the NICE committee accepts splitting the population in rituximab refractory patients and non-rituximab refractory patients, the CS should still have included a comparison with rituximab monotherapy for both populations as specified in the scope. Thirdly, the company included O-Benda as a comparator for rituximab refractory patients. However, in the response from NICE to comments on the draft scope, NICE clearly stated that "obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4).[28] Therefore, we believe that the submission currently does not present any relevant evidence for R-refractory patients.

3.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall survival
- progression-free survival
- overall response rate
- adverse effects of treatment
- health-related quality of life.

These outcomes were all assessed in the AUGMENT trial. The company included several additional outcomes (event-free survival, time to next anti-lymphoma treatment, time to next chemotherapy treatment, and response rate to next anti-lymphoma treatment) based on the AUGMENT trial. Therefore, all these outcomes are available for the comparison R^2 versus rituximab monotherapy.

All other comparisons rely on indirect comparisons. The company was not able to find any evidence providing a common comparator linking R^2 with any of the comparators of interest (apart from rituximab monotherapy, which was dismissed by the company). Therefore, the company performed a matching-adjusted indirect comparison (MAIC) to compare R^2 with R-CHOP and R-CVP in the non-

R-refractory population. For these analyses, only the outcomes OS, PFS, overall response rate (ORR) and complete response (CR) rate were used.

3.5 Other relevant factors

According to the company, R^2 represents an innovation in the management of patients with previously treated FL and MZL, because it is the first chemotherapy-free combination immunotherapy regimen licensed in this setting by the US Food and Drug Administration. The regimen is currently pending approval in the EU (CS, Document A, Section A16, pages 36-37; and Document B, Section B.2.12, pages 98-99).[1, 33]

There is a confidential simple discount PAS for lenalidomide ([redacted]) which applies to all current and future indications.

End-of-life criteria are not applicable for this appraisal (see CS, page 105).[1]

According to the company, no equality considerations have been identified or are anticipated (see CS, Document A, Section A3, page 8; and Document B, Section B.1.4, page 23).[1, 33]

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

4.1.1 Searches

Appendix D.1.1 of the CS provided details of the systematic search of the literature used to identify clinical effectiveness literature. It was reported that searches were conducted on 1 September 2017 and then updated on 4 April 2019. The ERG clarification letter asked whether the update searches had been conducted from database inception. In response, the company stated that it had been incorrect to report that the searches had been updated from September 2017 to 4 April 2019; the searches conducted on 4 April 2019 replaced the 2017 searches. "This was a de novo Clinical SLR conducted to replace the older Clinical SLR (with a cut off of September 2017), as some changes were made to the protocol and search strategies were made more extensive. All searches were conducted from database inception."[26] A summary of the resources searched is provided in Table 4.1.

Table 4.1: Resources for the clinical effectiveness and adverse reactions literature searches

Search strategy element	Resource	Host/Source	Date Range	Date searched
Electronic databases	MEDLINE	ProQuest	Not reported	4 April 2019
	Embase		Not reported	
	CENTRAL	Cochrane Library	Not reported	
	CDSR		Not reported	4 April 2019
Conference proceedings	EHA	Organisation websites, abstract books	2015-18	April 2019
	ICML		2013, 2015, 2017	
	ASCO		2015-2018	
	ASH		2014-2018	
	ESMO		2014-2018	
HTA Agencies	NICE	Organisation websites		April 2019
	CADTH			
	TGA			
Trials registries	ClinicalTrials.gov	ClinicalTrials.gov		April 2019
Manual searching of references of published systematic reviews, meta-analyses, and HTA documents was also conducted to identify potential publications that may not have been identified from the electronic searches.				
CENTRAL = Cochrane Central Register of Controlled Trials; CDSR = Cochrane Database Systematic Reviews; EHA = European Hematology Association; ICML = International Conference on Malignant Lymphomas; ASCO = American Society of Clinical Oncology; ASH = American Society of Hematology; ESMO = European Society for Medical Oncology; NICE = National Institute for Health and Care Excellence; CADTH = Canadian Agency for Drugs and Technologies in Health; TGA = Therapeutic Goods Administration Australia.				

ERG comment:

- The selection of databases searched was adequate, and searches were clearly reported and reproducible. The database name, host and date searched were provided. The date range of the searches was not reported.
- Searches conducted in September 2017 were reported in the CS, along with 'update' searches conducted in April 2019. In response to the ERG clarification letter, the company explained that

the April 2019 searches had replaced the September 2017 searches. Reporting the April 2019 searches would have been sufficient.

- An extensive range of resources additional to database searches was included in the SLR to identify further relevant studies and grey literature. Details of the resources searched, search strategies or search terms used, dates of searches, and results were not reported in the CS, but full details of the conference proceedings and HTA organisation website searches were provided in response to the ERG clarification letter.
- Accurate details of the MEDLINE segments searched were not reported. It is not clear if MEDLINE In-Process, Ahead of Print, and Daily Update were searched.
- Truncation and proximity operators were inconsistently used throughout. There were few synonyms used in the 2017 searches, and although there were more included in the 2019 'update' searches, they were still lacking.
- Comparators of interest were not included in the 2017 searches, but were included in the 2019 searches: prednisolone and cyclophosphamide.
- As study design filters were not included, both efficacy and safety evidence could be identified.
- The Cochrane Library searches did not report the database issue searched.
- The CS reported that ClinicalTrials.gov was searched for trials, but limited to "studies with results". In response to the ERG clarification letter, the company supplied full details of the ClinicalTrials.gov searches conducted in April 2019, which searched for "all studies".
- The PRISMA flow diagram (Figure 1) provided in the CS suggested that the 'update' searches of April 2019 were conducted from database inception, and replaced the original September 2017 search results. This was confirmed in the company response to clarification.
- A good range of conference proceedings and HTA organisation websites were searched, and although full details of these searches were not provided in the CS, they were provided in response to the ERG clarification letter.

4.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for randomised controlled trials (RCTs) and non-RCTs is presented in Table 4.2.

Table 4.2: Eligibility criteria (PICOS scope)

	Inclusion criteria	Exclusion criteria
Population (P)[a]	• Adults (≥18+) with relapsed and/or refractory FL • Adults (≥18+) with relapsed and/or refractory MZL • Any stage of disease	• Patients <18 years of age • Patients that do not have R/R FL/MZL
Intervention (I)	*Systemic induction (i.e., chemo-therapy, immunotherapy, chemo-immunotherapy) therapies recommended by NCCN/ESMO and deemed relevant to current clinical practice:* • Rituximab + bendamustine [R-B] • Rituximab + lenalidomide [R^2] • Rituximab + cyclophosphamide + vincristine + prednisone [R-CVP] • Rituximab + cyclophosphamide + doxorubicin + vincristine + prednisone [R-CHOP]	Any treatments that are not listed under the inclusion criteria

	Inclusion criteria	Exclusion criteria
	• Rituximab + chlorambucil [R-Chl] • Obinutuzumab + bendamustine [O-B] • Obinutuzumab + lenalidomide • Rituximab alone • Bendamustine alone • Lenalidomide alone • Idelalisib • Ibrutinib • Copanlisib • Tazemetostat • Rituximab + mitoxantrone + Chlorambucil + prednisone (R-MCP) • Rituximab + cyclophosphamide + doxorubicin + etoposide+ prednisone + interferon alpha (R-CHVPI)	
Comparators (C)[b]	• Any of the interventions listed in inclusion criteria OR fludarabine containing regimen • Placebo	Any treatments that are not listed under the inclusion criteria
Outcomes	• Survival (overall, progression-free, disease-free) • Response (overall, complete, partial) • Duration of treatment (median) • Duration of response • Quality of life: EORTC-QLQ-C30, EQ-5D, FACT-g and FACT-lym • Time to next lymphoma treatment • Adverse events of interest	Outcomes not included under inclusion criteria
Study design (S)	• Randomised controlled trials (RCTs) • Non-randomised clinical trials • Observational cohort studies (retrospective or prospective) • Systematic reviews and meta-analyses (for identification of primary studies only) • Single arms studies • Cross-sectional studies, case-control studies • Comparative studies	• Case series/case reports • Studies of non-original data • Non-systematic reviews • Comment, editorial, letter • Theses and dissertations • Non-human studies • Pharmacokinetic, pharmacodynamic, and bioequivalence studies
Publication type	Sample size ≥20 participants meeting the target population[c]	Sample size <20 participants meeting the target population[b]
Language	English language	Non-English

Source: CS, Appendix D1, Table 7.[1]
FL = follicular lymphoma; MZL = marginal zone lymphoma.
Notes: a) ≥70% of a mixed population needs to have R/R FL/MZL, or results need to be reported as subgroup data for the patient population of interest; b) only applicable to comparative studies; c) sample size limitation

	Inclusion criteria	Exclusion criteria
applies only to non-randomised studies. RCTs will be included regardless of sample size.		

ERG comment: Generally, the inclusion criteria are in line with the NICE scope. There are two small issues, both relating to outcomes. First, looking at inclusion criteria as formulated in the CS, it seems that only four specific quality of life instruments (EORTC-QLQ-C30, EQ-5D, FACT-g and FACT-lym) were included. Therefore, a paper comparing R^2 with R-chemo reporting the results for the SF-36 would be excluded. Second, only studies that reported 'adverse events of interest' were included. However, it is not specified what 'adverse events of interest' are. According to the ERG, all quality of life instruments and all adverse events should be eligible for inclusion. Nine studies were excluded because they did not include any relevant outcomes (CS, Appendix D, Figure 1, page 17).[34] However, the company did not provide a list with references of these studies; therefore, the ERG are unable to check whether any of these studies might be relevant.

4.1.3 Critique of data extraction

Data extraction of the selected relevant studies for the clinical evidence was performed by two independent reviewers and any discrepancies between reviewers were resolved by consensus and/or in conjunction with a third reviewer. The CS explains that when multiple sources of the same data were reported all sources were reviewed and reconciled (CS, Appendix D, page 15).[34]

ERG comment: The process of data extraction appears well conducted. The extraction by two independent reviewers minimises the risk of error and bias.

4.1.4 Quality assessment

In section D.5 of Appendix D of the CS,[34] the company lists the signalling questions that supported the risk of bias assessment of the trials AUGMENT and MAGNIFY, as follows:

- Was randomisation carried out appropriately?
- Was the concealment of treatment allocation adequate?
- Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?
- Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
- Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
- Is there any evidence to suggest that the authors measured more outcomes than they reported?
- Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

In the final statement regarding the quality assessment of the AUGMENT trial, the CS reports that '*Subsequently, this double-blind randomization method ensured low levels of bias in the AUGMENT study'*.[1] With regard to the to the quality assessment of the MAGNIFY trial, the CS states that *'(…) therefore, a lack of blinding was not thought to have a considerable effect on the outcome of the study. Furthermore, the results of interest for this submission are taken from the initial treatment period only and are therefore not affected by the open-label design'*.[1]

ERG comment: It is recommended that two reviewers perform risk of bias/quality assessment independently of each other to reduce the potential for any errors. This is not described in the CS.

Regarding the quality assessment of the AUGMENT trial, the ERG agrees that this is a good quality double-blind randomised trial. Regarding the MAGNIFY trial, the company only used data from the induction phase of the trial, i.e. before randomisation. Therefore, this study should be assessed as a single arm study, not an RCT. As such, the single arm from the MAGNIFY study is at high risk of bias.

4.1.5 Evidence synthesis

The company did not perform a meta-analysis to pool the two R^2 studies, AUGMENT and MAGNIFY.

ERG comment: The ERG agrees that this is justified because the MAGNIFY study, as used in the CS, did not have a comparator arm; and because there are important differences between the populations in the two studies. In particular MAGNIFY included both R-refractory and non-refractory patients but AUGMENT was only non-refractory patients, and there were differences regarding age, previous rituximab, refractory to last regimen, line of therapy, disease stage and Eastern Cooperative Oncology Group (ECOG) performance status.

The company did perform indirect comparisons because, according to the company, 'No head-to-head data are available for R^2 versus any of the comparators of interest to this submission; only R-mono was compared with R^2 within the AUGMENT RCT' (CS, Section b.2.9, page 67).[1] The ERG disagrees with this statement because, according to the NICE scope,[32] rituximab monotherapy is a relevant comparator; therefore, there are relevant head-to-head data available.

The company performed two matching-adjusted indirect comparisons (MAIC), one for the rituximab refractory population and one for the non-rituximab refractory population.

For non-rituximab refractory patients, the company included two comparators, both different types of rituximab in combination with chemotherapy:

- Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP)
- Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP)

For rituximab refractory patients, the company included one comparator:

- Established clinical management without lenalidomide, i.e. obinutuzumab in combination with bendamustine (O-Benda).

ERG comment: The ERG has several concerns with these comparators. Firstly, the NICE scope does not make a distinction in terms of patients being rituximab refractory or not. Therefore, the CS should have included a comparison of R^2 with rituximab monotherapy for all patients as specified in the scope. Secondly, even if the NICE committee accepts splitting the population in rituximab refractory patients and non-rituximab refractory patients, the CS should still have included a comparison with rituximab monotherapy for both populations as specified in the scope. Thirdly, the company included O-Benda as a comparator for rituximab refractory patients. However, in the response from NICE to comments on the draft scope, NICE clearly stated that "obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4)[28]). Therefore, the ERG believes that the submission currently does not present any relevant evidence for R-refractory patients.

Methods and results of the indirect comparison for the non-rituximab refractory population, R^2 versus R-CHOP and R-CVP, are discussed in Section 4.4 of this report.

Methods and results of the indirect comparison for the rituximab refractory population, R^2 versus O-Benda, will be ignored as this is not a relevant comparator according to NICE.[28]

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Included studies

The company identified three randomised controlled trials (RCTs) of the intervention of interest (lenalidomide in combination with rituximab, R^2): the AUGMENT trial,[35] the MAGNIFY trial,[36] and the ALLIANCE trial;[37] and one non-RCT: Tuscano 2014.[38] In this ERG report, the focus will be on the AUGMENT trial,[35] because this provides a head-to-head comparison of the intervention of interest (lenalidomide in combination with rituximab, R^2) versus a relevant comparator according to the NICE scope (rituximab monotherapy).

The other three studies of the intervention of interest (R^2) will be ignored in this report for the following reasons:

- The ALLIANCE trial[37] is a randomised, multicentre, Phase II study of R^2 versus lenalidomide monotherapy in patients with previously treated FL and prior rituximab. Lenalidomide monotherapy is not a relevant comparator according to the NICE scope. Therefore, only one arm of this trial is relevant.
- The MAGNIFY trial[36] is an ongoing, randomised, open-label, multicentre, Phase IIIb study of R^2 induction therapy followed by either R^2 maintenance therapy or R-mono maintenance therapy in patients with FL Grade 1–3b, MZL, or mantle cell lymphoma. Only patients who had stable disease (SD), partial response (PR), complete response (CR) or complete response unconfirmed (CRu) at the end of 12 cycles of initial therapy were randomised 1:1 to receive R^2 maintenance therapy or rituximab maintenance therapy. In the CS, the company only used data from the induction phase (before randomisation). However, this is single arm data, and as there is relevant RCT data from the AUGMENT trial, these data will be ignored in this report.
- Tuscano 2014[38] is a single-arm Phase II study evaluating the safety and efficacy of lenalidomide in combination with rituximab in patients with relapsed/refractory, indolent non-Hodgkin lymphoma (NHL), including 30 patients (22 FL, three MZL and five other NHL).

4.2.2 Methodology of the AUGMENT trial

The AUGMENT trial[35] is a randomised, double-blind, multicentre, Phase III study of R^2 versus rituximab plus placebo (R-mono) in non-R-refractory patients with FL Grade 1–3a or MZL. The study was conducted across 96 sites in 17 countries. The number of sites and patients from the UK have not been reported in the CS; but according to the clinical study report (CSR), the trial did not include any patients from the UK.

To be eligible for inclusion in the study, patients had to be aged ≥18 years, with histologically confirmed MZL or Grade 1, 2, or 3a FL (Grade 3b FL patients were excluded). Patients were required to have been previously treated with at least one systemic chemotherapy, immunotherapy or R-chemo. Initially, rituximab-naïve patients were included in the study; however, a protocol change required patients to have received at least two previous doses of rituximab. This change was carried out to ensure a study population that aligned with a population commonly seen in clinical practice. Furthermore, patients had to have documented relapsed/refractory FL or MZL; however, R-refractory patients were excluded (full inclusion and exclusion criteria are presented in Appendix M1 of the CS,[34] a summary is presented below in Table 4.3).

During the treatment period, patients underwent efficacy and safety assessments for a maximum of 12 cycles. Patients received oral lenalidomide or placebo at a starting dose of 10 mg (if creatine clearance (CrCl) ≥30 mL/min and <60 mL/min) or 20 mg (if CrCl ≥60 mL/min) once daily on Days 1 to 21 in each 28-day cycle, combined with four-weekly infusions of rituximab intravenously (IV) at a dose of 375 mg/m^2, followed by four additional doses on Day 1 of Cycles 2, 3, 4, and 5. Patients were stratified by prior rituximab treatment (yes vs. no), time since last anti-lymphoma therapy (≤2 vs. >2 years), and histology (FL vs. MZL), and then randomised 1:1 to R^2 or R-mono for 12 cycles. Treatment was terminated upon relapse or progression of disease, withdrawal of consent, or unacceptable toxicity.

Primary efficacy analyses were conducted on the ITT population, defined as all randomised patients. The primary endpoint of the study was PFS, as assessed by the Independent Review Committee (IRC) using a modification of the 2007 International Working Group Response Criteria (IWGRC [i.e. without a positron emission tomography scan]). Efficacy was assessed further in the ITT population through a number of secondary endpoints, including overall response rate (ORR), complete response (CR) rate, time to next anti-lymphoma treatment (TTNLT), duration of response (DOR), durable complete response rate (DCRR; defined as the proportion of patients that stayed in complete response for at least one year) and duration of complete response (DOCR). Safety analyses were conducted on the safety population, defined as all patients who received at least one dose of study treatment.

Pre-defined subgroup efficacy analyses were performed to compare treatments within the stratification factors, and between demographic and baseline characteristics. Table 4.3 presents a summary of the methodology for the AUGMENT trial.

Table 4.3: Summary of AUGMENT methodology

Trial Name	**AUGMENT**
Location	96 sites across 17 countries across North America, Europe, China and Brazil
Trial design	A multinational, randomised, double-blind, Phase III study Patients were randomised in a 1:1 ratio through an IVRS Randomisation was stratified by previous rituximab treatment (yes, no), time since last anti-lymphoma therapy (≤2, >2 years) and disease histology (FL, MZL)
Eligibility criteria for participants	**Inclusion criteria:** • Aged ≥18 years • Histologically confirmed MZL or Grade 1, 2, or 3a FL (CD20+ by flow cytometry or histochemistry) as assessed by investigator or local pathologist • Had to have been previously treated with at least one prior systemic chemotherapy, immunotherapy or R-chemo and had to have received at least two previous doses of rituximab: • Systemic therapy did not include local involved field radiotherapy for limited stage disease or *Helicobacter pylori* eradication • Prior investigational therapies were allowed provided the patient had received at least one prior systemic therapy • Had to have documented relapsed, refractory, or progressive disease (PD) after treatment with systemic therapy, and not be R-refractory • Rituximab-refractoriness was defined as did not respond (at least a PR) to rituximab or R-chemo therapy and/or time to disease progression <6 months after last rituximab dose • Rituximab-sensitive MZL or FL was defined as responded (at least a PR) to rituximab or R-chemo regimen therapy and time to disease progression ≥6 months after last rituximab dose

Trial Name	AUGMENT
	• Must have needed treatment for relapsed, progressed, or refractory disease as assessed by the investigator • Performance status ≤2 on the ECOG scale **Exclusion criteria:** • Life expectancy <6 months • Prior use of lenalidomide • Presence or history of central nervous system (CNS) involvement by lymphoma • Patients who were at a risk for a thromboembolic event and were not willing to take venous thromboembolism (VTE) prophylaxis
Settings and locations where the data were collected	An independent external DMC assessed ongoing safety throughout the study. The DMC conducted the planned interim futility analysis when an estimated 96 events per IRC review were reported. Response-related efficacy assessments were based on central review, including central radiology and clinical review by the IRC. Images received from investigators' sites were sent to the IRC, as well as relevant clinical information for haemato-oncology review.
Trial drugs	Lenalidomide 10 mg or 20 mg oral capsules[a] once daily on Days 1 to 21 of every 28-day Cycle up to 12 cycles combined with rituximab 375 mg/m^2 IV every week in Cycle 1 and on Day 1 of every 28-day Cycle from Cycles 2 through 5. Treatment continued until progression or unacceptable toxicity.
Permitted and disallowed medication	The following medications are prohibited during the study: • Systemic chronic corticosteroid at doses above 20 mg/day (prednisone/prednisolone or equivalent) during treatment phase. A seven-day washout period before Cycle 1 Day 1 study drug dosing was required for these patients • All investigational therapies (drug or otherwise) and anticancer therapies, other than lenalidomide or rituximab were prohibited during the entire Treatment Period of the study
Primary outcomes (including scoring methods and timings of assessments)	• PFS in relapsed/refractory indolent lymphoma patients, defined as the time from randomization to the first observation of disease progression, based on the modified 2007 IWGRC, or death due to any cause • Analysis was based on the IRC determination of disease progression

Other outcomes used in the economic model/specified in the scope	**Secondary endpoints**
	• To compare the safety of R^2 versus rituximab plus placebo
	• To compare the efficacy of R^2 versus rituximab plus placebo using other parameters of efficacy:
	  • DCRR, ORR, CR rate, DOR, and DOCR by the 2007 IWGRC without PET
	  • OS, EFS, and TTNLT
	Exploratory endpoints
	• To compare the effects of R^2 versus R-mono on:
	  • TTNCT and RTNLT
	  • CR/CRu rate in patients with FL based on the 1999 IWGRC
	  • PFS on next anti-lymphoma treatment (PFS2)
	  • HRQL as measured by the EORTC Quality of Life Questionnaire, Core 30 (QLQ-C30) and EuroQol Group's questionnaire 5 dimensions (EQ-5D-3L)

Pre-planned subgroups	Efficacy analyses were performed within a number of patient subgroups. These are described in Appendix M of the CS.

Trial Name	AUGMENT
Source: CS Table 4, pages 30-32. CR = complete response; CT = computerised tomography; DCRR = durable complete response rate; DMC = data monitoring committee; DOCR = duration of complete response; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; EFS = event-free survival; EORTC = European Organisation for Research and Treatment of Cancer; FL = follicular lymphoma; HRQL = health-related quality of life; IRC = Independent Review Committee; IVRS = interactive voice response system; IWGRC = International Working Group Response Criteria; MALT = mucosa-associated lymphoid tissue; MRI = magnetic resonance imaging; MZL = marginal zone lymphoma; ORR = overall response rate; PD = progressive disease; PR = partial response; R^2 = rituximab plus lenalidomide; R-chemo = rituximab-containing chemotherapy; R-mono = rituximab monotherapy; RTNLT = response rate to next anti-lymphoma treatment; TTNLT = time to next anti-lymphoma treatment; TTNCT = time to next chemotherapy treatment. Notes: [a] dose modification rules allowed for dosing down to 2.5 mg with Celgene supplying lenalidomide 2.5 mg, 5 mg, 10 mg, 15 mg, and 20 mg capsules.	

4.2.3 Baseline characteristics of the AUGMENT trial

Baseline characteristics for patients in the AUGMENT trial are presented in Table 4.4.

ERG comment: Of note, more patients in the R^2 arm than in the R-mono arm were female (58% vs. 46%), aged ≥65 years (46% vs. 41%) had Ann Arbor Stage III to IV disease (77% vs. 69%), FLIPI score ≥3 (39% vs. 30%), had an ECOG score of 1 or 2 (35% vs. 29%) and were refractory to the last prior regimen (17% vs. 14%). In addition, the company stated that for patients with MZL, baseline disease characteristics were imbalanced and favoured the R-mono arm (R^2 arm vs. R-mono arm): ECOG 0 (55% vs 72%); Ann Arbor Stage III to IV disease (77% vs. 56%); Ann Arbor Stage IV (65% vs. 41%); FLIPI score ≥3 (48% vs. 25%); B symptoms (13% vs. 3%); and high tumour burden per GELF criteria (65% vs. 56%). The ERG agrees with this and judged that the baseline characteristics for MZL patients may favour R-mono.

Table 4.4: Baseline demographic and disease characteristics, AUGMENT – ITT population

	FL		MZL		Total		Overall (n=358)
	R^2 (n=147)	R-mono (n=148)	R^2 (n=31)	R-mono (n=32)	R^2 (n=178)	R-mono (n=180)	
Male, n (%)	61 (41.5)	80 (54.1)	14 (45.2)	17 (53.1)	75 (42.1)	97 (53.9)	172 (48.0)
Median age, years (range)	62.0 (26.0-86.0)	61.0 (35.0-88.0)	68.0 (37.0-80.0)	66.0 (36.0-82.0)	64.0 (26.0-86.0)	62.0 (35.0-88.0)	62.5 (26.0-88.0)
Age distribution, n (%)							
<65	86 (58.5)	94 (63.5)	10 (32.3)	13 (40.6)	96 (53.9)	107 (59.4)	203 (56.7)
≥65	61 (41.5)	54 (36.5)	21 (67.7)	19 (59.4)	82 (46.1)	73 (40.6)	155 (43.3)
≥70	34 (23.1)	32 (21.6)	13 (41.9)	12 (37.5)	47 (26.4)	44 (24.4)	91 (25.4)
Race, white (%)	[redacted]	[redacted]	[redacted]	[redacted]	118 (66.3)	115 (63.9)	233 (65.1)
Histology (investigator review), n (%)							
FL	[redacted]	[redacted]	[redacted]	[redacted]	147 (82.6)	148 (82.2)	295 (82.4)
Grade 1	[redacted]	[redacted]	[redacted]	[redacted]	50 (28.1)	62 (34.4)	112 (31.3)
Grade 2	[redacted]	[redacted]	[redacted]	[redacted]	75 (42.1)	61 (33.9)	136 (38.0)
Grade 3a	[redacted]	[redacted]	[redacted]	[redacted]	22 (12.4)	25 (13.9)	47 (13.1)
MZL	N/A	N/A	31 (100.0)	32 (100.0)	31 (17.4)	32 (17.8)	63 (17.6)
MALT	N/A	N/A	14 (45.2)	16 (50.0)	14 (7.9)	16 (8.9)	30 (8.4)
Nodal	N/A	N/A	8 (25.8)	10 (31.3)	8 (4.5)	10 (5.6)	18 (5.0)
Splenic	N/A	N/A	9 (29.0)	6 (18.8)	9 (5.1)	6 (3.3)	15 (4.2)
Ann Arbor stage, n (%)							
I	13 (8.8)	13 (8.8)	2 (6.5)	5 (15.6)	15 (8.4)	18 (10.0)	33 (9.2)
II	21 (14.3)	29 (19.6)	5 (16.1)	9 (28.1)	26 (14.6)	38 (21.1)	64 (17.9)
III	69 (46.9)	60 (40.5)	4 (12.9)	5 (15.6)	73 (41.0)	65 (36.1)	138 (38.5)
IV	44 (29.9)	46 (31.1)	20 (64.5)	13 (40.6)	64 (36.0)	59 (32.8)	123 (34.4)

	FL		MZL		Total		Overall (n=358)
	R² (n=147)	R-mono (n=148)	R² (n=31)	R-mono (n=32)	R² (n=178)	R-mono (n=180)	
FLIPI category (derived), n (%)							
Low (0,1)	[redacted]	[redacted]	[redacted]	[redacted]	52 (29.2)	67 (37.2)	119 (33.2)
Intermediate (2)	[redacted]	[redacted]	[redacted]	[redacted]	55 (30.9)	58 (32.2)	113 (31.6)
High (≥3)	[redacted]	[redacted]	[redacted]	[redacted]	69 (38.8)	54 (30.0)	123 (34.4)
Baseline ECOG score, n (%)							
0	99 (67.3)	105 (70.9)	17 (54.8)	23 (71.9)	116 (65.2)	128 (71.1)	244 (68.2)
1	47 (32.0)	42 (28.4)	13 (41.9)	8 (25.0)	60 (33.7)	50 (27.8)	110 (30.7)
2	[redacted]	[redacted]	[redacted]	[redacted]	2 (1.1)	2 (1.1)	4 (1.1)
LDH elevated, n (%)							
Yes	39 (26.5)	43 (29.1)	6 (19.4)	6 (18.8)	45 (25.3)	49 (27.2)	94 (26.3)
No	107 (72.8)	105 (70.9)	25 (80.6)	26 (81.3)	132 (74.2)	131 (72.8)	263 (73.5)
High tumour burden (GELF criteria)							
Yes	77 (52.4)	68 (45.9)	20 (64.5)	18 (56.3)	97 (54.5)	86 (47.8)	183 (51.1)
No	70 (47.6)	80 (54.1)	11 (35.5)	14 (43.8)	81 (45.5)	94 (52.2)	175 (48.9)
Prior anti-lymphoma regimens							
1	[redacted]	[redacted]	[redacted]	[redacted]	102 (57.3)	97 (53.9)	199 (55.6)
>1	[redacted]	[redacted]	[redacted]	[redacted]	76 (42.7)	83 (46.1)	159 (44.4)
Refractory to last prior regimen							
Yes	26 (17.7)	25 (16.9)	4 (12.9)	1 (3.1)	30 (16.9)	26 (14.4)	56 (15.6)
No	121 (82.3)	123 (83.1)	27 (87.1)	31 (96.9)	148 (83.1)	154 (85.6)	302 (84.4)
POD24[a], n (%)							
Yes	[redacted]	[redacted]	[redacted]	[redacted]	56 (31.5)	61 (33.9)	117 (32.7)
No	[redacted]	[redacted]	[redacted]	[redacted]	122 (68.5)	118 (65.6)	240 (67.0)

	FL		MZL		Total		Overall (n=358)
	R^2 (n=147)	R-mono (n=148)	R^2 (n=31)	R-mono (n=32)	R^2 (n=178)	R-mono (n=180)	

Source: CS, Table 5, pages 34-35.

ECOG = Eastern Cooperative Oncology Group; FL = follicular lymphoma; FLIPI = follicular lymphoma international prognostic index; GELF = Groupe d'Etude des Lymphomes Folliculaires; LDH = lactate dehydrogenase; MZL = marginal zone lymphoma; MALT = mucosa associated lymphatic tissue; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo.

Notes: [a]) POD24 is defined as relapse within two years of initial chemoimmunotherapy.

4.2.4 Statistical analyses of the AUGMENT trial

The primary outcome of AUGMENT was PFS. The primary analysis was performed in the ITT population using outcomes assessed by the IRC using a modified version of the 2007 IWGRC. Analyses were performed using both FDA and European Medicines Agency (EMA) censoring rules for PFS but only the EMA censoring rule analyses for the ITT population were presented in the main body of the CS. Safety assessments for the study were conducted on the safety population.

Table 4.5 presents the hypothesis and associated statistical analysis methods adopted in the AUGMENT trial. PFS was defined as the time from the date of randomisation to the first observation of documents disease progression or death from any cause, whichever occurred first. The analysis compared Kaplan-Meier survival curves using a log-rank test (one sided $p < 0.025$) and a Cox proportional hazards model. OS was also analysed using Kaplan-Meier estimates of OS.

Overall response rate (ORR) was defined as the proportion of patients with best response of at least PR without administration of new anti-lymphoma therapy. Complete response (CR) was the proportion of patients with a best response of CR during the study without administration of new anti-lymphoma therapy. ORR and CR were compared between treatment groups using a stratified Cochran-Mantel-Haenszel (CMH) test stratified by the randomisation stratification factors.

Planned subgroup analyses included the randomisation stratification factors previous rituximab treatment (yes, no), time since last anti-lymphoma therapy (≤ 2, >2 years), and histology (FL, MZL) and also age (<65, ≥65 years); gender (male, female); race (White; Other races); region (US, EU, Asia-Pacific region and Brazil); FLIPI (<3, ≥3) for FL patients only; number of prior anti-lymphoma regimens (1, >1); Ann Arbor stage at enrolment (I to II, III to IV); prior rituximab-containing chemotherapy regimen (yes, no); refractory to last prior regimen (defined as <PR or PD within six months from last systemic regimen) (yes, no); High tumour burden per Groupe d'Etude des Lymphomes Folliculaires (GELF) criteria (yes, no); chemo-resistant (<PR or PD within six months from last chemotherapy) (yes, no) or ECOG performance status ≥2 [yes; no])

ERG comment: The statistical analysis of the trial used appropriate methods and the ERG does not have any concerns.

Table 4.5: Summary of statistical analyses

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
AUGMENT	The primary objective of the study was to compare the efficacy of R^2 to R-mono. Efficacy determination was based on PFS as the primary endpoint. The AUGMENT study was considered positive if the R^2 group was significantly superior to the rituximab group for the primary endpoint.	The analysis of the primary endpoint was planned when approximately 193 IRC-assessed PFS events were reached. The cut-off date for database lock was prespecified before database lock. KM estimates of PFS were provided, and the KM product limit method was used to estimate the survivorship function for PFS. Event rates at specific time points were estimated from KM curves. Medians together with two-sided 95% CIs were provided. The resulting PFS estimates were presented graphically.	Based on the rate of accrual anticipated in this study and 5% annual dropout rate, it was estimated that approximately 350 patients would be randomised in a 1:1 ratio to the two treatment arms and that PFS would be reached at 43 months. The basis for the power and sample size determination was a test of the equality of the overall time-to-event (i.e. PFS) curves between experimental and control treatment groups using a stratified log-rank test.	**EMA censoring rules** Event: • Death before first PD assessment while on study • Death between adequate assessment visits • All progressions and deaths, regardless of whether they occurred after next anti-lymphoma therapy or after ≥2 missed scheduled assessments Censored: • Patients with no baseline assessment were censored at randomisation • Patients who did not progress or die and those that discontinued for any reason other than death or progression will be censored on the date of their last adequate assessment with evidence of no progression • Patients who died or progressed after more than one missed visit will be censored at the date of their last adequate assessment that revealed no progression

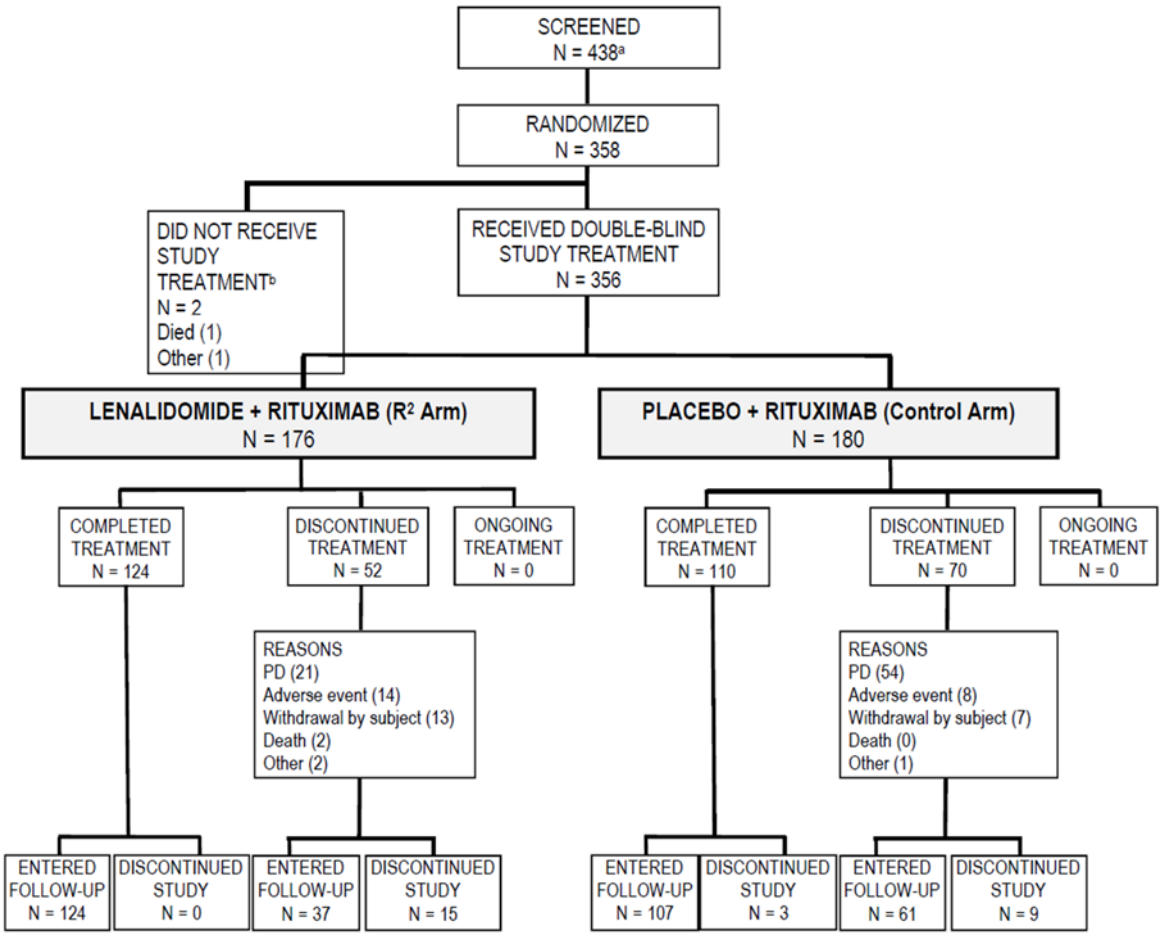
Source: CS, Table 8, pages 46-48.
CI = confidence interval; EMA = European Medicines Agency; IEE = induction efficacy population; IRC = Independent Review Committee; ITT = intention-to-treat; KM = Kaplan–Meier; ORR = overall response rate; PD = progressive disease; PFS = progression-free survival; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo.

4.2.5 Results of the AUGMENT trial

The data presented in the CS are based on the 22 June 2018 data cut-off for the primary analysis. Efficacy analyses were conducted in the ITT population and based on data from IRC review, using the modified 2007 IWGRC. EMA censoring rules were applied to the analyses.

At the time of the data cut-off (22 June 2018) more patients in the R^2 arm had completed treatment compared with the R-mono arm. In the R^2 arm, 124 patients (70.5%) had completed treatment, 52 patients (29.5%) had discontinued treatment, and no patients were ongoing with treatment. In the R-mono arm, 110 patients (61.1%) had completed treatment, 70 patients (38.9%) had discontinued treatment, and no patients were ongoing with study treatment (see Figure 4.1).

Figure 4.1: CONSORT diagram of patient flow during the AUGMENT trial



Source: CS, Appendix D4, pages 59-60.

FL = follicular lymphoma; MZL = marginal zone lymphoma; PD = progressive disease.

Notes: [a]) in total, 438 patients were screened for study participation, of which 18 patients (4.1%) were screened twice. Of the total 456 screens, 98 were screen failures primarily due to failure of inclusion/exclusion criteria (96.9%). Screen failures either did not meet inclusion criteria (n=70) and/or met at least one exclusion criterion (n=28); [b]) two patients randomised to the R^2 arm did not receive study medication: one patient with MZL died due to septic shock after randomisation but prior to receiving the first dose of study treatment and one patient with FL discontinued due to Grade 2 dyspnoea on Cycle 1 Day 1, prior to administration of the first dose of study drug.

The overall median follow-up time for surviving patients in the ITT Population was 28.30 months (range: 0.1 to 51.3 months); this was comparable between FL and MZL patients.

Table 4.6 presents a summary of the main results from the AUGMENT trial. Results for FL and MZL separately are reported in Appendix 1 of this report.

Table 4.6: Summary of results from the AUGMENT trial: ITT population

Endpoint	**Overall**	
	R^2 (n=178)	**R-mono (n=180)**
Median OS, months (95% CI)[a]	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.61 (0.33, 1.13)[b]	
Median PFS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted][b]	
Best response, n (%)		
ORR (CR+PR)	138 (77.5)	96 (53.3)
95% CI[d]	70.7, 83.4	45.8, 60.8
p-value	<0.0001[e]	
CR rate	60 (33.7)	33 (18.3)
95% CI[d]	26.8, 41.2	13.0, 24.8
p-value	0.001[e]	
PR	78 (43.8)	63 (35.0)
SD	20 (11.2)	55 (30.6)
PD/ death	7 (3.9)	23 (12.8)
No evidence of disease	3 (1.7)	4 (2.2)
Unknown/ND/Missing	10 (5.6)	2 (1.1)
Median TTNLT, months (95% CI)[a]	NE (NE, NE)	32.2 (23.2, NE)
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	0.54 (0.38, 0.78)[b]	
p-value	0.0007[g]	
Median EFS, months (95% CI)[a]	27.6 (22.1, NE)	13.9 (11.4, 16.7)
Hazard ratio (95% CI)	0.51 (0.38 to 0.67)[b]	
p-value	<0.0001[g]	
Median TTNCT, months (95% CI)[a]	NE (NE, NE)	NE (NE, NE)
TTNCT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	0.50 (0.32, 0.78)[b]	
p-value	0.0017[g]	
RTNLT		
ORR, n (% [95% CI][d])	28 (57.1 [42.2, 71.2])	29 (36.3 [25.8, 47.8])
p-value	0.0282[f]	
CR, n (% [95% CI][d])	15 (30.6 [18.3, 45.4])	13 (16.3 [8.9, 26.2])
p-value	0.0775[f]	
DCRR, n/N (%)	[redacted]	[redacted]

95% CI[d]	[redacted]	[redacted]
p-value	[redacted][e]	
N, Median DOR, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)[c]	0.53 (0.36 to 0.79)	
p-value[e]	0.0015	
N, Median DOCR, months (95% CI)[a]	60, NE (25.3, NE)	33, NE (13.8, NE)
Hazard ratio (95% CI)[h]	[redacted]	
p-value	[redacted]	

Source: CS, Table 10 and 11, pages 52 and 55-56.
CI = confidence interval; CR = complete response; DCRR = durable complete response rate, DOCR = duration of complete response; DOR = duration of response; EFS = event-free survival; FL = follicular lymphoma; IRC = Independent Review Committee; ITT = intent-to-treat; MZL = marginal zone lymphoma; ND = not done; NE = not estimable; ORR = overall response rate; OS = overall survival; PD = progressive disease; PR = partial response; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo; RTNLT = response rate to next anti-lymphoma treatment; SD = stable disease; TTNLT = time to next anti-lymphoma treatment; TTNCT = time to next anti-lymphoma chemotherapy treatment.
Notes: [a]) median estimate is from Kaplan–Meier analysis; [b]) from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last anti-lymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c]) from Cox proportional hazard model; [d]) exact confidence interval for binomial distribution; [e]) from CMH test adjusting for the three stratification factors; [f]) from Fisher-Exact test; [g]) from log-rank test adjusting for the three stratification factors; [h]) from log-rank test; [i]) exact confidence interval for binomial distribution.

Overall, R^2 showed favourable results when compared to R-mono for PFS with a greater median PFS ([redacted] months; HR [redacted] (95% CI: [redacted]). However, there was no evidence of a difference in overall survival (OS) with a hazard ratio (HR) of 0.61 (95% CI: 0.33 to 1.13) for patients treated with R^2 compared to R-mono. At the time of the analysis the OS data was immature with 16 deaths on R^2 and 26 deaths on R-mono at the time of the analysis. Overall response rate (ORR) was significantly greater for R^2 compared with R-mono (78% vs. 53%; p<0.0001). The complete response (CR) rate was also greater for the R^2 arm compared with R-mono (34% vs. 18%; p=0.001).

ERG comment: As can be seen from Tables A1.1 and A1.2 (see Appendix 1 of this report), results for R^2 versus R-mono in MZL patients are generally less favourable for R^2 than in FL patients. However, it is important to note that PFS outcomes in the MZL subgroup are difficult to interpret because of the small sample size (31 patients in the R^2 arm and 32 patients in the R-mono arm) and imbalance in baseline prognostic factors (as discussed in Section B.2.3 of the CS and section 4.2.3 of this report).

Health-related quality of life (HRQoL)

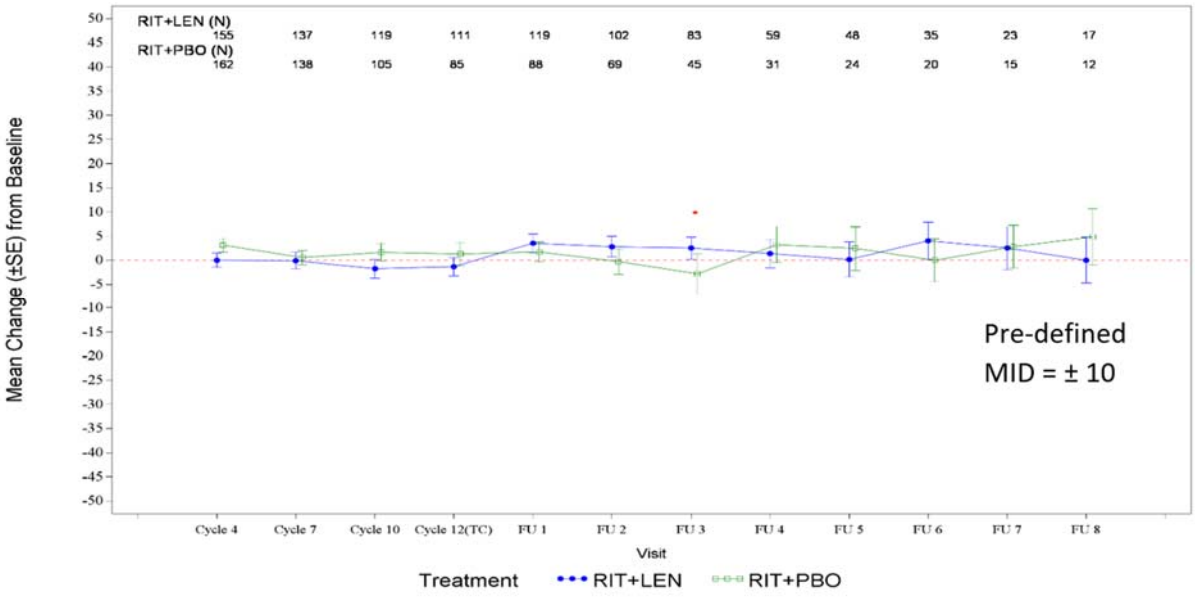
HRQoL was assessed using the EORTC QLQ-C30 and EuroQol Five Dimension Three Level (EQ-5D-3L) questionnaire. The global health status/quality of life (GHS/QoL) domain of the QLQ-C30 was chosen as the primary patient reported outcome of interest.

Primary HRQoL analyses were performed on the HRQoL-evaluable population, defined as patients in the ITT population who had a GHS/QoL domain score at baseline and at least one post-baseline assessment. The ITT population was also analysed, but only to assess the HRQoL compliance rates. The HRQoL-evaluable population comprised of 338 patients (94% of the ITT population), including 165 patients receiving R^2 and 173 patients receiving R-mono.

A minimal important difference (MID) of a ≥10-point change from baseline at the individual patient level was used to define the proportion of patients reporting a meaningful difference in QOL for any given domain of the EORTC QLQ-C30.

Based on the results from the cross-sectional analysis (within- and between-group difference in mean change from baseline score at each post-baseline assessment visit), no clinically meaningful change from baseline in the GHS/QoL domain of the QLQ-C30 was observed across any of the post-baseline assessment visits, regardless of treatment group (See Figure 4.2). Between-group differences in mean changes were small and not clinically meaningful across all assessment visits and did not differ between FL and MZL patients. Furthermore, change from baseline scores over time, based on the cross-sectional assessment, did not differ meaningfully by response status, occurrence of Grade 3/4 AEs, and occurrence of any neutropenia. The longitudinal assessment also indicated no statistically significant or clinically meaningful difference in LS mean changes from baseline between treatment groups across all timepoints; and no change exceeded the MID threshold.

Figure 4.2: Cross-sectional assessment of global health status/quality of life changes from baseline



Source: CS, Appendix P, Figure 22, page 237.
FU = follow-up; MID = minimally important difference; Len = lenalidomide; PBO = placebo; Rit = rituximab; SE = standard error; TC = treatment completion.

4.2.6 Adverse events

Adverse event data from the AUGMENT trial were taken from the 22 June 2018 database cut-off; safety analyses were conducted in the safety population.

Overall, the median lenalidomide/placebo treatment duration was [redacted] months for the R^2 arm and [redacted] months for the R-mono arm. The median rituximab treatment duration was also similar between the R^2 and R-mono arms ([redacted] vs. [redacted], respectively).

A summary of the treatment-emergent adverse event (TEAEs) during AUGMENT for the total population (FL and MZL) is presented in Table 4.7. TEAEs were reported in 174 patients (99%) in the R^2 arm and 173 patients (96%) in the R-mono arm. More patients in the R^2 arm (69%) experienced a Grade 3 or 4 TEAE compared with those in the R-mono arm (32%), and two patients in each treatment arm reported a Grade 5 TEAE. Additionally, a greater proportion of patients reported serious adverse events in the R^2 arm (26%) compared with those in the R-mono arm (14%). Separate tables for FL and MZL patients are presented in Appendix 1 of this report.

Table 4.7: Summary of treatment-emergent adverse events in AUGMENT: Safety population

	Total population (FL + MZL)	
	R^2 (n=176)	**R-mono (n=180)**
Number of patients (%)		
Any TEAE	174 (98.9)	173 (96.1)
Len related	159 (90.3)	118 (65.6)
R related	132 (75.0)	105 (58.3)
Grade 3–4 TEAE	121 (68.8)	58 (32.2)
Len related	101 (57.4)	38 (21.1)
R related	57 (32.4)	19 (10.6)
Grade 5 TEAE	2 (1.1)	2 (1.1)
Any SAE	45 (25.6)	25 (13.9)
Len related	23 (13.1)	8 (4.4)
R related	13 (7.4)	3 (1.7)
Any TEAE leading to dose reduction of Len/Pbo	46 (26.1)	6 (3.3)
Any TEAE leading to dose interruption of Len/Pbo	112 (63.6)	47 (26.1)
Any TEAE leading to dose interruption of R	60 (34.1)	37 (20.6)
Any TEAE leading to discontinuation of Len/Pbo	15 (8.5)	9 (5.0)
Any TEAE leading to discontinuation of R	6 (3.4)	2 (1.1)

Source: CS, Table 21, page 94 and Clarification Letter, Table 6, page 21.
Len = lenalidomide; Pbo = placebo; R = rituximab; R2 = lenalidomide + rituximab; R mono = placebo, rituximab + placebo; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

In the safety population, TEAEs that occurred more frequently (≥10% difference) in the R^2 arm than the R-mono arm included the following: neutropenia (58% vs. 22%), diarrhoea (31% vs. 23%), constipation (26% vs. 14%), cough (23% vs. 17%), upper respiratory tract infection (18% vs. 13%) and leukopenia (20% vs. 9%) (see Table 4.8).

The difference in the number of Grade 3 or 4 TEAEs between treatment arms (shown in Table 4.7) was largely driven by Grade 3 or 4 events of neutropenia and leukopenia. Neutropenia occurred in 88 patients (50%) in the R^2 arm compared with 23 patients (13%) in the R-mono arm, and leukopenia occurred in 12 patients (7%) in the R^2 arm compared with three patients (2%) in the R-mono arm.

The most common TEAEs, occurring in more than 10% of patients, are presented in Table 4.8 below. Separate adverse events tables for FL and MZL patients are presented in Appendix 1 of this report.

Table 4.8: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT: Safety population (FL and MZL)

	Total population (FL + MZL)	
	R² (n=176)	**R-mono (n=180)**
Number of patients (%)		
Blood and lymphatic system disorders	118 (67.0)	58 (32.2)
Neutropenia	102 (58.0)	40 (22.2)
Leukopenia	36 (20.5)	17 (9.4)
Anaemia	28 (15.9)	8 (4.4)
Thrombocytopenia	26 (14.8)	8 (4.4)
Gastrointestinal disorders	115 (65.3)	88 (48.9)
Diarrhoea	55 (31.3)	41 (22.8)
Constipation	46 (26.1)	25 (13.9)
Abdominal pain	22 (12.5)	16 (8.9)
Nausea	20 (11.4)	23 (12.8)
Infections and infestations	110 (62.5)	88 (48.9)
URTI	32 (18.2)	23 (12.8)
Nasopharyngitis	13 (7.4)	18 (10.0)
General disorders and administration site conditions	98 (55.7)	89 (49.4)
Fatigue	38 (21.6)	33 (18.3)
Pyrexia	37 (21.0)	27 (15.0)
Asthenia	24 (13.6)	19 (10.6)
Oedema peripheral	23 (13.1)	16 (8.9)
Skin and subcutaneous tissue disorders	89 (50.6)	43 (23.9)
Pruritus	21 (11.9)	7 (3.9)
Rash	19 (10.8)	7 (3.9)
Musculoskeletal and connective tissue disorders	73 (41.5)	58 (32.2)
Muscle spasms	23 (13.1)	9 (5.0)
Back pain	14 (8.0)	18 (10.0)
Respiratory, thoracic and mediastinal disorders	73 (41.5)	65 (36.1)
Cough	40 (22.7)	31 (17.2)
Dyspnoea	19 (10.8)	8 (4.4)
Investigations	60 (34.1)	50 (27.8)
Alanine aminotransferase increased	18 (10.2)	15 (8.3)
Metabolism and nutrition disorders	58 (33.0)	40 (22.2)
Decreased appetite	23 (13.1)	11 (6.1)
Nervous system disorders	58 (33.0)	39 (21.7)
Headache	26 (14.8)	17 (9.4)
Injury, poisoning and procedural complications	42 (23.9)	40 (22.2)
Infusion related reaction	26 (14.8)	24 (13.3)

	Total population (FL + MZL)	
	R^2 (n=176)	**R-mono (n=180)**
Eye disorders	28 (15.9)	14 (7.8)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	26 (14.8)	9 (5.0)
Tumour flare	19 (10.8)	1 (0.6)
Psychiatric disorders	24 (13.6)	20 (11.1)
Cardiac disorders	21 (11.9)	17 (9.4)
Vascular disorders	21 (11.9)	22 (12.2)
Source: CS, Appendix F, Table 31, pages 63-64 and Clarification Letter, Table 7, pages 22-23. R^2 = lenalidomide + rituximab; R-placebo = rituximab + placebo; URTI = upper respiratory tract infection.		

ERG comment: As shown in Table 4.7, R^2 was associated with more grade 3-4 TEAEs and SAEs when compared to R-mono, especially lenalidomide/placebo related adverse events; but rituximab-related grade 3-4 TEAEs and SAEs were also more frequent in the R^2 arm than in the R-mono arm. R^2 was also associated with more TEAEs leading to dose reductions, dose interruptions and discontinuations of lenalidomide/placebo or rituximab when compared to R-mono. Adverse events are generally the same for FL and MZL patients; however, AEs for MZL patients are based on small numbers.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The systematic literature review (SLR) performed by the company identified 45 studies (13 RCTs and 32 non-RCTs). According to the company, 39 studies were considered not relevant for the submission because they did not investigate comparators of interest (lenalidomide (1x), obinutuzumab plus lenalidomide (1x), idelalisib (4x), copanlisib (2x), ibrutinib (3x), rituximab plus bendamustine (6x), other bendamustine-containing regimens (1x), rituximab monotherapy (15x), bendamustine monotherapy (5x), and tazemetostat (1x)). Therefore, the company included a total of six relevant studies.

ERG comment: As explained in Section 3.3 of this report, rituximab monotherapy is a relevant comparator for this appraisal according to the NICE scope. Therefore, the 15 studies investigating rituximab monotherapy should have been included. However, as there is a trial with a head-to-head comparison of R^2 with rituximab monotherapy, the 15 rituximab monotherapy studies can probably be ignored.

Of the six relevant studies identified by the company, there were five relevant RCTs (AUGMENT (R^2),[35] MAGNIFY (R^2),[36] ALLIANCE (R^2),[37] Van Oers (R-CHOP)[39] and GADOLIN (O-Benda)[40]) and one relevant non-RCT (Tuscano 2014[38] (R^2)). The SLR found no studies for the relevant comparator R-CVP.

ERG comment: The four R^2 studies were discussed in Section 4.1.2 of this report. This ERG report will focus on the AUGMENT trial, because this provides a head-to-head comparison of the intervention of interest (lenalidomide in combination with rituximab, R^2) versus a relevant comparator according to the NICE scope (rituximab monotherapy). The study by Van Oers et al. (2006), was relevant for the indirect comparison using published data and will be discussed in Section 4.4.1 of this report. The GADOLIN study was used by the company for an indirect comparison of R^2 with O-Benda. However, as explained in Sections 3.3 and 4.4 of this report, O-Benda is not considered by NICE to be a relevant comparator for this appraisal; therefore, this study will be ignored.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

The company performed two types of indirect comparisons. First, the company performed an indirect comparison with data from published evidence. This included comparisons of R^2 with:

- R-CHOP for non-rituximab refractory patients, based on comparator data from a study by Van Oers et al. (2006)[39] comparing R-CHOP with CHOP (only the R-CHOP arm was used in the analyses).
- Established clinical management without lenalidomide, i.e. O-Benda for rituximab refractory patients, based on comparator data from a study by Sehn et al. (2016)[40] comparing O-Benda with bendamustine monotherapy (only the O-Benda arm was used in the analyses).

ERG comment: As explained in Section 3.3 of this report, NICE does not consider O-Benda a relevant comparator for disease that is refractory to rituximab. Therefore, this comparison will be ignored.

In the response from NICE to comments on the draft scope, NICE stated that "obinutuzumab in combination with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4).[28] When NICE recommends a drug for use within the Cancer Drugs Fund (CDF), NICE considers that there is plausible potential for the drug to satisfy the criteria for routine commissioning, but there is significant remaining clinical uncertainty which needs more investigation, through data collection in the NHS or clinical studies.[27] This means that the cost effectiveness of drugs recommended for use within the CDF has not yet been established. Therefore, any comparisons of effectiveness or cost effectiveness with CDF-drugs are equally uncertain.

Second, the company performed an indirect comparison with data from the Haematological Malignancy Research Network (HMRN). This included a comparison of R^2 with:

- Pooled data for R-CHOP/R-CVP for non-rituximab refractory patients.

There was no data for R-refractory patients receiving O-Benda in the HMRN database (due to this regimen only being recently available) and so this data source was not used for this population.

ERG comment: The company stated that 'Due to small patient numbers for non-R-refractory patients receiving R-CHOP and R-CVP in the HMRN database, clinical expectation that R-CHOP and R-CVP would have similar efficacy in a relapsed/refractory setting and empirical data demonstrating this to be the case, efficacy analyses compared R^2 to the pooled R-CHOP/R-CVP' (CS, page 84).[1] The ERG was not convinced by this statement, and asked the company to provide further clarification (Clarification Letter, Question A14).[26] The company responded that 'Data for R-CHOP and R-CVP have been pooled given clinical feedback that it is not unreasonable to assume similar efficacy between R-CHOP and R-CVP in the relapsed/refractory setting, and HMRN clinical data supporting this'.[26] However, looking at the Advisory Board document provided by the company,[19] no such statement is included; therefore, it is not clear how this clinical feedback was obtained. In addition, clinicians did advise that '[redacted]', while '[redacted]'.[19] This suggests that R-CHOP and R-CVP are generally considered for different types of patients, making a comparison of the effectiveness of the two drugs problematic. Clinicians also advised that '[redacted]'.[19]

The company also provided data from the HMRN database, to show that R-CHOP and R-CVP have similar effectiveness. However, these data are based on small numbers of patients (63 in total; [redacted] for

R-CHOP and [redacted] for R-CVP). Analyses of OS and PFS using Cox proportional hazards models showed no significant difference between treatments after adjusting for other covariates (age, prior lines of therapy, early relapse, stage, nodal sites and prior rituximab). However, an analysis of a small sample which shows no statistically significant differences between the two treatments does not mean that one can infer they are equivalent and can be combined for further indirect comparisons. Analysis of a larger dataset with sufficient statistical power could lead to a different conclusion. The one covariate that was consistently related to outcome was age, which suggests that R-CVP will be more often considered for elderly patients and R-CHOP will be more often considered for younger patients; which means that the drugs are generally considered for different populations, making a comparison problematic. In conclusion, the ERG does not think the company has presented convincing evidence suggesting that R-CHOP and R-CVP have similar clinical effectiveness.

In the next two sections a critique of the two types of MAIC will be presented: using published evidence and using HMRN data.

4.4.1 MAIC comparing R^2 with R-CHOP based on published evidence.

Table 4.9 shows a list of potential effect modifiers/prognostic variables (EM/PVs) that would ideally be adjusted for in a matching-adjusted indirect comparison (MAIC), as identified and validated by external clinical experts consulted by the company.[19]

Table 4.9: Potential EM/PVs that would ideally be adjusted for in a MAIC

Characteristic	highest priority	Included in MAIC	Comments
Previous exposure to rituximab	Yes		Not included in MAIC Was 0% in Van Oers
FLIPI components: - Age (median if mean no reported) - Ann Arbor Stage (III-IV) - Nodal sites (>4) - High LDH	Yes	 Yes Yes	 No data reported in Van Oers Not included in MAIC
Refractory to last therapy	Yes	Yes	
Prior lines of therapy 1 vs. 2 vs. >2	Yes	Yes (2 and 3+)	One prior line of therapy was not included
FLIPI risk group (low vs. intermediate vs. high)	Yes	Yes (medium and high)	Low FLIPI risk was not included
FLIPI2+ components: - Serum beta-2 microglobulin high - Bone marrow involvement - Diameter of largest node >6 cm - Haemoglobin <12 dL/L			 No data reported in Van Oers Not included in MAIC No data reported in Van Oers No data reported in Van Oers
Time from last treatment			No data reported in Van Oers
POD24			No data reported in Van Oers
ECOG performance status (0–1 vs. 2+)			No data reported in Van Oers
Presence of B-symptoms			Not included in MAIC
Source: CS, Section B.2.9, pages 70-71.			

EM = effect modifiers; ESS = effective sample size; FL = follicular lymphoma; FLIPI = Follicular Lymphoma International Prognostic Index; MAIC = matching-adjusted indirect comparison; MZL = marginal zone lymphoma; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.
Notes: Adjusted N is the sum of the absolute weights. The patient characteristics presented are the potential EM/PVs that were included in the matching. The following potential EM/PVs had data for all included studies but were dropped from the matching to maintain a sufficiently large effective sample size for subsequent analysis: % previous rituximab exposure. The ESS and adjusted N including these variables were 4.2 and 0.1.

The company stated that 'if the adjustment resulted in an expected sample size and/or adjusted number of patients that was too small for analysis, then the list of variables used for adjustment was reduced before analysis. This was done to maintain the maximum number of the most clinically important variables in the adjustment. Several combinations of variables were explored. However, note that excluding known imbalanced covariates from matching may result in populations with differing levels of effect modifiers/prognostic variables on each treatment, which can bias the analysis results' (CS, page 71).[1]

Clinical advisors consulted by the company,[19] agreed that the most significant factor to be considered for MAIC of AUGMENT and MAGNIFY compared with comparator studies was prior rituximab exposure. Other important factors noted by advisors were FLIPI score, age, refractoriness to last therapy, duration of prior response and number of prior therapies. 'If inclusion of one or more of these factors in the MAIC is not possible, particularly with respect to prior rituximab experience, or where their application sufficiently reduces effective sample size, the credibility of comparison of the rituximab-non-refractory patient data from AUGMENT/MAGNIFY with published data for R-CHOP and R-bendamustine would be limited.'[19]

ERG comment: Clinical advisors agreed that the most significant factor to be considered for MAIC was prior rituximab exposure, yet this could not be included in the MAIC as none of the patients in Van Oers had prior rituximab. Therefore, the credibility of comparison of the rituximab non-refractory patient data from AUGMENT with published data for R-CHOP is limited, according to the clinical advisors consulted by the company. Previous rituximab use was one of the major exclusion criteria in the study by Van Oers et al. (2006).[39] That means that all patients in Van Oers et al. are 100% rituximab-naïve and that the study is not reflective of UK practice, as acknowledged by the company (CS, page 101).[1] Several covariates were not included in the MAIC because data were not reported in the study by Van Oers et al. (2006).[39] Although this is through no fault of the company, it affects the reliability of the MAIC results as all possible covariates present in both studies should be adjusted for.

Standard methods for MAIC were used as recommended in NICE DSU TSD report 18.[41] Individual patient data (IPD) from AUGMENT and summary data from the Van Oers et al. (2006) study[39] (for rituximab-naïve FL patients only) were used for the comparisons in the non-R-refractory population. The IPD from AUGMENT was matched to the R-CHOP data to ensure similar baseline characteristics using recommended weighting methods. The matching used the maximum set of covariates (based on what was available in both studies but excluding previous rituximab exposure).

For the analysis of OS and PFS using the matched data, pseudo-IPD data were generated from the published KM curves using the Guyot method for digitising curves.[42] This data was compared to the IPD survival data for R^2 using a number of statistical methods: KM curves, a Cox proportional hazards model; and different parametric survival models (exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma). The proportional hazards assumption and underlying assumptions of the parametric models were assessed.

The results of the matching for the EMs/PVs included in the matching are provided in Table 4.10 for all covariates included.

Table 4.10: Patient characteristics, observed and match-adjusted for the non-R-refractory population (FL and MZL), comparing R^2 (AUGMENT) and R-CHOP (Van Oers 2006)

Characteristic	AUGMENT (R^2) (n=178)	Van Oers (R-CHOP) (n=234)	Adjusted R^2 (n=78.8)
Patient characteristics			
% refractory	16.9	16.0	16.0
% Ann Arbor stage III-IV	77.0	100.0	100.0
% FLIPI medium	30.9	33.0	33.0
% FLIPI high	38.8	37.0	37.0
% 2 prior lines of therapy	17.4	22.0	22.0
% 3+ prior lines of therapy	25.3	0.0	0.0
Age	62.3	54.0	54.0
Outcomes			
OS	Not estimable	NR	NR
HR (95% CI)		[redacted]	
PFS (N, median (95% CI))	178, 39.4 months (NR)	234, 33.1 months (NR)	78.8, 30.4 months (NR)
HR (95% CI)		[redacted]	

Source: CS, Appendix D2, Table 15, page 36.
EM = effect modifiers; ESS = effective sample size; FL = follicular lymphoma; FLIPI = Follicular Lymphoma International Prognostic Index; MZL = marginal zone lymphoma; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.
Notes: Adjusted N is the sum of the absolute weights. The patient characteristics presented are the potential EM/PVs that were included in the matching. The following potential EM/PVs had data for all included studies but were dropped from the matching to maintain a sufficiently large effective sample size for subsequent analysis: % previous rituximab exposure. The ESS and adjusted N including these variables were 4.2 and 0.1.

ERG comment: The comparison of R^2 versus R-CHOP was adjusted for the variables listed in Table 4.10, i.e. percentage of patients that were refractory, Ann Arbor score and FLIPI score, prior lines of therapy and age. Most of these variables were already reasonably balanced between group, except the percentage of patients with three or more prior lines of therapy, which was 0% in the study by Van Oers and 25% in AUGMENT. The MAIC results for the comparison of R^2 versus R-CHOP are only applicable to the population of Van Oers et al. (2006).[39] This means, patients with rituximab-naïve FL only, all patients had one or two prior lines of therapy (none had three or more), and all patients had Ann Arbor stage III-IV.

As mentioned previously, the most significant factor according to clinical experts to be considered for the MAIC of R^2 (AUGMENT) compared with R-CHOP (Van Oers) was prior rituximab exposure; but this was not included in the MAIC because all patients in the comparator study (Van Oers et al. (2006)) were rituximab-naïve. Another important factor noted by clinical experts was duration of prior response; this was also not included as a covariate in the MAIC. Therefore, the credibility of the MAIC is limited and results are not representative for the UK patient population. The company also concluded that the 100% rituximab-naïve population in Van Oers is not reflective of UK practice and used data from UK HMRN in the economic base-case analysis instead. Therefore, the indirect comparison using HMRN data will be critiqued next.

4.4.2 Indirect comparison of R^2 with R-CHOP/R-CVP based on HMRN data.

The company performed an indirect comparison with data from the Haematological Malignancy Research Network (HMRN). This included a comparison of R^2 with pooled data for R-CHOP/R-CVP for non-rituximab refractory patients.

'Due to small patient numbers for non-R-refractory patients receiving R-CHOP and R-CVP in the HMRN database, clinical expectation that R-CHOP and R-CVP would have similar efficacy in a relapsed/refractory setting and empirical data demonstrating this to be the case, efficacy analyses compared R^2 to the pooled R-CHOP/R-CVP' (CS, page 84).[1] As explained in section 4.4 of this report, the ERG does not think the company has presented any convincing evidence suggesting that R-CHOP and R-CVP have similar clinical effectiveness. The ERG believes the treatments are generally considered for different populations and their effectiveness is therefore difficult to compare.

There were 63 patients identified as receiving either R-CVP or R-CHOP as second- or later-line therapy. Comparisons were made for three time to event outcomes collected within the AUGMENT clinical study (OS, TTNLT and PFS). The definition of TTNLT as used for the HMRN analysis is time to documentation of new anti-lymphoma treatment from 'baseline'. The definition of PFS as used for the HMRN analysis is time from 'baseline' to disease progression (including transformation to diffuse large B-cell lymphoma) or death due to any cause and the definition of OS was time from start of treatment to date of death or if still alive censored at 18 December 2018.

The HMRN is a population-based cohort, established in 2004, comprising a total population of 3.8 million people covering the former adjacent UK Cancer Networks of Yorkshire and the Humber & Yorkshire Coast. The HMRN identified [redacted] patients who had received ≥1 prior line of chemotherapy for treatment of FL and were identified as being non-R-refractory or R-refractory after each treatment line. For the subgroup of patients who were non-R-refractory, [redacted] patients received R-CVP and [redacted] patients received R-CHOP as a second or later line therapy, although most patients ([redacted]%) received these treatments in second-line. Patients could be included in both treatment subgroups if they had received both treatments in different lines of therapy, for example, R-CHOP in second-line and R-CVP in third-line. The HMRN dataset only includes FL patients, not MZL patients (CS, Section B.3.3, page 134).[1]

The baseline characteristics that were commonly collected by the HMRN and the AUGMENT study are presented in Table 4.11.

Table 4.11: Covariates commonly collected across AUGMENT and HMRN datasets

Data source	HMRN	AUGMENT
Treatment	**R-CVP/R-CHOP (2L+ population)**	**R^2**
N	[redacted]	[redacted]
Age (years):		
Median	[redacted]	[redacted]
Range	[redacted]	[redacted]
n (%) Age >=60yrs	[redacted]	[redacted]
n (%) Age >=65yrs	[redacted]	[redacted]
Sex, n, %		
n (%) Males	[redacted]	[redacted]

Data source	HMRN	AUGMENT
n (%) Females	[redacted]	[redacted]
Number of prior systemic anti-lymphoma regimens:		
n (%) 1	[redacted]	[redacted]
n (%) 2	[redacted]	[redacted]
n (%) ≥ 3	[redacted]	[redacted]
Prior rituximab treatment, n (%)	[redacted]	[redacted]
POD24[a], n (%)	[redacted]	[redacted]
Fully Staged, n (%)	[redacted]	[redacted]
Bone marrow involved, n (%)	[redacted]	[redacted]
Nodal sites		
n (%) ≤4	[redacted]	[redacted]
n (%) >4	[redacted]	[redacted]
Bulky disease [b]	[redacted]	[redacted]
Stage		
n (%) I	[redacted]	[redacted]
n (%) II	[redacted]	[redacted]
n (%) III	[redacted]	[redacted]
n (%) IV	[redacted]	[redacted]

Source: CS, Appendix D3, Table 28, page 55.
2L+ = second or later line therapy; HMRN = Haematological Malignancy Research Network; NA = not applicable; R^2 = rituximab plus lenalidomide; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP = rituximab plus cyclophosphamide vincristine prednisolone.
[a]) POD24 is defined as relapse within two years of initial chemoimmunotherapy.
[b]) Bulky disease has different definitions in AUGMENT and the HMRN dataset. AUGMENT: At least one lesion that is ≥ 7 cm or at least 3 lesions with 3 cm or larger in the longest diameter by investigator review. HMRN: At least one lesion that is ≥ 10 cm.

The same list of potential modifiers/prognostic variables discussed previously in the context of the ITC with published data, was used to identify the matching variables for this comparison. Therefore, Table 4.12 shows the same list of potential effect modifiers/prognostic variables (EM/PVs) that would ideally be adjusted for in a MAIC, as identified and validated by external clinical experts consulted by the company.[19]

Table 4.12: Potential EM/PVs that would ideally be adjusted for in a MAIC

Characteristic	highest priority	Included in MAIC	Comments
Previous exposure to rituximab	Yes	Yes	
FLIPI components: - Age (mean, or median if mean no reported, or % >60 years if neither reported) - Ann Arbor Stage (III-IV) - Nodal sites (>4) - High LDH	Yes	Yes Yes Yes	Included as: % Age ≥60yrs Not collected in HMRN
Refractory to last therapy	Yes	No	Not included in MAIC
Prior lines of therapy 1 vs. 2 vs. >2	Yes	Yes	
FLIPI risk group (low vs. intermediate vs. high)	Yes	No	Not collected in HMRN
FLIPI2+ components: - Serum beta-2 microglobulin high - Bone marrow involvement - Diameter of largest node >6 cm - Haemoglobin <12 dL/L			 Not included in MAIC Not included in MAIC Not included in MAIC Not included in MAIC
Time from last treatment			Not included in MAIC
POD24	Yes	Yes	
ECOG performance status (0–1 vs. 2+)			Not included in MAIC
Presence of B-symptoms			Not included in MAIC

Source: CS, Section B.2.9, pages 70-71.
EM = effect modifiers; ESS = effective sample size; FL = follicular lymphoma; FLIPI = Follicular Lymphoma International Prognostic Index; MAIC = matching-adjusted indirect comparison; MZL = marginal zone lymphoma; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.
Notes: Adjusted N is the sum of the absolute weights. The patient characteristics presented are the potential EM/PVs that were included in the matching. The following potential EM/PVs had data for all included studies but were dropped from the matching to maintain a sufficiently large effective sample size for subsequent analysis: % previous rituximab exposure. The ESS and adjusted N including these variables were 4.2 and 0.1.

As can be seen from Table 4.12, matching was performed for the following variables:

- Age ≥60 years (FLIPI component)
- Ann Arbor Stage III-IV (FLIPI component)
- Nodal sites >4 (FLIPI component)
- Prior rituximab treatment
- Prior lines of therapy (1 vs. 2 vs. >2)
- POD24 status

The company stated that 'A key treatment effect modifier/prognostic factor that was not collected by the HMRN was the FLIPI risk category. However, three of the four FLIPI components were collected (only LDH was not collected)' (CS, pages 85-86).[1] Another key treatment effect modifier/prognostic

factor that was not included in the MAIC was 'refractory to last therapy', it is unclear why this factor was not included. In addition, all FLIPI2+ components (Serum beta-2 microglobulin high; bone marrow involvement; diameter of largest node >6 cm; and haemoglobin <12 dL/L), time from last treatment, ECOG performance status (0–1 vs. 2+), and presence of B-symptoms were not included in the MAIC.

Regarding ECOG performance status, the company states that ECOG PS 'was dropped from the MAICs because there were very few ECOG PS 2+ patients in AUGMENT/MAGNIFY, and the comparator studies also either had a small number of ECOG PS 2+ patients (hence were balanced) or did not report these data' (CS, page 71).[1] It was not reported whether ECOG PS was reported in the HMRN dataset.

The company was asked why 'sex' and 'bone marrow involved' were not included in the matching. The company responded that 'Sex was not identified as being a potential prognostic factor and/or treatment effect modifier in the list of variables that was validated by external clinical experts and was therefore not included as a matching variable' (Response to Clarification, Question A17a).[26] The company agreed that 'bone marrow involved' should have been considered as a matching variable given that it was identified as being a potential prognostic factor and/or treatment effect modifier.[26] In response to Question A17b, the company performed the comparison to R-CVP/R-CHOP with additional adjustment for bone marrow involvement, and concluded that the addition of this extra variable has had little impact on the results.[26]

In conclusion, several potential treatment effect modifiers/prognostic factors were not included in the MAIC; some because data were not reported in HMRN (FLIPI risk group, LDH,), some because the company regarded it not relevant (sex), and some for reasons that are not clear ('refractory to last therapy', all FLIPI2+ components (Serum beta-2 microglobulin high; bone marrow involvement; diameter of largest node >6 cm; and haemoglobin <12 dL/L), time from last treatment, ECOG performance status (0–1 vs. 2+), and presence of B-symptoms).

The main concerns are the same as for the previous MAIC (Section 4.1.1), i.e. the set of covariates included in the MAIC does not reflect the complete set of all possible covariates which affects the reliability of the OS and PFS results. Several covariates were not included in the MAIC because data were not reported in the HMRN dataset. Although this is through no fault of the company, it is a serious limitation which affects the reliability of the MAIC results.

As stated in Section 4.4.1 of this report, the credibility of the MAIC relies on the inclusion of all relevant treatment effect modifiers/prognostic factors. DSU report TSD 18[41] states that, 'An unanchored MAIC or STC effectively assumes that absolute outcomes can be predicted from the covariates; that is, it assumes that all effect modifiers and prognostic factors are accounted for. This assumption is very strong, and largely considered impossible to meet. Failure of this assumption leads to an unknown amount of bias in the unanchored estimate'.[41] As can be seen from the list of covariates included in the MAIC, it is clear that several treatment effect modifiers/prognostic factors were not included in the MAIC, including some that were considered key treatment effect modifiers/prognostic factors by the clinicians consulted by the company (FLIPI risk group, and 'refractory to last therapy').

The results of the matching for the EMs/PVs included in the matching are provided in Table 4.13 for all covariates included.

Table 4.13: Patient characteristics, observed and match-adjusted for the non-R-refractory population, comparing R^2 (AUGMENT) and R-CHOP/R-CVP (HMRN)

Characteristic	AUGMENT (R^2) (n=178)	HMRN (R-CHOP/ R-CVP) (n=63)	Adjusted R^2 (n=[redacted])
Patient characteristics			
% Prior rituximab	[redacted]	[redacted]	[redacted]
% Age ≥60yrs	[redacted]	[redacted]	[redacted]
% Ann Arbor stage III-IV	[redacted]	[redacted]	[redacted]
% Nodal sites ≤4	[redacted]	[redacted]	[redacted]
% 1 prior lines of therapy	[redacted]	[redacted]	[redacted]
% 2 prior lines of therapy	[redacted]	[redacted]	[redacted]
% Early relapse	[redacted]	[redacted]	[redacted]
Outcomes			
OS	Not estimable	63, [redacted] (NR)	NR
PFS (N, median (95% CI))	178, 39.4 months (NR)	63, [redacted] (NR)	NR

Source: CS, Appendix D3, Table 29, page 57.
EM = effect modifiers; ESS = effective sample size; FLIPI = Follicular Lymphoma International Prognostic Index; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

ERG comment: The comparison of R^2 versus R-CHOP/R-CVP was adjusted for the variables listed in Table 4.13, i.e. percentage of patients that had prior rituximab, age, Ann Arbor score and FLIPI score, nodal sites, prior lines of therapy and early relapse. The resulting population for the comparison of R^2 versus R-CHOP/R-CVP, are patients based on the baseline characteristics of patients in the HMRN dataset. As mentioned previously, two key treatment effect modifiers/prognostic factors (FLIPI risk group, and 'refractory to last therapy') were not included in the matching process. In addition, several covariates were not included in the MAIC because they were not reported in the HMRN dataset. Therefore, the credibility of the MAIC is limited.

Results of the MAIC are presented in Kaplan-Meier curves for OS, PFS, and TTNLT (CS, Figures 17-19, pages 88-90).[1] Hazard ratios (HRs) from the Cox Proportional-Hazard models comparing R^2 and R-CHOP/R-CVP are reproduced in Table 4.14. R^2 had a significant improvement in OS and TTNLT compared to R-CHOP/R-CVP and a benefit for TTNLT, but no evidence of a difference in PFS.

Table 4.14: Results from Cox Proportional Hazard models comparing R^2 and R-CVP/R-CHOP

Outcome	R^2, adjusted N	R-CHOP/R-CVP, N	HR (95% CI)[a]
OS	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]
TTNLT	[redacted]	[redacted]	[redacted]

CI = confidence interval; HR = hazard ratio; N = number of patients; OS = overall survival; PFS = progression-free survival; R^2 = rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP = rituximab plus cyclophosphamide, vincristine, prednisolone; TTNLT = time to next anti-lymphoma treatment.
[a]) bootstrapped CI.

The company should have presented crude unadjusted differences alongside the MAIC estimates, in line with the recommendations in NICE Decision Support Unit (DSU) technical support document (TSD) 18[41] to enable comparisons between the adjusted MAIC and unadjusted results. No such estimates have been presented, apart from the Kaplan-Meier curves in Figures 17-19 (CS, pages 88-90).[1]

NICE DSU TSD report 18 lists several themes that should be considered and addressed explicitly when reporting population-adjusted analyses (See TSD 18, pages 64-65).[41] In Appendix 2 these themes are reproduced with an ERG comment how they were addressed in this submission. As can be seen from Appendix 2 not all themes were addressed in the CS.

In conclusion the results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were excluded from the matching models, small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources, The analysis also used an unanchored MAIC involving two single treatment arms from different studies, as there was no relevant comparative trial data. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No further additional work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches. A good range of resources were searched and the searches were transparent and reproducible. One set of searches was conducted to identify both efficacy and safety evidence. Separate searches were conducted to identify cost effectiveness studies, health-related quality of life studies, and healthcare resource use data.

The company submission included six studies that were deemed relevant by the company. Four studies evaluated R^2, one of these was an RCT of R^2 versus R-monotherapy (the AUGMENT trial[35]), the other three[36-38] did not include relevant comparators according to the NICE scope. The remaining two studies evaluated R-CHOP versus CHOP (Van Oers et al., 2006[39]) and O-Benda versus bendamustine monotherapy (the GADOLIN trial[40]). The trial by Van Oers et al. (2006)[39] was used by the company for an unanchored indirect comparison (using individual arms of different studies) of R^2 versus R-CHOP. However, the study only included rituximab-naïve patients and was therefore not representative for the UK patient population. The GADOLIN study[40] was used by the company for an unanchored indirect comparison of R^2 with O-Benda. However, as explained in Sections 3.3 and 4.4 of this report, O-Benda is not considered by NICE to be a relevant comparator for this appraisal; therefore, this study was ignored in this report.

In conclusion, the CS included one relevant study, for the comparison of R^2 versus R-monotherapy: the AUGMENT trial.[35] All patients in this trial were non-R-refractory. In addition, the company performed an unanchored indirect comparison of R^2 versus R-CHOP and R-CVP, using data for R^2 from the AUGMENT trial and pooled data for R-CHOP/R-CVP from the HMRN database.

The AUGMENT trial[35] is a randomised, double-blind, multicentre, Phase III study of R^2 versus rituximab plus placebo (R-mono) in non-R-refractory patients with FL Grade 1–3a or MZL. The study was conducted across 96 sites in 17 countries. The trial did not include any patients from the UK. The primary efficacy analyses were conducted on the ITT population, defined as all randomised patients. The primary endpoint of the study was PFS, as assessed by the Independent Review Committee (IRC).

Results from the AUGMENT trial show favourable results for R^2 when compared to R-mono in terms PFS with a greater median PFS ([redacted] vs. [redacted] months; HR [redacted] (95% CI: [redacted]). However, there was no evidence of a difference in OS with a HR of 0.61 (95% CI: 0.33 to 1.13) for patients treated

with R^2 compared to R-mono. At the time of the analysis the OS data was immature with 16 deaths on R^2 and 26 deaths on R-mono at the time of the analysis. Overall response rate (ORR) was significantly greater for R^2 compared with R-mono (78% vs. 53%; p<0.0001). The complete response (CR) rate was also greater for the R^2 arm compared with R-mono (34% vs. 18%; p=0.001). Results for R^2 versus R-mono in MZL patients were generally less favourable for R^2 than in FL patients. However, it is important to note that PFS outcomes in the MZL subgroup are difficult to interpret because of the small sample size (63 patients in total) and imbalance in baseline prognostic factors. In terms of health-related quality of life, no clinically meaningful change from baseline in the GHS/QoL domain of the QLQ-C30 was observed across any of the post-baseline assessment visits, regardless of treatment group. Between-group differences in mean changes were small and not clinically meaningful across all assessment visits and did not differ between FL and MZL patients.

R^2 was associated with more grade 3-4 TEAEs and SAEs when compared to R-mono, especially lenalidomide/placebo related adverse events; but rituximab-related grade 3-4 TEAEs and SAEs were also more frequent in the R^2 arm than in the R-mono arm. R^2 was also associated with more TEAEs leading to dose reductions, dose interruptions and discontinuations of lenalidomide/placebo or rituximab when compared to R-mono. Adverse events are generally the same for FL and MZL patients; however, AEs for MZL patients are based on small numbers.

The company performed three unanchored indirect comparisons, two using data from published evidence and one using data from HMRN:

- R^2 versus R-CHOP for non-rituximab refractory patients, based on comparator data from a study by Van Oers et al. 2006[39] comparing R-CHOP with CHOP (only the R-CHOP arm was used in the analyses).
- R^2 versus established clinical management without lenalidomide, i.e. O-Benda for rituximab refractory patients, based on comparator data from a study by Sehn et al. 2016[40] comparing O-Benda with bendamustine monotherapy (only the O-Benda arm was used in the analyses).
- R^2 versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN.

As mentioned above, the two unanchored indirect comparisons using published evidence have been ignored in this report. R^2 versus R-CHOP, because the study by Van Oers is not representative for UK patients, and R^2 versus O-Benda because O-Benda is not a relevant comparison for this appraisal according to NICE.

Results from the remaining MAIC (R^2 versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN) show a significant improvement in OS (HR = [redacted] (95% CI: [redacted]) and TTNLT (HR = [redacted] (95% CI: [redacted]) compared to R-CHOP/R-CVP, but no evidence of a difference in PFS (HR = [redacted] (95% CI: [redacted]).

The results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were excluded from the matching models, small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources, The analysis also used an unanchored MAIC involving two single treatment arms from different studies, as there was no relevant comparative trial data. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

The company conducted searches for cost effectiveness, health-related quality of life and healthcare resource use evidence. A good range of databases, conference proceedings and additional resources were searched. The company submission and clarification response provided sufficient detail for the ERG to be able to appraise the searches conducted by the company.

5.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Appendices G, H and I of the CS report the literature searches used to identify cost effectiveness, health-related quality of life and healthcare resource use studies. Separate sets of searches were run for each. Searches were conducted on 8 February 2019. A summary of the sources searched is provided in Table 5.1. The CS reported that targeted literature searches were conducted to identify adverse event disutility values, FL and MZL prognosis studies, and data on response rates, OS and PFS: these targeted searches were not provided. The company described how these data were identified via targeted literature searches in their response to the ERG clarification letter.

Table 5.1: Resources for the cost effectiveness, health-related quality of life and healthcare resource use literature searches

Search strategy element	Resource	Host/Source	Date Range	Date Searched
Electronic databases	MEDLINE	Embase.com	Not reported	8 February 2019
	Embase			
	MEDLINE In-Process	PubMed	Not reported	8 February 2019
	EconLit	EBSCO	Not reported	8 February 2019
	NHS EED	CRD interface	Not reported	8 February 2019
	HTA		Not reported	8 February 2019
Conference proceedings	ISPOR International	http://www.ispor.org/heor-resources/presentations-database/search	2017, 2018	February 2019
	ISPOR European	http://www.ispor.org/heor-resources/presentations-database/search	2017, 2018	February 2019
	ASH	http://www.hematology.org/Annual-Meeting/Archive.aspx	2017, 2018	February 2019
	EHA	https://ehaweb.org/congress/previous-congresses/	2017, 2018	February 2019
	ICML	http://www.lymphcon.ch/icml/website/icml-abstracts-books/icml-abstract-books-1981-2011.html	2015, 2017	February 2019

Search strategy element	Resource	Host/Source	Date Range	Date Searched
	ASCO	https://meetinglibrary.asco.org/browse-meetings/	2017, 2018	February 2019
HTA Agencies	NICE	https://www.nice.org.uk/		February 2019
	SMC	https://www.scottishmedicines.org.uk/		February 2019
	AWMSG	http://www.awmsg.org/		February 2019
	HAS	https://www.has-sante.fr/portail/jcms/r_1455081/en/home-page?portal=r_1455081		February 2019
	SLV	https://legemiddelverket.no/English https://www.legemiddelsok.no/		February 2019
Bibliographic searches of key systematic review and meta-analysis articles were conducted to ensure that initial searches captured all the relevant economic studies				

HTA = Health Technology Assessment Database; NHS EED = NHS Economic Evaluation Database; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; ASH = American Society of Hematology; EHA = European Hematology Association; ICML = International Conference on Malignant Lymphoma; ASCO = American Society of Clinical Oncology; NICE = National Institute for Health and Care Excellence; SMC = Scottish Medicine Consortium; AWMSG = All Wales Medicines Strategy Group; HAS = The Haute Autorité de Santé; SLV = Statens legemiddelverk.

ERG comment:

- MEDLINE and Embase were searched simultaneously using embase.com. This approach is not recommended. A simultaneous multi-file search such as this should include both MeSH and EMTREE subject headings to ensure that all subject indexing terms are searched; however, all of the economic search strategies only included EMTREE terms which may have impaired how well the strategies performed.
- There were no details about which MEDLINE segments were searched (Table 35, Table 44 and Table 54 in Appendix G of the CS).[34]
- Date ranges were not reported for any of the economic related database searches.
- The CS reported that MEDLINE In-Process was searched using PubMed (Table 36, Table 45 and Table 55). This is inaccurate, as the search limit used in PubMed identifies 'Ahead of print' and recently added records, not in-process records: (publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatusaheadofprint). Therefore in-process records were actually excluded from the company's PubMed search.
- The company reported searching NHS EED and the HTA database via the Cochrane Library using the CRD search interface. This is incorrect as NHS EED and HTA are no longer available on the Cochrane Library or have anything to do with Cochrane. The company conducted the NHS EED and HTA searches via the Centre for Reviews and Dissemination (CRD) interface, and misreported using the Cochrane Library.
- Truncation and proximity operators were used more often in the cost effectiveness searches than in the clinical effectiveness searches. As with the clinical effectiveness searches, there were too few synonyms. However, the 'syn' operator was included, and embase.com enables automatic synonym searches when this operator is added to an EMTREE term. The ERG does not have access to Embase.com to test the impact of this on search performance.
- The search strategies used in MEDLINE In-Process (PubMed), EconLit, and NHS EED/HTA only included a population facet of search terms, and so were sensitive enough to identify studies for all of the economic sections (cost effectiveness, health-related quality of life and healthcare resource

use). The embase.com search strategies included an additional facet of search terms for each of the economic sections (cost effectiveness, health-related quality of life and healthcare resource use); three separate searches were conducted in embase.com.

- It is not clear if the search facets used to identify cost effectiveness, health-related quality of life and healthcare resource use were based on validated search filters, such as those published on the ISSG Search Filters Resource website: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/
- A good range of conference proceedings and HTA organisation websites were searched, and although full details of these searches were not provided in the CS, they were provided in response to the ERG clarification letter.
- Targeted literature reviews were referred to in the CS, but no details were reported. In response to the ERG clarification letter the company provided details of the targeted literature reviews, and how adverse event disutility values, FL and MZL prognosis studies, and data on response rates, OS and PFS were identified. Data were identified by investigating the clinical systematic literature review results, reviewing previous NICE technology appraisals, and a targeted literature search of PubMed.

5.1.2 Inclusion/exclusion criteria used in the study selection

In- and exclusion criteria for the review on cost effectiveness studies, utilities and costs and resource use are presented in Table 39 of Appendix G, Table 48 of appendix H, and Table 58 of Appendix I of the CS, repectively.[34]

ERG comment: The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies.

5.1.3 Included/excluded studies in the cost effectiveness review

In total, 24 cost effectiveness studies met the pre-defined eligibility criteria.[13, 20, 22, 23, 43-62] These were extracted from 31 publications of which 22 full publications and nine HTA submissions. Details of these studies were provided in Tables 23 and 24 of the CS. The search for utility studies resulted in 38 included studies, for which details and references were provided in Table 49 of Appendix H of the CS.[34] The search for costs and resource use resulted in 17 included studies, for which details and references were provided in Table of Appendix I of the CS.[34]

ERG comment: The rationales for excluding CE studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

5.1.4 Conclusions of the cost effectiveness review

The CS provides an overview of the included cost effectiveness, utility and resource use and costs studies, but no specific conclusion was formulated.

ERG comment: Eligibility criteria were suitable for the SLR performed. However, it was not fully clear to the ERG how the information obtained from the SLR was implemented in the de novo analysis. For instance, the company stated in B.3.1 of the CS that they had identified four economic evaluations that had a UK perspective and were of potential value to inform this submission. They then stated that 'more details of how these evaluations have informed the de novo analysis are discussed in Section B.3.2.'.[1] However, Section B.3.2. of the CS does not contain any information on the use of these evaluations.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.2: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS)
Model	Partitioned survival model	Makes use of the PFS and OS data directly, ensuring that estimated survival outcomes versus observed outcomes are matched.	B.3.2
States and events	Progression-free, post-progression, death		B.3.2
Comparators	Non-rituximab-refractory patients: R-CHOP and R-CVP rituximab-refractory patients: O-Benda	Expert opinion	B.3.2
Population	The patient population considered in the model is, adult patients with previously treated FL or MZL (pooled). The model is split into two subpopulations: non-rituximab-refractory and rituximab-refractory patients.	In line with the proposed licence. FL and MZL populations were pooled due to the similar prognosis of FL and MZL patients, and the difficulty in sourcing MZL-specific data	B.3.2
Treatment effectiveness	Non-rituximab-refractory: Unanchored MAIC using AUGMENT and HMRN Rituximab-refractory: Unanchored MAIC using MAGNIFY and GADOLIN		B.3.3
Adverse events	Grade 3 and 4 based on trial data		B.3.3
Health related QoL	EQ-5D-3L data from AUGMENT	NICE reference case	B.3.4
Resource utilisation and costs	NHS and Personal Social Services	NICE reference case	B.3.2
Discount rates	3.5% discount rate was used for utilities and costs	NICE reference case	B.3.2.
Subgroups	non-rituximab-refractory and rituximab-refractory patients		B.3.9
Sensitivity analysis	Probabilistic and deterministic sensitivity analyses and scenario analyses	NICE reference case	B.3.8

FL, follicular lymphoma; MZL, marginal zone lymphoma; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; O-Benda, obinutuzumab plus bendamustine; MAIC, matching-adjusted indirect comparison; HMRN, Haematological Malignancy Research Network.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.3: NICE reference case checklist

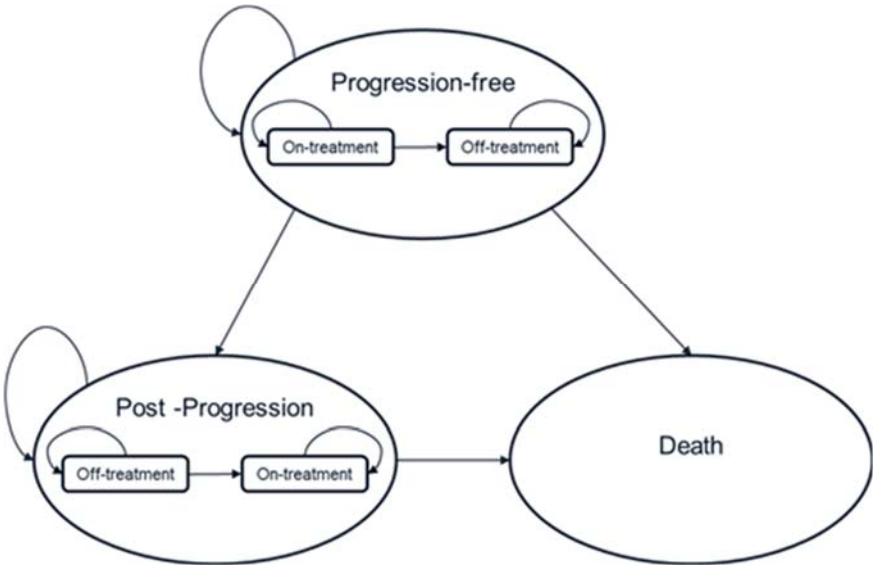
Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope	Yes, although divided in non-rituximab-refractory and rituximab-refractory patients	
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Non-rituximab-refractory patients: R-CHOP and R-CVP Rituximab-refractory patients: O-Benda	R-mono was not included in the evaluation while it was listed in the scope. The company added a comparison of R^2 and R-mono in response to clarification questions. NICE have explicitly stated that O-Benda is not considered a relevant comparator for disease that is refractory to rituximab. The ERG report does not contain information on this comparator.
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	
Synthesis of evidence in outcomes	Systematic review (SLR)	Yes	
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of	Time-trade off or standard gamble	Yes	

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
changes in HRQoL			
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	Yes	
NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; SLR = systematic literature review			

5.2.2 Model structure

A cohort-level partitioned survival model (PSM) was developed with three health states: progression-free (PF), post-progression (PP) and death. The company argued that a PSM was more appropriate than a state transition model (STM) because of a lack of data on post progression survival (PPS). According to the company, the relevant comparators for this submission are not included in the head-to-head study with R^2 (AUGMENT); therefore, the data available for informing PPS for the comparators are reduced to available published data or alternative sources. All patients start 'on treatment' in the PF health state. Subsequently, patients either remain on treatment or come off treatment before progressing or dying per cycle. Within PP, patients can have a treatment-free interval before receiving subsequent therapy. Patients in the PP on treatment health state remain in this health state until they die. The model was programmed in Microsoft Excel.

Figure 5.1: Model structure



Source: Based on Figure 21 of the CS

ERG comment: The main concern of the ERG relates to the use of a PSM instead of a STM. The use of a PSM instead of a STM was justified by the lack of data of relevant comparators in the head-to-head study with R^2 to inform a state transition model. Despite the potential limitations of a state transition model, a partitioned survival analysis has several limitations related to the extrapolation (as mentioned in NICE DSU TSD 19[63]). The ERG requested a scenario analysis using a STM as a scenario, as recommended in TSD 19, which the company did not deliver. The company argued that because of the weight of the limitations in the STM approach, combined with the specifics of the data available for this decision problem, constructing a state transition model is not applicable for this submission. The ERG acknowledges that every model approach has its limitations, and that the lack of data for the R-CHOP and R-CVP posed a problem populating a STM. However, the lack of a structural link between endpoints in a PSM may lead to biased extrapolations.[63] Therefore, according to the ERG, and in line with recommendations from TSD 19, STM should be used alongside PSM to assess the plausibility of extrapolations, if only for the comparison in the pivotal trial.

5.2.3 Population

R^2 does not currently have a UK marketing authorisation. The patient population considered in the model is in line with the proposed license: adult patients with previously treated FL or MZL. Due to the similar prognosis of FL and MZL patients, and the difficulty in sourcing MZL-specific data, FL and MZL populations were pooled throughout the economic analysis. Non-rituximab refractory patients and rituximab refractory patients were modelled separately because the company assumed the relevant comparators for these patients would be different. The patient cohort considered in the model varies per population. The patient starting age and gender were matched to the data source used for the comparator arms (for non-R refractory patients this was the HMRN: mean age [redacted] years, percentage female [redacted]). Body surface area (BSA) data were taken from individual patients in the AUGMENT study (mean BSA 1.85 m^2).

ERG comment: The main concern of the ERG relates to pooling the FL and MZL populations throughout the economic analysis. In response to clarification question B1 the company provided an overview of the population the evidence in the economic analysis was based on. All evidence of the comparators was based on datasets that only contained patient with FL, while the AUGMENT trial

contained patients with FL and MZL. The AUGMENT trial was used as the source for utilities for R^2 as well as the comparators, and as the source of subsequent treatments for R^2. Furthermore, the company provided exploratory post-hoc analyses which investigated the impact of the histology (MZL/FL) on the outcome of PFS in the AUGMENT trial data to justify that the prognosis and comparative effectiveness are similar for FL and MZL. These analyses showed that neither the interaction term between the randomised treatment arm and histology, nor histology were statistically significant (p-value >0.05). The company argued that clinicians during the expert meeting stated resource use for FL and MZL patients was similar. Analysis of AUGMENT quality of life data showed that if histology was included in the mixed effects regression model used for the utilities, this resulted in a mean utility difference of 0.03 for MZL patients, however this was not statistically significant (p=0.145). The company provided an FL-only scenario analysis (discussed in section 5.2.11).

5.2.4 Interventions and comparators

R^2 does not currently have a UK marketing authorisation. The R^2 dosing regimen within the model is lenalidomide 20 mg orally once daily on days 1–21 of repeated 28-day cycles for up to 12 cycles of treatment. Rituximab is given as 375 mg/m2 every week in Cycle 1 (days 1, 8, 15 and 22) and Day 1 of every 28-day cycle for Cycles 2–5. This is in line with the recommended dose in the SmPC.[18] Patients with moderate renal impairment start on a dose of 10 mg of lenalidomide if CrCl is ≥30 ml/min but <60 ml/min. These criteria were met by [redacted]% of patients in AUGMENT and [redacted]% in MAGNIFY (R-refractory population), and these proportions are used to inform the starting dose in the model for the non-R-refractory and R-refractory populations, respectively.

In AUGMENT R^2 is compared to R-mono. The company states that according to clinical experts, R-mono is rarely used in the relapsed/refractory setting in UK clinical practice.[31] Instead, comparators for R^2 in the non-R-refractory population are rituximab in combination with chemotherapy; predominantly R-CHOP and R-CVP. Experts also stated that R-Benda is primarily used in a first-line setting and clinicians are reluctant to re-challenge relapsed/refractory patients with bendamustine in subsequent lines of therapy.[19] Therefore, R-mono and R-Benda were not considered relevant comparators for the non-R-refractory population. For the R-refractory population the company states that clinical experts believe that O-Benda has largely replaced use of bendamustine.[19]

ERG comment: The main concerns of the ERG relate to: a) the inclusion of O-Benda as a comparator while NICE have explicitly stated it is not considered a relevant comparator for disease that is refractory to rituximab, b) omitting R-mono as a comparator (based on expert opinion) although listed in the scope and given the direct evidence available.

a) The ERG did not include O-Benda in her review as NICE has explicitly stated it is not considered a relevant comparator for disease that is R-refractory.

b) In response to question B3 the company provided an analysis of R^2 versus R-mono based on the AUGMENT trial data.

5.2.5 Perspective, time horizon and discounting

The analysis took an NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length of 28 days with a 40-year time horizon and half cycle-correction is applied.

ERG comment: In the CS, the company stated that a 40-year time horizon was used. The model output showed this was in fact a life time horizon.

5.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for intervention and comparators was the AUGMENT study[64] for R^2 and HMRN data[65] for R-CHOP and R-CVP. The AUGMENT study is a Phase III, multicentre, double-blind, randomised study comparing R^2 versus R-mono in patients with non-R-refractory/relapsed FL or MZL. Only data from the R^2 arm and from the 22 June 2018 data cut-off were used in the model. The HMRN is a population-based cohort covering the Yorkshire and Humber & Yorkshire Cancer Networks for all patients newly diagnosed with a haematological malignancy between 2004 and 2016. No data on MZL patients was available in the HMRN.

The Phase III study by van Oers et al. (2006)[39] on R-CHOP was not used in the base-case analysis because all patients were R-naïve, which was not thought to be reflective of current clinical practice in the UK. Also, with prior rituximab exposure being an important effect modifier, matching with the R^2 arm of AUGMENT data would be hampered. The van Oers et al. study data were used in a scenario analysis. For R-CVP, no trial-based evidence was found.

As the company considered OS and PFS in HMRN to be similar between R-CHOP and R-CVP, and clinical opinion suggested that in the relapsed/refractory setting it would not be unreasonable to assume the efficacy of R-CHOP and R-CVP to be similar, HMRN data for R-CHOP and R-CVP were pooled. Data from AUGMENT (n=103) were then matched to the pooled data from HMRN for R-CHOP and R-CVP (n=63). For the economic model, this implied that the comparisons of R^2 vs. R-CHOP and R-CVP had identical outcomes for effectiveness (LYs and QALYs) and only differed with respect to costs.

Parametric survival curves were fitted to the matched patient level data from AUGMENT and HRMN and were then used to extrapolate survival beyond study follow-up. Survival analysis was performed for OS, PFS, TTNLT, and ToT (time on treatment). The CS mentioned four criteria for selection of the curves: 1) proportional hazards assumption based on log cumulative hazard plots 2) visual inspection, Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) 3) clinical plausibility, and 4) implausible curve crossings (e.g. OS moving below TTNLT) before 15 years of follow-up.

PFS and ToT data were used to determine the number of patients staying in the PF (on and off treatment) health states. PFS, TTNLT and OS data were used to determine the number of patients transitioning to the PP (on and off treatment) health states. The number of patients transitioning to the death state was derived using OS data.

The curves were adjusted for treatment waning, which was assumed to occur at five years. After this time point, the comparator hazard of progressing or dying was applied to the R^2 arm. This five-year time point was selected for the base-case as the company stated it to be consistent with previous NICE submissions in the same disease area (TA472[20] and TA137[59]).

Overall survival

As the log-cumulative hazard versus time plot for OS suggested that the proportional hazards assumption was violated, stratified models were used. Although AIC and BIC indicated that the exponential distribution fitted best on average, the Weibull distribution was selected for the base-case analysis. The company did not explain why the exponential distribution was not used, but stated that AIC/BIC for Weibull suggested a reasonable fit, and Weibull was also used in TA137[59]. The curves were adjusted for general population mortality (age and gender matched) so overall survival in the model would not exceed survival in the general population.

For the R-mono comparison which was added upon request of the ERG in the response to clarification, the company chose Weibull for both arms again. No rationale or diagnostic plots were provided but statistical fit based on AIC/BIC was best for exponential, like in the R-CHOP/R-CVP comparisons.

Progression free survival

As the KM curves for R^2 and R-CHOP/R-CVP at first slightly diverge but then converge and even overlap, this was suggestive of a non-constant treatment effect. This was confirmed by the log-cumulative hazard plot, which was non-parallel. The company then decided to model the PFS for R^2 using the KM data until the maximum follow-up of 46.7 months, and applied the comparator hazard to extrapolate further. In this way, the company stated in the CS, the relative treatment effect of R^2 vs. R-CHOP/R-CVP based on the MAIC was accurately reflected. Parametric curves were still fitted to each arm to be able to test assumptions used in the base-case. Based on the AIC/BIC, the Weibull distribution makes the best fit to the R-CHOP/R-CVP data whereas the exponential and log-logistic distributions seem to fit the R^2 data best. Nevertheless, the company chose to use the generalised gamma curve in the base-case, because the Weibull would cross the TTNLT curve in R-CHOP/R-CVP at approximately eight years, which would be clinically implausible since it would be unlikely that patients have their next treatment prior to progression in clinical practice.

Finally, the curves were adjusted to ensure that long-term PFS estimates would not be higher than TTNLT or OS.

For the R-mono comparison, a simpler approach was taken, using log-logistic for both arms.

Time to next anti-lymphoma treatment

From the cumulative hazard plot, proportional hazards seemed reasonable but not definitive, and therefore stratified models were used, with unstratified models explored in a scenario. Based on the AIC/BIC, the exponential distribution best fitted the R^2 data, and the log-normal distribution fitted best to R-CHOP/R-CVP. However, as the exponential distribution would result in crossing of PFS and TTNLT around seven years, the company chose the log-normal distribution for the base-case analysis for both arms.

Finally, in line with what was done for PFS, the curves were adjusted to ensure they would not be higher than OS.

For the R-mono comparison, the generalised gamma was used for both arms. AIC and BIC were provided in the model for a series of distributions, but the choice for generalised gamma was not further justified.

Time on treatment

ToT data were used to determine the proportion of patients on treatment to calculate overall drug costs. Parametric survival curves were fitted to the ToT data which, however, produced a poor fit. Therefore, the company chose to use the KM data directly in the model, and maximum treatment durations were used to cap ToT

For the R-mono comparison the same approach was used, that is, KM data were used.

ERG comment: The main concerns of the ERG relate to: a) the uncertainty introduced by the indirect comparison of R^2 with R-CHOP and R-CVP based on only 63 patients - which seems to be underlined by b) the counterintuitive results for the R-mono comparison c) the lack of justification for the choice of time-point at which treatment effect ends d) the seemingly arbitrary way of selecting the curves used

for extrapolating, which seems mainly guided by trying to avoid implausible curve crossings e) in particular the choice of the OS curves and f) the PFS curves, but also g) TTNLT curves. In addition, h) an error was found when running the scenario using van Oers data for efficacy.

a) The ERG has serious doubts about how trustworthy the results of the indirect treatment comparisons with R-CHOP and R-CVP are, given that HMRN data are based on only 63 patients in total, were collected much earlier than the data from AUGMENT, and consist of two treatment regimens which may not be as similar as assumed. Although the ERG appreciates that R-CHOP and R-CVP data were pooled to obtain a larger sample size, it is still small and the pooling may have introduced additional bias as the KM curves from the HMRN report[65] show a rather consistent difference in favour of R-CHOP, which may be a result of the fact that the target population for R-CHOP is the younger and fitter group, enhancing efficacy. Furthermore, data collection for patients in HMRN started much earlier (from 2004 onwards), and a time effect interfering with the treatment effect cannot be ruled out, given the continuous improvements in clinical practice. These changes in clinical practice may be illustrated by the fact that in the modelled subsequent therapies, the proportion of targeted therapies was 0% for the HMRN R-chemo cohort and 6.7% for the R^2 arm in AUGMENT. The uncertainty associated with the indirect comparison was not captured in the model and as such cannot be quantified but its impact is likely substantial.

b) In their response to clarification (question A7b), the company stated that R-CHOP and R-CVP are considered more effective than R-mono.[26] One would expect the model to confirm this. However, in the additional analysis that the company provided upon request of the ERG, the ICER of R^2 versus R-mono was substantially higher at £22,580 vs £11,471 for R^2 versus R-CHOP. This was predominantly caused by the fact that LYs and QALYs for R^2 were lower, while costs were higher. So, when using data from the direct comparison as per AUGMENT, R^2 was more costly and less effective than when using results from the MAIC. This again raises the question whether the indirect comparison provided valid results, as the MAIC seems to inflate efficacy and lower costs for R^2. The model does not accommodate quantification of this uncertainty and so the ERG cannot provide an estimate of its potential impact.

c) The company assumed treatment waning to start at five years, based on previous STAs. Upon the ERG's request in the clarification phase to further justify this choice of timepoint, the company replied that neither TA472[20] or TA137[59] appeared to present evidence to support their assumptions, even though treatment effect was a key uncertainty in these appraisals, having a large impact on the ICERs. The company further argues that five years is considered conservative as the immunomodulatory effect of lenalidomide could promote a longer treatment effect versus R-chemo's. The company also argued that choice of time point did not have a huge impact on the results when tested at three or 10 years. The ERG considers the company's choice of time point to be rather arbitrary and a shorter or longer duration of treatment effectiveness may be equally likely. As in the company base-case, the ERG ran scenarios varying the time point to three and seven years.

d) The company proposed a systematic way of selecting the parametric curves for extrapolating, consisting of four steps i.e. 1) proportional hazards assumption based on log cumulative hazard plots 2) visual inspection, AIC and BIC 3) clinical plausibility, and 4) implausible curve crossings (e.g. OS moving below TTNLT) before 15 years of follow-up. In the actual selection, however, it is difficult to see how these criteria were handled. For OS, PFS, and TTNLT, the CS states that 'all curves fit the data reasonably well'. Avoiding implausible curve crossing seemed to be the main argument for selection.

e) For OS, the company's argument for choosing the Weibull distribution over the better fitting exponential distribution was that the AIC/BIC for Weibull suggested a reasonable fit, and Weibull was also used in TA137[59] on R-mono. Given the company's claim that R^2 is essentially different from R-mono, the ERG is not convinced that OS in R^2 would be logically comparable to OS in TA137. Given the criteria that the company stated to have taken into account for selecting the curves, it is not clear to the ERG how the Weibull could be preferred over the exponential distribution. The ERG base-case used the exponential distribution for both arms. For the R-mono comparison the same argument applied and so the ERG base-case also incorporated the exponential distribution for OS.

f) The company interpreted the slight divergence and subsequent convergence/overlap of the KM curves for R^2 and R-CHOP/R-CVP as a non-constant treatment effect. They then decided to model PFS for R^2 using the KM data until the maximum follow-up after which the comparator hazard was applied to extrapolate further. In this way, the CS stated, the relative treatment effect of R^2 vs. R-CHOP/R-CVP based on the MAIC was accurately reflected. The ERG fails to see why and how this way of modelling PFS would accurately reflect the, as the company stated "non-constant", relative treatment effect. The overlap of the KM curves may as well be indicative of the absence of a treatment effect. At the end of follow-up, the KM for R^2 is higher than any of the estimated survival curves. See Figure 5.2 with the parametric PFS curves alongside the KM data + comparator hazard that was actually used in the base-case. The ERG considers this approach to favour R^2 even though the parametric PFS curves for R-CHOP/R-CVP are mostly higher than those for R^2, in particular from the five-year point onwards. Also, choosing the point of last follow-up as a starting point to extrapolate further is quite arbitrary. At any other timepoint, the start of the extrapolation could have been substantially different. Furthermore, near the tail of the KM-curve the number of patients approaches zero (exact numbers difficult to see from the CS) which would increase the uncertainty of the extrapolation that follows.

   For R-CHOP/R-CVP, the generalised gamma that was used in the company base-case appears to underestimate PFS in the first year (Figure 5.3) and remains lower than most of the other non-parametric curves. Given that generalised gamma does not provide the best statistical fit, the ERG considers this a sub-optimal choice. As the company advocated separate model types for the two treatment arms, which seems reasonable, the ERG base-case includes the log-logistic curve for R^2 (as hazard appears to be non-constant from the log-cumulative hazard plot) and Weibull for R-CHOP/R-CVP, going by AIC and BIC as the main criteria.

   For the R-mono comparison, the selected log-logistic curve did not seem to fit very well to the R-mono arm. No justification was provided for choosing log-logistic. In a scenario, the ERG explored the use of the generalised gamma, which fitted R-mono better (but was a worse fit to R^2).

g) For TTNLT, the main reason to select the log-normal curve was because the exponential curve would cause crossing of TTNLT and PFS arms. However, as the log-normal distribution did not fit the R^2 data very well, and the crossing of curves would not actually take place but was corrected for, the ERG considered the exponential curve equally suitable. The choice of TTNLT curve was however not too influential as there are no consequences for OS and utility scores in the company base-case were high throughout, making TTNLT mostly about the timing of the one-off subsequent treatment costs and a slight utility decrement. For the R-mono comparison, in the absence of any diagnostic plots, it was difficult to see whether the generalised gamma would be the optimal choice, but the ERG considered that given AIC and BIC for the various parametric survival curves, there

may not be one model type that would make a good fit to both arms and so the company base-case was left unchanged at this point.

h) When running the scenario where efficacy data from van Oers were used, the model returned error values. This was caused by a dot versus comma issue (possibly specific to the version of Excel used to run the model in) in the parameters for the parametric survival curves. In the ERG base-case, this error was fixed.

Figure 5.2: PFS curves for R^2 with in addition the KM + comparator hazard curve that was used in base-case

Source: adapted from company model[26] – KM + comparator curve added

Figure 5.3: PFS curves for R-CHOP/R-CVP (GenGamma used in company base-case, log-logistic in ERG base-case)

Source: adapted from company model[1] – changed line presentation styles

5.2.7 Adverse events

The main sources of evidence on treatment-related adverse events used for intervention and comparators are the AUGMENT[64] and RELEVANCE[66] trials, because of a lack of safety data from HMRN. RELEVANCE is a Phase III study comparing R^2 with R-chemotherapy (R-CHOP, R-CVP and R-Benda) for patients with previously untreated FL. AUGMENT was used in the base-case analysis for R^2, and RELEVANCE was used for R-CHOP and R-CVP with incidence adjusted for relative incidence of R^2 in AUGMENT compared to R^2 in RELEVANCE:

Comparator AE incidence = ($AE_{comparator}$ incidence in RELEVANCE/AE_{R2} incidence in RELEVANCE) x AE_{R2} incidence in AUGMENT.

So, the incidence of R-CHOP and R-CVP AEs were adjusted for any possible differences in R^2 AEs between AUGMENT and RELEVANCE.

Grade 3/4 AEs with incidence of greater than 2% in either treatment were considered. If any reported AEs for R-CHOP/R-CVP were >2% incidence, they were also included for R^2. Any AEs reported in AUGMENT that were used in the model, but were not reported for the comparator, were assumed 0% incidence for the comparator and not costed for.

In a scenario, AEs for the comparators were taken from van Oers et al. (2006)[39] which concerned a relapsed/refractory population. As van Oers et al. was a study on R-CHOP and no data on AEs in R-CVP were available, in this scenario the R-CHOP AE incidences were also applied to the R-CVP comparator.

Furthermore, AE incidence for maintenance treatment and autologous stem cell transplant (ASCT) were also considered. The incidence of AEs for rituximab maintenance were taken from van Oers et al. (2010)[67] and were neutropenia (11.5%) and infection (19.7%). In line with NG52 NHL guidelines,[68] the

only post ASCT AE for the model was febrile neutropenia, for 98.3% of patients undergoing ASCT as taken from Leger et al.[69]

ERG comment: The main concerns of the ERG relate to: a) the omission of AEs related to ASCT and subsequent R-mono therapy in the R^2 arm b) the RELEVANCE population being exclusively patients that were previously untreated.

a) For patients in the R-CHOP/R-CVP arm undergoing ASCT and R-mono as subsequent therapy, AEs related to these treatments were accounted for and costed in the model. Also, a small utility decrement was applied for these AEs. For the R^2 arm, even though these therapies were also observed here (be it to a lesser extent), related AEs were not accounted for. The ERG considered this to be inconsistent and corrected for it in the ERG base-case.

b) AE incidences in the company base-case were taken from the RELEVANCE trial[66] which concerned a population of previously untreated patients. Data from a relapsed/refractory population were only used in a scenario, with lower AE incidences than in the base-case. The ERG questions the applicability of RELEVANCE for the present STA. On the one hand, it may be the case that previously untreated patients have fewer side effects than a relapsed/refractory population, since they have not built up any intolerances. On the other hand, one would expect that a population receiving second-line treatment might be a special selection in the sense that those patients who experienced severe AEs in first-line will not be eligible for second-line. Either way, the ERG feels it is important to seriously consider the scenario provided by the company. Therefore, the ERG included it as one of their scenarios.

5.2.8 Health-related quality of life

The utility values were estimated for the health states PF, and PP off and on treatment using EQ-5D-3L data collected in AUGMENT. A covariate selection process was used to select the appropriate mixed effects utility model as input for the economic model. The final covariates included in the model were health state (PF versus PP), next anti-lymphoma treatment, treatment, baseline utility, previous rituximab exposure, refractory to last prior regimen and number of prior therapies. The R^2 arm had a utility increment of 0.011 compared with the R-mono arm for all health states. However, given that this difference was minimal and not statistically significant, the company used the same utility values based on R^2 in the model.

Health-related quality of life data identified in the review

According to the CS, the SLR identified a total of 38 studies from 53 publications, including 12 HTAs and one observational study. Out of these, the company considered the utility values of the studies of Wild et al., Pereira et al. and TA472 most relevant.[20, 70, 71]

Health state utility values

The utility values resulting from the mixed effects model were used to inform the health states in the model for all treatments, and utility values from the study of Wild et al.[70] were tested in a scenario analysis. However, the disease characteristics that were used to derive utility values from the mixed effects model were population-dependent, and therefore, the utility values for R^2 versus R-CHOP/R-CVP and R^2 versus R-mono are slightly different. A summary of all utility values used in the model is provided in Table 5.4. The company stated in the CS that the mean utility value for the PF state was generally consistent with those reported in the three studies selected from the SLR, with the exception of the lower PF utility value of Pereira et al.[71] The mean utility values for post-progression were higher based on AUGMENT trial data compared with the other studies.

Table 5.4: Health state utility values

State	Utility value R^2 versus R-CHOP/CVP	Utility value R^2 versus R-mono	Reference	Justification
PF	0.863	0.847	Section B.3.4, page 177 and 182 of the CS.	EQ-5D values derived from a relevant patient population and model specific health states.
PP (off treatment)	0.837	0.821		
PP (on treatment)	0.808	0.792		
Source: based on Table 47 of the CS				

Adverse event related disutility values

Utility decrements for grade 3 and 4 AEs were applied in the model for the expected duration of each AE, based on literature and previous appraisals. See Table 5.5 for details on the AE utility decrements, durations and sources.

Table 5.5: Adverse event related disutility values

Adverse event	Disutility value	Duration (days)	Source for disutility	Source for duration
Neutropenia	0.090	15.09	Nafees et al. (2008)[72]	TA306[73]
Leukopenia	0.119	13.96	TA513 (assumed to be the same as anaemia)[23]	TA306[73]
Anaemia	0.119	16.07	Swinburn et al. (2010)[74]	TA306[73]
Pneumonia	0.200	14.00	Beusterien et al. (2010)[75]	TA306[73]
Lymphocyte count decreased	0.100	34.00	Stein et al. (2018)[76]	Assumed maximum of all Grade 3/4 AEs
Lymphopenia	0.100	34.00	Stein et al. (2018)[76]	TA306[73]
Febrile neutropenia	0.150	7.14	Lloyd et al. (2006)	TA306[73]
White blood cell count decreased	0.100	34.00	Stein et al. (2018)[76]	Assumed maximum of all Grade 3/4 AEs
Diarrhoea	0.048	34.00	Nafees et al. (2008)[72]	Assumed maximum of all Grade 3/4 AEs
Thrombocytopenia	0.108	23.23	Tolley et al. (2013)[77]	TA306[73]
Hypokalaemia	0.124	34.00	TA423[78]	Assumed maximum of all Grade 3/4 AEs

Adverse event	Disutility value	Duration (days)	Source for disutility	Source for duration
Pulmonary embolism	0.124	34.00	TA423[78]	Assumed maximum of all Grade 3/4 AEs
Infusion-related reaction	0.195	34.00	Tolley et al. (2013)[77]	Assumed maximum of all Grade 3/4 AEs
Nausea and emesis	0.048	6.00	Nafees et al. (2008)[72]	TA306[73]
Allergic reaction	0.098	34.00	Hannouf et al. (2012)[79]	Assumed maximum of all Grade 3/4 AEs
Hypotension	0.057	8.00	Hannouf et al. (2012)[79]	TA306[73]
Fatigue	0.073	31.50	Nafees et al. (2008)[72]	TA306[73]
Alopecia	0.045	34.00	Nafees et al. (2008)[72]	Assumed maximum of all Grade 3/4 AEs
Infection	0.195	34.00	Tolley et al. (2013)[77]	Assumed maximum of all Grade 3/4 AEs
Sepsis	0.267	34.00	Hannouf et al. (2012)[79]	Assumed maximum of all grade ¾ AEs
Abdominal pain	0.069	17.00	Doyle et al. (2008)[80]	TA306[73]
Acute kidney injury	0.270	29.75	TA306[73]	TA306[73]
Source: Based on Table 44 of the CS.				

ERG comment: The main concerns of the ERG relate to: a) the high utility values for the PF and PP off treatment and the PP on treatment health states; and b) the modest utility decrement for progressed disease;

a) Utility values for the PF (0.863 for R^2 versus R-CHOP/R-CVP and 0.847 for R^2 versus R-mono) and PP (off treatment 0.837 for R^2 versus R-CHOP/R-CVP and 0.821 for R^2 versus R-mono, on treatment 0.808 for R^2 versus R-CHOP/R-CVP and 0.792 for R^2 versus R-mono) health states were higher than the utility reported for the general population (0.80 for age category 55-64).[81] Utility scores higher than in the general population seem quite unlikely in patients with treated FL or MZL. In addition, these utility values were also higher than reported in the literature for this population.[70, 71] Also, the company decided to go with the slightly higher utilities from the R^2 arm in AUGMENT, even though there was not a significant difference between R^2 and R-mono. The ERG capped the utility values in its base-case to general population norms, which had a low impact but slightly increased the ICER.

b) The utility difference between the PF health state and the PP off treatment and PP on treatment health states were -0.026 and -0.056 respectively in the R^2 versus R-CHOP/R-CVP comparison and

respectively -0.026 and -0.055 in the R^2 versus R-mono comparison. This seems modest given the difference in utility value between these health states reported in the literature, which show differences up to -0.27.[71] The ERG judges that a larger utility difference between PF and PP health states would be more plausible, and explored this in a scenario analysis using utility values of Wild et al. (0.62) and Pereira et al. (0.45) for both PP health states. For R^2 versus R-CHOP and R/CVP, this substantially increased the ICER, while for R^2 versus R-mono the ICER decreased.

5.2.9 Resources and costs

The cost categories included in the model were costs associated with treatment (drug acquisition costs including subsequent therapies, drug administration costs including subsequent therapies, costs associated with treatment-related AEs), disease monitoring costs and costs associated with end of life care.

Unit prices were based on the National Health Service (NHS) reference prices,[82] Personal Social Services Research Unit (PSSRU),[83] the Monthly Index of Medical Specialities (MIMS)[84-90] and the Electronic Market Information Tool (eMIT)[91].

Resource use and costs data identified in the review

According to Appendix I of the CS[1], the SLR identified 17 studies of which 14 reported UK relevant resource use and cost information. The CS did not state which of these studies the company considered to be consistent with the NICE reference case and appropriate for the economic model.

Drug acquisition costs (with PAS)

For lenalidomide, dosing data had been taken directly from AUGMENT (non-R-refractory population) to align the drug costs with the efficacy data because according to the company, dose reductions for lenalidomide can occur. To capture the impact of treatment reductions or missed treatment cycles over time on costs, the observed number of patients on each dosage at every cycle was combined with the unit drug costs to calculate a weighted cost per cycle. This cost was then multiplied by the proportion of patients eligible for treatment who receive treatment in that cycle (based on ToT KM curves and the mean treatment cycle length). To align with the costing method applied for lenalidomide, the same method was applied to calculate rituximab costs for the R^2 arm. The use of mean relative dose intensities (RDIs) were explored in scenario analyses, using values of [redacted] and [redacted] for rituximab and lenalidomide in the R^2 arm of AUGMENT, respectively (Table 5.6).

The proportion of patients eligible for treatment who receive treatment in each arm in the rituximab monotherapy arm of AUGMENT was applied to all comparators in the model, in order to similarly align the costing of the comparators to the study dosing methods described above for R^2. A mean dose intensity value of 87.5% was assumed in scenario analyses across all individual chemotherapies within the R-chemotherapy comparator regimens. No dose intensity value was applied to rituximab within R-chemotherapy combinations or R-maintenance, because dose reductions were not recommended for rituximab. For BSA dependent treatments, the company applied the method of moments technique to IPD from AUGMENT to calculate the average number of vials that would be required to satisfy one administration of treatment. Other methods, such as dose banding and using the minimum cost per mg for each treatment (no wastage), were explored in scenario analyses.

Patients in the R^2 received allopurinol (in the first treatment cycle only) and filgrastim as concomitant treatments. Rituximab maintenance was given every three months up to two years or until disease progression to patients who responded to R-chemotherapy induction treatment.

Table 5.6: Treatment acquisition costs

Treatment	Size	Cost per pack	Source
Lenalidomide (with PAS)	21 x 2.5 mg tablets	[redacted]	MIMS (Revlimid)[85]
	21 x 5 mg tablets	[redacted]	
	21 x 10 mg tablets	[redacted]	
	21 x 15 mg tablets	[redacted]	
	21 x 20 mg tablets	[redacted]	
Rituximab	2 x 100 mg vials	£349.25	MIMS (MabThera)[84]
	1 x 500 mg vial	£873.15	
	1 x 1,400 mg (SC)	£1,344.65	
	2 x 100 mg vials	£314.33	MIMS (Rixathon)[89]
	1 x 500 mg vial	£785.84	
Cyclophosphamide	1 x 1,000 mg vial	£13.47	eMIT[91]
	1 x 2,000 mg vial	£27.50	
	1 x 500 mg vial	£8.31	
Doxorubicin	1 x 10 mg vial	£4.48	
	1 x 200 mg vial	£15.59	
	1 x 50 mg vial	£17.78	
Vincristine	5 x 1 mg vials	£11.59	
	5 x 2 mg vials	£17.82	
	5 x 5 mg vials	£99.00	
Prednisolone	28 x 1 mg tablets	£0.17	
	28 x 2.5 mg tablets	£0.59	
	30 x 20 mg tablets	£4.17	
	56 x 25 mg tablets	£20.25	
	28 x 5 mg tablets	£0.27	

Source: based on Table 49 of the CS.
eMIT = electronic market information tool; MIMS = Monthly Index of Medical Specialities; PAS = patient access scheme; SC = subcutaneous

Administration costs

Drug administration costs were based on NHS reference costs tariffs, pharmacy costs for the preparation of the infusion, and NHS transport costs.[82] For rituximab combination chemotherapies, a cost of £374.52 was applied at first administration of each cycle, followed by a cost of £312.34 for subsequent administrations per cycle. For simpler chemotherapies such as in the R^2 arm, first administration per cycle cost £309.22 and £312.34 for subsequent administrations per cycle. For all infusion treatments, pharmacy costs were applied assuming a 15-minute infusion preparation time based on TA243[21] and £48 per hour for hospital-based scientific and professional staff from PSSRU costs.[83] NHS transport costs were assumed in 30% of patients and were applied to all administrations in the model.

Treatment-specific monitoring

Costs of a full blood count were added to each treatment cycle for lenalidomide per visit to monitor the dose-limiting toxicities of neutropenia and thrombocytopenia.

Health state costs

Table 5.7 presents the costs that are included in the economic model per health state.

Table 5.7: Health state related costs

Health state	Costs	Cost components considered	Reference
PF (on treatment)	Drug acquisition	**R²:** Cycle 1: £[redacted] Cycles 2-5: £[redacted] Cycles 6-12: £[redacted] **R-CHOP:** £1,216 per cycle **R-CVP:** £1,200 per cycle **R-mono:** Cycle 1: £4,680 Cycles 2-5: £1,170	Table 48, page 187 of the CS
	Drug administration	**R²:** Cycle 1: £1,348 Cycles 2-5: £335 **R-CHOP/R-CVP:** £400 per cycle **R-mono:** Cycle 1: £1,348 Cycles 2-5: £335	Table 48, page 187 of the CS
	Maintenance/ASCT	R-maintenance: £1,345 (SC), £1,170 (IV)	Table 48, page 187 of the CS
		ASCT: £35,558	Table 59, page 202 of the CS
	Disease monitoring	£254,95 per month	Table 57, page 200
	Adverse events	£1,832 (**R² non-R-refractory**) £3,604 (**R-CHOP**) £2,754 (**R-CVP**) £462 (**R-mono**) £1,773 (R² R-refractory) £370 (R-maintenance) £6,336 (ASCT)	Table 61, page 204 of the CS
PF (off treatment)	Disease monitoring	£83.09 per month	Table 57, page 200 of the CS

Health state	Costs	Cost components considered	Reference
PP (on treatment)	Disease monitoring	£232.17 per month	Table 57, page 200 of the CS
	Subsequent treatments	£5,195 (**R^2**) £8,371 (**R-CHOP/R-CVP**)	Table 62, page 206 of the CS
PP (off treatment)	Disease monitoring	£58.04 per month	Table 57, page 200 of the CS
Death	Terminal care	£6,362	Page 206 of the CS

Source: Based on Table 56 of the CS
CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP = cyclo-phosphamide, vincristine and prednisolone; R = rituximab; R^2 = lenalidomide plus rituximab.

Disease monitoring

Disease monitoring resource use costs were assumed to be similar to those in previous FL submissions[20, 21, 23] and were split by health state. Patients were assumed to have monthly haematologist visits and diagnostic tests with a CT scan every six months in the induction phase of the PF health state. In the maintenance phase of the PF health state, follow-up visits (based on ESMO guidelines) were reduced to one visit every three months and one annual CT scan, and in the post-maintenance phase to one visit every four months without CT scan. In the PP state, higher visit and diagnostic test frequencies were assumed (both monthly). Although resource use information for MZL was limited, similar tests and frequencies were suggested in the MZL ESMO guidelines[29] and therefore disease monitoring costs were assumed to be identical to FL.

Stem cell transplant (pre-progression)

Patients that were fit and young enough and who relapse early but who are not refractory to induction therapy were considered for consolidation with ASCT. In the economic model, the company applied ASCT to [redacted]% of patients in R-CHOP. For R-CVP, 0% ASCT was applied as R-CVP was considered unlikely to be used as an induction regimen prior to ASCT. For R^2, as clinicians suggested that it was unlikely that ASCT would be offered post R^2 in clinical practice and ASCT was also not offered to patients after R^2 within the AUGMENT protocol, 0% ASCT was applied as well. The cost of ASCT was based on NHL guidance uplifted to 2018 costs and included £35,558.15.[25, 83] The NHS reference cost (£18,520.20) for ASCT was used in a scenario analysis.

Adverse event related costs and costs of terminal care

The frequency of grade 3-4 AEs that occurred in ≥2% of patients was applied to the incidence rate for each treatment to obtain a one-off upfront cost to each treatment arm in the model.

Furthermore, to reflect the costs of terminal care, a one-off cost of £6,361.77 was applied in the model when a patient died. This cost was based on the average cost derived from the Round et al. (2015) modelling study,[92] which estimated the cost of cancer care during the final phases of life.

Total AE costs per treatment are shown in Table 5.8.

Table 5.8: Total AE costs per treatment

Treatment	Total costs
Non-R-refractory	
R^2	£1,831.71
R-CHOP	£3,604.13
R-CVP	£2,753.56
R-mono	£462,41
R-refractory	
R^2	£1,773.94
Post-induction	
R-maintenance	£369.95
ASCT	£6,400.93

Source: based on Table 61 of the CS.
AE = adverse event; ASCT = autologous stem-cell transplant; CHOP = cyclophosphamide, doxorubicin hydro-chloride, vincristine and prednisolone; CVP = cyclophosphamide, vincristine and prednisolone; R = rituximab; R^2 = lenalidomide plus rituximab.

Costs of subsequent treatments

Subsequent treatments were included in the model as an average one-off cost to patients entering the PP (on treatment) health state, derived using TTNLT data. Costs for patients in the R^2 arm were derived from subsequent treatments from AUGMENT. The total subsequent treatment data from the pooled R-chemotherapies in the HMRN database were used for R-CHOP and R-CVP. The company also conducted a scenario analysis in which the costs were equalised by applying the subsequent treatment costs of the comparator arm to R^2. The mean duration of subsequent treatments was based on HMRN data, with AUGMENT mean durations used in a scenario analysis.

ERG comment: The main concerns of the ERG relate to: a) subsequent treatments that were included as a one-off cost and were therefore potentially underestimated; b) the source used for the proportion of patients who receive subsequent treatment after R-CHOP/R-CVP to determine subsequent treatment costs; and c) the omission of data observed in AUGMENT to inform pre-progression ASCT in the R^2 arm.

a) Subsequent treatments were included in the model as a one-off cost to those patients entering the PP on treatment health state. The company costed for observed incidences of subsequent treatments from the data sources, which for R2 had a much shorter follow-up than for R-CHOP/R-CVP. The ERG is concerned that because of the limited follow-up in AUGMENT as compared to HMRN, this assumption does not reflect clinical practice and subsequent treatment costs for R^2 in the economic model are therefore likely to be underestimated. Although subsequent treatment duration in the model lasts no longer than a maximum of 130.3 days, patients in the PP on treatment health state remain in this health state until they die, and the relatively high age-adjusted utility value corresponding to this health state is assumed over this whole time span. The ERG is concerned about the fact that subsequent treatment costs, in contrast to the utilities, are not counted over the remaining time that patients stay in the PP on treatment health state. As patients in the R^2 arm

remain in the PP on treatment health state for a longer time on average, applying subsequent treatment costs as one-off possibly favoured R^2.

b) To calculate subsequent treatment costs, the company based the proportion of patients receiving subsequent treatment after R-CHOP/R-CVP on the total subsequent treatment data from the pooled R-chemotherapies from HMRN because of its larger sample size (n=129) compared to the HMRN R-CHOP/R-CVP cohort (n=67). However, the ERG judges that, in line with the treatment effectiveness, the R-CHOP/R-CVP cohort should be used to calculate subsequent treatment costs and applied this to the ERG base-case for the comparison with R-CHOP and R-CVP. This resulted in slightly lower subsequent treatment costs for R-CHOP and R-CVP and a slightly higher ICER, although the impact was small. In addition, the ERG also explored the impact of equal subsequent treatment costs between R^2 and R-CHOP, R-CVP and R-mono, which resulted in a large increase of the ICER.

c) The company assumed the percentage of post-induction (but pre-progression) ASCTs in R^2 to be zero, because it was not protocolised in AUGMENT and clinicians considered it unlikely that patients would receive ASCT post R^2. The ERG was unable to find any report of actual incidence of ASCT performed post R^2 in AUGMENT, but would have liked to see a scenario using observed frequencies, as clinical practice may sometimes contrast with protocols and clinical opinion. If the observed frequency was non-zero, this would increase the ICER for R^2 compared to R-CHOP.

5.2.10 Cost effectiveness results

R^2 versus R-CHOP and R-CVP

In the deterministic base-case analysis, total LYs and QALYs gained were larger for R^2 than for R-CHOP and R-CVP. Incremental QALYs ([redacted]) were mainly driven by QALY gains in the PP (off treatment) health state. Total costs were also higher for R^2 than for R-CHOP and R-CVP. Incremental costs ([redacted] and [redacted] respectively) mainly resulted from higher drug acquisition (induction) costs. The deterministic incremental cost effectiveness ratio (ICER) amounted to £11,471 per QALY gained for R^2 versus R-CHOP and £16,814 for QALY gained for R^2 versus R-CVP (see Table 5.9).

R^2 versus R-mono (added by the company after the clarification phase)

In the deterministic base-case analysis, total LYs and QALYs gained were larger for R^2 than for R-mono. Incremental QALYs ([redacted]) were mainly driven by QALY gains in the PF health state. Total costs were also higher for R^2 than for R-mono. Incremental costs ([redacted]) mainly resulted from higher drug acquisition (induction) costs. The deterministic cost effectiveness ratio (ICER) amounted to £22,580 per QALY gained (see Table 5.9).

Table 5.9: Company's deterministic base-case results

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
R^2 versus R-CHOP							
R^2	£[redacted]	[redacted]	[redacted]	--	--	--	--
R-CHOP	£[redacted]	[redacted]	[redacted]	£[redacted]	[redacted]	[redacted]	£11,471
R^2 versus R-CVP							
R^2	£[redacted]	[redacted]	[redacted]				
R-CVP	£[redacted]	[redacted]	[redacted]	£[redacted]	[redacted]	[redacted]	£16,814

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
R^2 versus R-mono							
R^2	[redacted]	[redacted]	[redacted]	--	--	--	--
R-mono	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£22,580
Source: Based on Table 64 of the CS ICER = incremental cost effectiveness ratio; LYG = life years gained; QALY = quality adjusted life year.							

5.2.11 Sensitivity analyses

The company performed a probabilistic sensitivity analysis (PSA) and deterministic sensitivity analysis (DSA) to show the uncertainty surrounding the base-case results.

R^2 versus R-CHOP and R-CVP

Compared with the deterministic results, the PSA with 1,000 iterations showed lower incremental QALYs and costs for both R-CHOP and R-CVP, which resulted in increased ICERs of £13,443 (versus R-CHOP) and £20,896 (versus R-CVP) (see Table 5.10). The cost effectiveness acceptability curve in the economic model showed that R^2 had an 82% (versus R-CHOP) and 72% (versus R-CVP) probability of being cost effective at a willingness-to-pay (WTP) threshold of £30,000.

The company performed DSAs by varying key model parameters between their upper and lower limits of the confidence intervals. For R^2 versus R-CHOP, the ICER was most sensitive to the cost of ASCT, the total subsequent treatment costs for R-CHOP and the proportion of patients who receive SCT. For R^2 versus R-CVP, the ICER was most sensitive to the total subsequent treatment costs for R-CVP (including ASCT costs) and resource use costs. For both comparisons, in none of the DSAs the ICER exceeded the WTP threshold of £30,000.

R^2 versus R-mono (added by the company after the clarification phase)

For R^2 versus R-mono, the company only provided basic deterministic results and the PSA and DSA were performed by the ERG. Compared with the deterministic results, the PSA with 1,000 iterations showed lower incremental QALYs and costs, which resulted in an increased ICER of £26,116) (see Table 5.10). The cost effectiveness acceptability curve in the economic model showed that R^2 had a 69% probability of being cost effective at a willingness to pay (WTP) threshold of £30,000.

The company performed DSAs by varying key model parameters between their upper and lower limits of the confidence intervals. The ICER was most sensitive to the total subsequent treatment costs for R^2 and R-mono and the frequency of haematologist visits post progression. In none of the DSAs the ICER exceeded the WTP threshold of £30,000.

Table 5.10: Company's base-case results (probabilistic, 1,000 iterations)

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
R^2 versus R-CHOP					
R^2	£[redacted]	[redacted]	--	--	--
R-CHOP	£[redacted]	[redacted]	£[redacted]	[redacted]	£13,443
R^2 versus R-CVP					
R^2	£[redacted]	[redacted]			
R-CVP	£[redacted]	[redacted]	£[redacted]	[redacted]	£20,896

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
R^2 versus R-mono					
R^2	£[redacted]	[redacted]	--	--	--
R-mono	£[redacted]	[redacted]	£[redacted]	[redacted]	£26,116
Source: Based on Table 64 of the CS. ICER = incremental cost effectiveness ratio; QALY = quality adjusted life year.					

Scenario analyses

The company conducted several scenario analyses. The results for R^2 versus R-CHOP showed ICERs ranging between £4,398 and £14,891 per QALY gained, excluding the scenarios assessing different time horizons. The three most influential scenarios that decreased the ICER were using an exponential distribution for R^2 ToT (£4,398), a 0.0% discount rate for QALYs (£8,174) and using a log-normal distribution for R^2 ToT (£8,312).

The results for R^2 versus R-CVP showed ICERs ranging between £9,731 and £20,636 per QALY gained, excluding the scenarios assessing different time horizons. The three most influential scenarios that increased the ICER were a 6.0% discount rate for QALYs (20,636), applying the comparator hazard to R^2 arms after three years (£20,471) and using a Gompertz distribution for R^2 PFS (£20,413). The three most influential scenarios that decreased the ICER were using exponential (£9,731) and log-normal (£13,650) distributions for R^2 ToT and a 0.0% discount rate for QALYs (11,976). For R^2 versus R-mono, the ICERs ranged between £12,125 and £43,814, excluding the scenarios assessing different time horizons. The three most influential scenarios that increased the ICER were applying the comparator hazard to R^2 arms after three years (£43,814), using a 6.0% discount rate for QALYs (£27,613) and applying the same subsequent treatment costs for R^2 and R-mono (£24,951). The three most influential scenarios that decreased the ICER were applying the comparator hazard to R^2 arms after 10 years (£12,125), using an exponential distribution for R^2 ToT (£13,845) and using a 0.0% discount rate for QALYs (£16,391).

ERG comment: The main comments of the ERG relate to: a) the inability to perform a fully incremental analysis in the model; b) unstable PSA results; and c) the additional scenario analysis for the FL-only population.

a) The PSA did not enable simultaneous calculation of outcomes for more than two comparators and representation of multiple comparators in the cost effectiveness acceptability curve (CEAC). Therefore, the ERG created three separate model files. Furthermore, compared with the company's deterministic base-case results, probabilistic incremental QALYs are lower, likely caused by non-linearity of the model.

b) The ERG twice performed a PSA with 10,000 iterations to test its stability, but increasing the number of iterations did not stabilise the results.

c) An additional scenario analysis for the FL-only population was provided by the company in response to clarification. The FL-only scenario resulted in increased deterministic ICERs of £15,909 and £23,746 for the R-CHOP and R-CVP comparisons, respectively, making it the most influential scenario. For the R-mono comparison, using FL-only data lowered the ICER to £20,310.

Therefore, given that the pooling of FL and MZL population appeared to have a substantial impact, the ERG included the FL-only scenario in their exploratory analyses.

5.2.12 Model validation and face validity check

Face validity

The model structure and its appropriateness to reflect the clinical pathway, notably the decision to split the progressed disease health state up into on- and off-treatment, were validated in an advisory board consisting of six clinicians and two UK economic experts. These further validated the extrapolation of survival beyond the trial period, the indirect treatment comparison, the use of clinical validity of utilities derived from AUGMENT versus those in the literature and subsequent treatment usage.

Internal validity

Distributions to estimate PFS, OS and TTNLT were chosen such that no implausible curve crossing occurred. A health economist that was not involved in model development reviewed the model for coding errors, inconsistencies and input plausibility. Several extreme value checks were also performed and sub-modules of the model were tested.

Cross validity

No cross validity checking of the model was reported by the company, although the company did state that the chosen modelling approach of partitioned survival analysis with the health states of PF, PP and dead was in line with "the majority of economic evaluations found in the SLR".[26] However, the company then diverted from this path by adding additional health states (splitting progression by whether patients were on or off treatment given that TTNLT was considered a better endpoint than PFS).

External validity

Model predictions for PFS, OS, TTNLT were compared with the respective KM data from AUGMENT and found mostly in line, with the notable exception of 1-year PFS for R-CHOP/R-CVP that was under-estimated in the model compared to the observed data. According to the company, from two years onwards model predictions were more aligned with observations for this outcome. Comparisons with other trials were not made because no other datasets were available.

Predictive validity

No predictive validity checking was reported by the company.

ERG comment: The main concerns of the ERG relate to: a) limited information available on the company's validation efforts based on the CS and b) concerns regarding external validity.

a) The company provided limited information on its validation efforts. In response to the clarification letter, however, the company provided the meeting report of the advisory board[26] and the filled in Assessment of the Validation Status of Health Economic Decision Models (AdViSHE) tool.[93] The latter shed more light on especially the internal validation of the company's model, which was performed to a good standard. The advisory board meeting report supported some model approaches and assumptions, but not all: for instance, the model structure including the on- and off-treatment division was not corroborated, and neither was the choice of distributions for R-CHOP/R-CVP OS and PFS.

b) External validation exercised by the company found that R-CHOP/R-CVP PFS, OS and TTNLT at one year were under-estimated. Whilst these extrapolations stabilised from two years onwards to be more aligned with the observed data, this under-estimation may still have an impact on cost effectiveness estimates, as explored in the treatment effectiveness section. Furthermore, it is not clear whether these extrapolations have been validated by experts as the expert meeting minutes only contained a statement regarding (the comparison with) R-mono.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.11 summarises the main issues highlighted by the ERG in Section 5.2 of this report, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table 5.11: Main ERG critique of company's submitted economic evaluation

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
Partitioned survival analysis, no alternative results from state transition model provided for comparison	+/-	No	Requested but not provided
Population, interventions and comparators, perspective and time horizon (sections 5.2.3-5.2.5)			
O-Benda not a relevant comparator, while R-mono left out	NA	Base-case	R-mono analysis provided by company upon request
MZL and FL populations were pooled throughout the analyses, as assumed to be similar by the company	+/-	Scenario	Scenario provided upon request
Treatment effectiveness and extrapolation (section 5.2.6)			
Indirect comparison seems to inflate R^2 efficacy and lower costs relative to R^2 in direct comparison based on AUGMENT	+	No	No
Substantial uncertainty concerning extrapolation of PFS curves. Company base-case not based on best fit, nor solid other justification	+	Base-case (MJ), scenarios	Scenarios
Curves for OS extrapolation do not provide best fit, choice is not sufficiently justified	+/-	Base-case (MJ), scenarios	Scenarios
Curves for TTNLT extrapolation do not provide best fit, choice is not sufficiently justified	+/-	Base-case (MJ), scenarios	Scenarios
Cut-off for treatment effectiveness at 5 years not supported by evidence	+/-	Scenarios	Scenarios
Adverse events (section 5.2.7)			
Incidence for adverse events in R-CHOP and R-CVP taken from published source on a previously untreated population	+/-	Scenario	Scenario
AEs (costs and utility decrements) related to subsequent treatments were omitted for R^2	+	Base-case (FV)	No

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Health-related quality of life (section 5.2.8)			
Utility scores for all health states are likely high	+	Base-case (FV)	Yes, scenarios allow for alternative values
Utility decrement post progression low	+	Scenarios	
Resources and costs (section 5.2.9)			
One-off costs for subsequent treatment likely underestimates R^2 costs	+	Scenario	Scenario using same subsequent treatment costs
Incidences of subsequent treatments for R-CHOP and R-CVP were taken from the mixed R-chemo group of HMRN, which likely is an overestimate	+	Base-case (FV)	No
Consolidation ASCT in R^2 arm assumed zero, data on observed number of ASCTs was not provided in CS	+	No	No
Cost effectiveness analyses (sections 5.2.10 and 5.2.11)			
Discrepancy between probabilistic and deterministic results	+/-	No	No
PSA does not allow for full incremental analysis	+/-	No	No

Footnotes: [a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; ICER = incremental cost effectiveness ratio; MJ = matters of judgement; NA = not applicable.

Based on all considerations in Section 5.2 of this report (summarised in Table 5.11), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016[94])

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred).

The adjustments apply to the R-CHOP and R-CVP comparisons. For the R-mono comparison, these adjustments may be different or do not apply. In the list below, when nothing is mentioned on R-mono, this implies that this particular adjustment was similarly applied to the R-mono comparison.

Fixing errors

1. Error cells when using 'van Oers' as input for R-CHOP efficacy (section 5.2.6).
   The ERG replaced dots by commas in the van Oers parameters for curves.

Fixing violations

2. AEs related to subsequent treatments not accounted for in R^2 arm (section 5.2.7).
   The ERG included costs and utility decrement for AEs related to ASCT and rituximab subsequent treatment in R^2 arm like in the comparator arm.
3. Subsequent treatment rates for R-CHOP/R-CVP taken from mixed R-chemo population (section 5.2.9). The ERG used pooled R-CVP/R-CHOP subsequent treatment rates instead of R-chemo. (Not applicable in the R-mono comparison)
4. Utilities in all health states were higher than or comparable to general population levels (section 5.2.8). The ERG capped utilities at the general population level

Matters of judgment

5. Weibull distributions for OS do not provide the best fit, reasons for selecting unclear. (section 5.2.6). The ERG used the exponential distribution to extrapolate OS in both arms
6. KM+comparator hazard approach likely overestimates PFS in R^2 (section 5.2.6).
   The ERG used log-logistic for PFS in R^2 and Weibull for PFS in the comparator (not applied to R-mono comparison)
7. Lognormal curves for extrapolating TTNLT appear to be suboptimal (section 5.2.6).
   The ERG used log-logistic for TTNLT both arms (not applied to R-mono comparison)

Table 6.1 shows the combined effect of all abovementioned adjustments simultaneously, resulting in the (deterministic) ERG base-case. The 'fixing error' adjustments were combined and the other ERG analyses were performed also incorporating these 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

5.3.1 ERG base-case results

The results of the deterministic ERG base-case are shown in Tables 5.12 and 5.13. The fully incremental analysis could only be performed for R^2, R-CHOP and R-CVP since when compared to R-mono, the effectiveness and costs of R^2 would be different. The ERG wishes to emphasise that all ERG analyses (except those for the R-mono comparison) are conditional upon the MAIC results for which uncertainty could not be quantified or incorporated in the economic model. For R-CHOP, deterministic incremental

costs were [redacted] and incremental QALYs [redacted] which resulted in an ICER of £15,505. Main drivers for the increased ICER compared to company base-case were the alternative OS and PFS curves, and to a lesser extent the use of only R-CHOP/R-CVP data for subsequent treatment rates, instead of pooled R-chemo data in the company base-case. For R-CVP, incremental costs were [redacted] and incremental QALYs were [redacted] (identical to the R-CHOP comparison), which resulted in an ICER of £21,759 which was driven by the same factors as in the R-CHOP comparison. Finally, for the R-mono deterministic comparison, incremental costs were [redacted] and incremental QALYs were [redacted] with a resulting ICER of £27,372. Main drivers were the cap of utilities to the level of the general population, and the use of alternative OS curves.

The fully incremental analysis showed R-CHOP to be strictly dominated, and the relevant comparison would be R^2 versus R-CVP. R-CHOP and R-CVP serve different populations however, and the ERG has already commented on the fact that pooling of R-CHOP and R-CVP may not be justified. If the assumption of equality does not hold, a fully incremental analysis based on a zero difference in QALYs between R-CHOP and R-CVP may not be indicated.

Compared with the deterministic base-case results, the ERG PSA with 1,000 iterations resulted in lower incremental costs but also in lower incremental QALYs, with consistently lower ICERs as a result, for all comparisons (see Table 5.14). For the R-CHOP comparison, the difference was quite modest with a probabilistic ICER of £15,818, but for R-CVP and R-mono it was more pronounced (probabilistic ICERs of £23,367 and £29,010, respectively). However, in the company base-case the differences between deterministic and probabilistic analyses were even larger, more than £4,000 in the R-CVP comparison for instance, and the cost effectiveness planes of the company base-case showed a number of extreme outliers concerning QALYs which were not observed in the ERG analyses (see Figure 5.4). This would imply the possibility that the QALY outliers in the company base-case may have been caused by the chosen distributions for extrapolating.

The cost effectiveness acceptability curves showed that compared to R-CHOP, R^2 approximately had an 83% and 90% probability of being cost effective at willingness-to-pay (WTP) thresholds of £30,000 and £50,000, respectively (See Figure 5.5). These percentages were lower for the R-CVP comparison; 68% and 84% (See Figure 5.6). For R-mono they were 54% and 77% (See Figure 5.7).

Table 5.12: ERG pairwise deterministic base-case results

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case for R^2 versus R-CHOP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,505
R-CHOP	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-CVP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,759
R-CVP	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-mono					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,372
R-mono	[redacted]	[redacted]			

ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Table 5.13: ERG fully incremental and pairwise deterministic base-case results for R2, R-CHOP and R-CVP (ICER compared to next relevant alternative)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
R-CVP	[redacted]	[redacted]			
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,759
ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year					

Table 5.14: ERG probabilistic base-case results

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case for R^2 versus R-CHOP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,818
R-CHOP	[redacted]	[redacted]			
Probabilistic ERG base-case for R^2 versus R-CVP					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,367
R-CVP	[redacted]	[redacted]			
Probabilistic ERG base-case for R^2 versus R-mono					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,010
R-mono	[redacted]	[redacted]			
ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year					

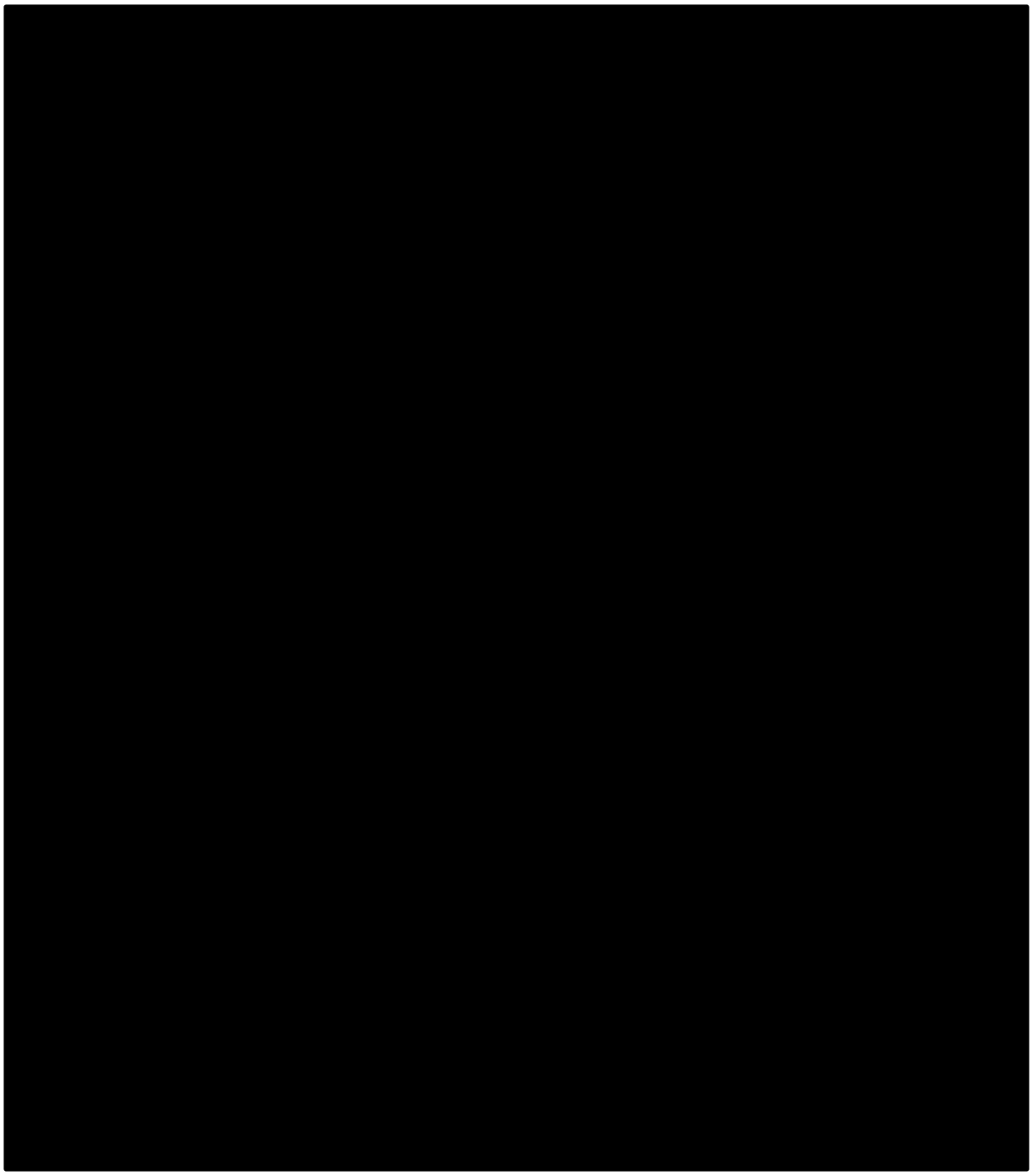


Figure 5.4: Cost effectiveness planes (1,000 iterations) for company and ERG base-case

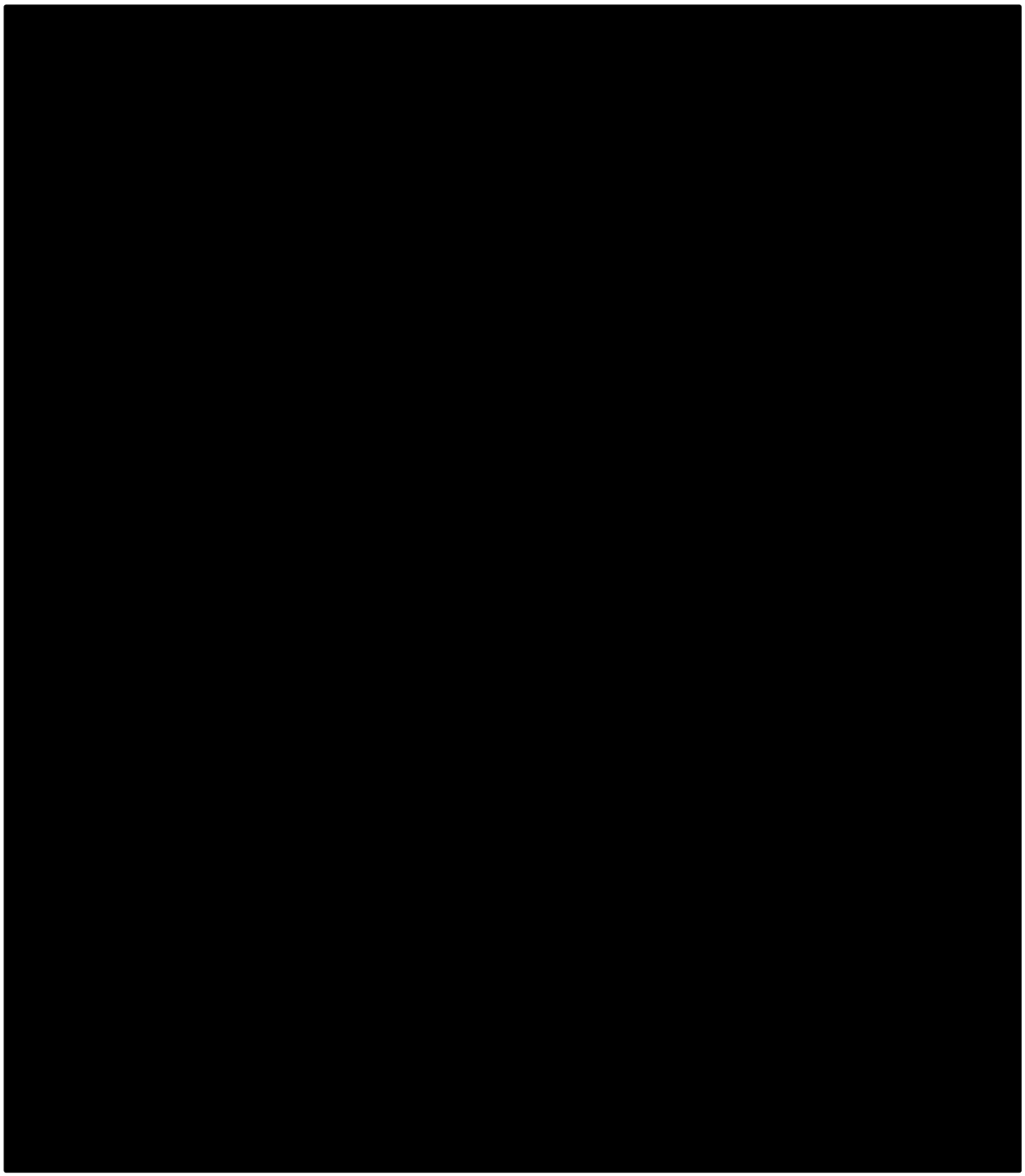


Figure 5.5: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP

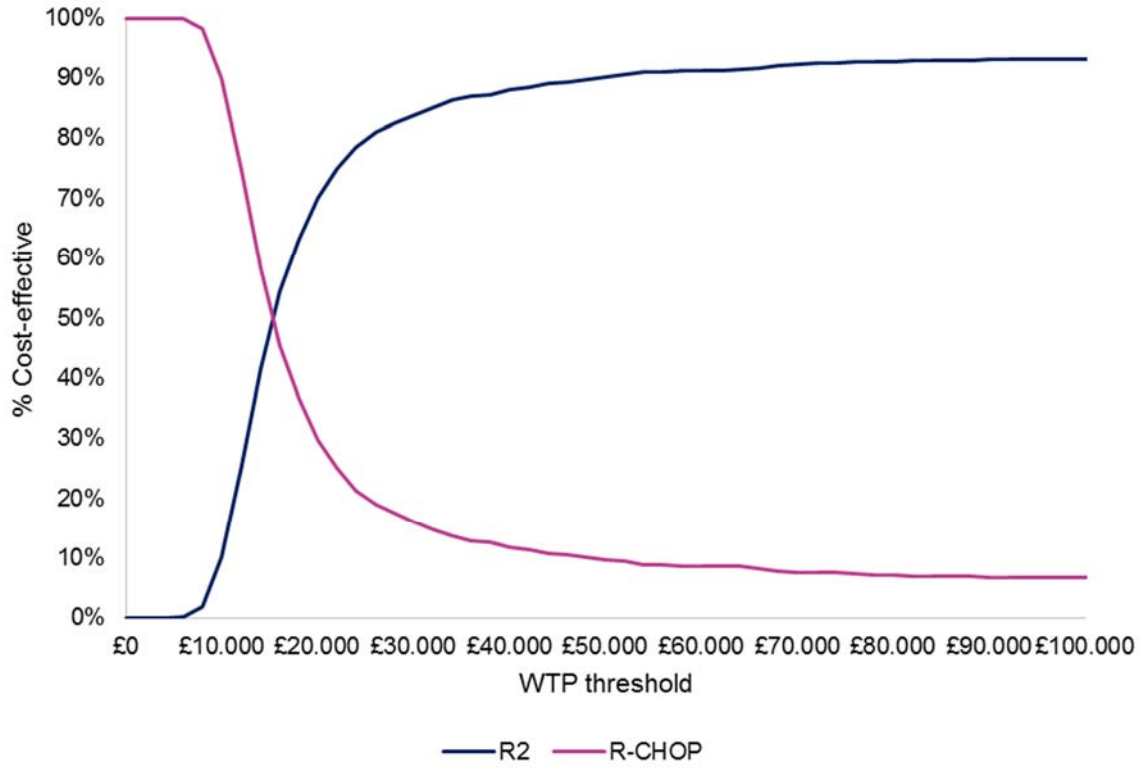


Figure 5.6: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP

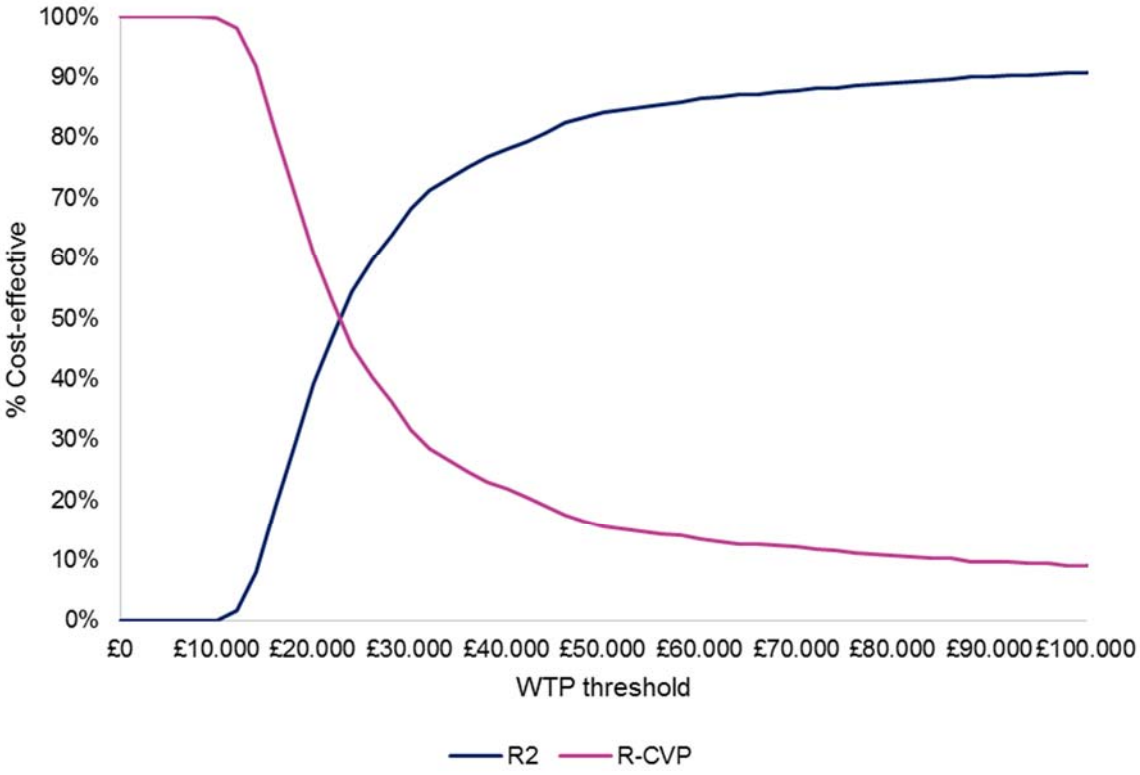
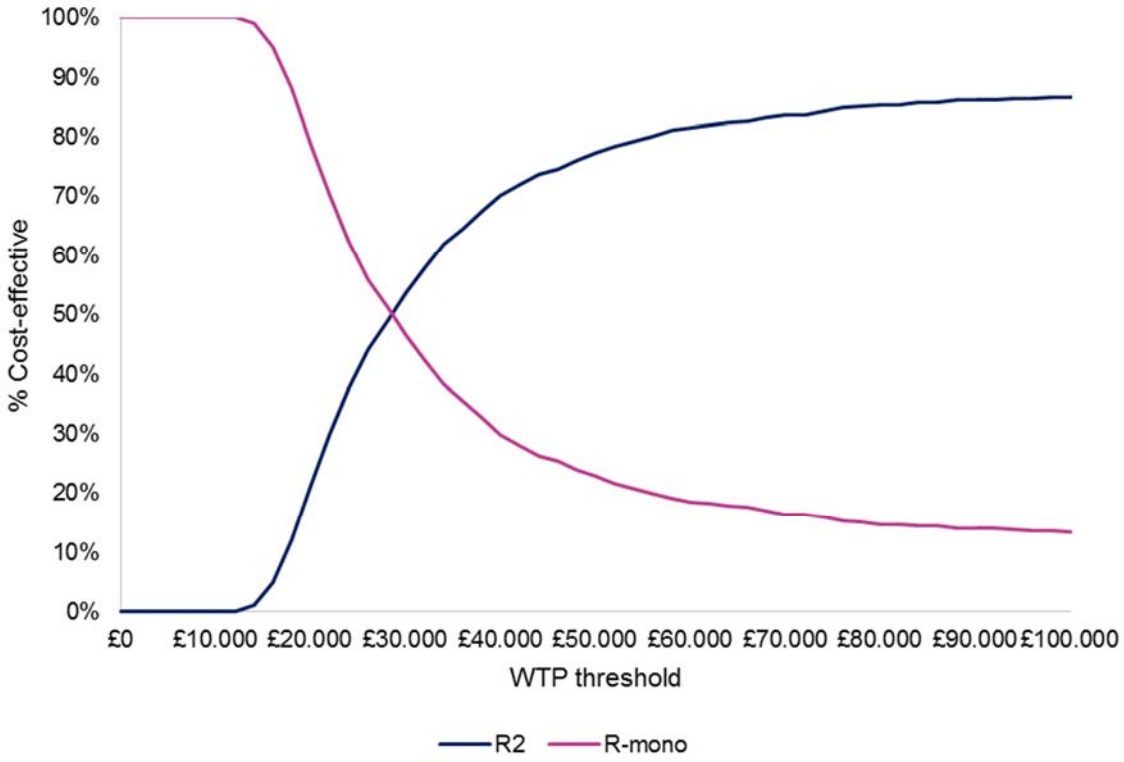


Figure 5.7: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono



5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These were all performed using the ERG base-case. Results are presented in Table 6.2 in Section 6 of this report.

Exploratory analyses using the ERG base-case:

1. Alternative PFS distributions: use Weibull for PFS both arms (for the R-mono comparison, generalised gamma was used as the alternative PFS distribution) (section 5.2.6)
2. Alternative PFS distributions: use exponential For PFS R^2 and Weibull for PFS comparator (not applied to R-mono comparison) (section 5.2.6)
3. Treatment waning effect after three-year cut-off (section 5.2.6)
4. Treatment waning effect after seven-year cut-off (section 5.2.6)
5. Adverse events for comparator taken from van Oers et al. (2006)[39] (Not applicable in R-mono comparison) (section 5.2.7)
6. FL-only population (section 5.2.3)
7. Apply same subsequent treatment costs for R^2 as for R-CHOP/R-CVP (Not applicable in R-mono comparison) (section 5.2.9)
8. Alternative utilities taken from Wild et al. (2006)[70] 0.805 for PF, 0.736 for PP off treatment, and 0.62 for PP on treatment (section 5.2.8)
9. Source for R-CHOP efficacy taken from van Oers et al. (Not applicable in R-mono comparison) (section 5.2.6)
10. Alternative utilities taken for PP states taken from Pereira et al. (2010)[95] 0.45 for both PP states. (section 5.2.8)

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

Separate sets of searches were conducted to identify cost effectiveness studies, health-related quality of life studies and healthcare resource use evidence. The CS provided clear, transparent and reproducible searches. A good range of databases and additional resources were searched.

The company submission was largely in line with the NICE reference case. The CS partly deviated from the scope, however, where it concerned the comparators modelled. More specifically, R-mono was excluded while direct evidence existed for R^2 versus R-mono, and in the refractory population O-Benda was the sole comparator while NICE had explicitly stated it was not a relevant comparator for this appraisal.

The ERG had concerns about the appropriateness of the partitioned survival model approach and its superiority over a state transition model and would have liked to see both approaches properly explored.

The ERG was concerned about the company pooling MZL and FL populations in the model, assuming they are similar. The ICER for the company's FL-only scenario was substantially higher for the R-CHOP and R-CVP comparisons. This raises serious doubts about the validity of this assumption, and the ERG considered this to be a relevant source of uncertainty.

The main concern of the ERG was the questionable trustworthiness of R^2 efficacy resulting from the indirect comparison, which seemed to be inflated relative to the direct comparison data from AUGMENT. Although the ERG did not have the necessary data to quantify this uncertainty, it may have lowered the ICER substantially.

The ERG had concerns about the way survival curves were selected. The choice of OS curve was mainly based on a previous STA. In particular the choice of PFS curves was not sufficiently justified and appeared sub-optimal, with a likely overestimation of PFS in the R^2 arm, and substantial underestimation of PFS in the first year for R-CHOP and R-CVP. This matter was exacerbated by the high utility values for all health states. The ERG considered these to be potentially overestimated, being higher than or comparable to those in the general population. With utilities remaining high throughout the model, any adjustment in survival curves only had little impact on the ICER, as a high utility post-progression implied there was hardly any penalty on progression in terms of cost effectiveness.

The ERG considered the source used to inform the model concerning AE incidences for R-CHOP and R-CVP to be likely biased, being based on a previously untreated population.

With respect to costs and resource use, the ERG considered the costs of subsequent treatment for R-CHOP and R-CVP to be likely overestimated, as they were based on a mixed R-chemo population from HMRN, while also data specific to R-CHOP and R-CVP separately were available from this source. This was adjusted for in the ERG base-case. The ERG was also concerned about the fact that in the post-progression on treatment phase, there would be a one-off cost for subsequent treatments only, which may be not be reflective of the long-term situation in this health state. As patients in the R^2 arm remain in this health state for a longer time on average, applying costs as one-off possibly favoured R^2.

The ERG made various adjustments to the company base-case. The probabilistic ERG base-case for R^2 versus R-CHOP was £15,818 per QALY gained (based on 1,000 iterations). For R^2 versus R-CVP, the ICER was £23,367 and for R^2 versus R-mono, it was £29,010.

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost-effectiveness estimates. For the R-CHOP/R-CVP comparisons, using R-CHOP and R-CVP efficacy from van Oers et al. would lower the ICER substantially, £8,251 for R^2 versus R-

CHOP and £13,315 for R^2 versus R-CVP. Alternative assumptions regarding lowered utilities in the PP health states had the most significant upward impact, increasing the ICER to £33,626 for R^2 versus R-CHOP and £47,281 for R^2 versus R-CVP. For the R-mono comparison, lowering the PP health state utility had the opposite effect, lowering the ICER to £17,826. Another influential scenario was the change of time-point where treatment waning starts to three years (instead of five years in base-case). This increased the ICER to £40,543.

In conclusion, even though the ERG base-case ICER for R-CHOP was below £20,000, the uncertainty around the cost effectiveness of R^2 is substantial, mainly caused by the possible bias introduced by the indirect treatment comparison, which could not be accounted for in the ERG analyses. The ICER for R-CVP is higher and suffers from the same uncertainty. The R-mono analysis is based on a direct comparison, but is also surrounded by substantial uncertainty, as the ICER is rather sensitive to, for instance, the time-point at which treatment waning starts and utilities in the PP health state.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

6.1 Analyses undertaken by the ERG

In Section 5.3 the ERG base-case was presented, which was based on various changes compared to the company base-case. Tables 6.1 to 6.3 show how individual changes impact the results plus the combined effect of all changes simultaneously, for the R-CHOP, R-CVP, and R-mono comparators, respectively. The exploratory scenario analyses are presented in Tables 6.4 to 6.6 respectively. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.1 to 6.6 correspond to the analyses numbers reported in Section 5.3 of this report. The submitted model files contain technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment). The ERG wishes to emphasise that all ERG analyses (except the R-mono comparison) are conditional upon the MAIC results for which uncertainty could not be quantified or incorporated in the economic model.

Although the tables below report pairwise comparisons only, R-CHOP and R-CVP could also be compared to R^2 in a fully incremental analysis. However, as R-CHOP and R-CVP are by assumption equally effective, and R-CHOP is always the more costly strategy given the higher rate of ASCT performed in the R-CHOP patient population, it is not to be expected that there will be any shifts in the relative comparisons within the fully incremental analysis. Therefore, in practice, the relevant comparison will be R^2 versus R-CVP. For R-mono, a full incremental analysis on the scenarios is not applicable, as a different set of scenarios was performed here, and, more importantly, because costs and QALYs in R^2 are different in this comparison.

Table 6.1: Deterministic ERG base-case for R2 versus R-CHOP comparison

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,471
R-CHOP	[redacted]	[redacted]			
Fixing violations (1, include AEs related to subs Tx in R^2)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,544
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx instead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,206
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,977
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,345
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R^2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,429
R-CHOP	[redacted]	[redacted]			

Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,484
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,505
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,818
R-CHOP	[redacted]	[redacted]			

Table 6.2: Deterministic ERG base-case for R2 versus R-CVP comparison

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,814
R-CVP	[redacted]	[redacted]			
Fixing violations (1, include AEs related to subs Tx in R^2)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,888
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx instead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,549
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,557
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,304
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R^2 and Weibull for PFS compator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,875
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,867
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,759
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,367
R-CVP	[redacted]	[redacted]			

Table 6.3: Deterministic ERG base-case for R2 versus R-mono comparison

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,580
R-mono	[redacted]	[redacted]			
Fixing violations (1, include AEs related to subs Tx in R^2)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,673
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,054
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,318
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,372
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,010
R-mono	[redacted]	[redacted]			

Table 6.4: Deterministic scenario analyses (conditional on ERG base-case) for R2 versus R-CHOP

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,505
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,632
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,915
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,018
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,654
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,270

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
R-CHOP	[redacted]	[redacted]			
FL-only population					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,680
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,640
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,526
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£8,251
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£33,626
R-CHOP	[redacted]	[redacted]			

Table 6.5: Deterministic scenario analyses (conditional on ERG base-case) for R2 versus R-CVP

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,759
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,887
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,167
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,562
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,523
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,618
R-CVP	[redacted]	[redacted]			
FL-only population					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,841

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,899
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,227
R-CVP	[redacted]	[redacted]			
Source for R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,315
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£47,281
R-CVP	[redacted]	[redacted]			

Table 6.6: Deterministic scenario analyses (conditional on ERG base-case) for R2 versus R-mono

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,372
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,206
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£40,543
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,091
R-mono	[redacted]	[redacted]			
FL-only population					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,936
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,263
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,349
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,826

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
R-mono	[redacted]	[redacted]			

7. REFERENCES

[1] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Document B. Company evidence submission to National Institute for Health and Care Excellence. Single technology appraisal (STA)*: Celgene Ltd, 2019

[2] Haematological Malignancy Research Network (HMRN). Incidence [Internet]. York: Haematological Malignancy Research Network, Epidemiology & Statistics Group (ECSG) at the University of York, 2019 [accessed 1.5.19]. Available from: https://www.hmrn.org/statistics/incidence

[3] Haematological Malignancy Research Network (HMRN). *Patient's age and treatment for haematological malignancy: a report from the Haematological Malignancy Research Network (HMRN) [Internet]*. York: Haematological Malignancy Research Network, 2014 [accessed 1.5.19] Available from: https://www.hmrn.org/Download.aspx?target=document&id=41

[4] Haematological Malignancy Research Network (HMRN). Survival: follicular lymphoma [Internet]. York: Haematological Malignancy Research Network, Epidemiology & Statistics Group (ECSG) at the University of York, 2019 [accessed 1.5.19]. Available from: https://www.hmrn.org/statistics/survival

[5] Cancer Research UK. Non-Hodgkin lymphoma: survival [Internet]. Cancer Research UK, 2018 [accessed 1.5.19]. Available from: https://about-cancer.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/survival

[6] Provencio M, Sabin P, Gomez-Codina J, Torrente M, Calvo V, Llanos M, et al. Impact of treatment in long-term survival patients with follicular lymphoma: a Spanish Lymphoma Oncology Group registry. *PLoS One* 2017;12(5):e0177204.

[7] Johnson PW, Rohatiner AZ, Whelan JS, Price CG, Love S, Lim J, et al. Patterns of survival in patients with recurrent follicular lymphoma: a 20-year study from a single center. *J Clin Oncol* 1995;13(1):140-7.

[8] 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.

[9] El-Zimaity M. Marginal zone B-cell lymphoma [Internet]. Medscape, 2014 [accessed 1.5.19]. Available from: https://emedicine.medscape.com/article/1610599-overview

[10] Olszewski AJ, Castillo JJ. Survival of patients with marginal zone lymphoma: analysis of the Surveillance, Epidemiology, and End Results database. *Cancer* 2013;119(3):629-38.

[11] Denlinger NM, Epperla N, William BM. Management of relapsed/refractory marginal zone lymphoma: focus on ibrutinib. *Cancer Manag Res* 2018;10:615-24.

[12] Office of National Statistics (ONS). Cancer Registration Statistics, England, 2017. (Updated: 26 April 2019) [Internet]. Office of National Statistics (ONS), 2019 [accessed 1.5.19]. Available from: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland/2017/finalreferencetables.xlsx

[13] Wang HI, Roman E, Crouch S, Aas E, Burton C, Patmore R, et al. A generic model for follicular lymphoma: predicting cost, life expectancy, and quality-adjusted-life-year using UK population-based observational data. *Value Health* 2018;21(10):1176-85.

[14] Lymphoma Action. MALT lymphoma (extranodal marginal zone lymphoma) [Internet]. Lymphoma Action, 2018 [accessed 1.5.19]. Available from: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-08/LYMweb0218MALT2017v1.2.pdf

[15] Lymphoma Action. Splenic marginal zone lymphoma [Internet]. Lymphoma Action, 2018 [accessed 1.5.19]. Available from: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-06/LYMweb0220SplenicMZ2017v1.2.pdf

[16] Lymphoma Action. Nodal marginal zone lymphoma [Internet]. Lymphoma Action, 2019 [accessed 1.5.19]. Available from: https://lymphoma-action.org.uk/sites/default/files/media/documents/2018-06/LYMweb0219NodalMZ2017v1.2.pdf

[17] Chiu H, Trisal P, Bjorklund C, Carrancio S, Torano EG, Guarinos C, et al. Combination lenalidomide-rituximab immunotherapy activates anti-tumour immunity and induces tumour cell death by complementary mechanisms of action in follicular lymphoma. *Br J Haematol* 2019;185(2):240-253.

[18] Celgene. *Draft summary of product characteristics. Revlimid® (Lenalidomide)*: Celgene, 2019

[19] Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0086) March 2019. (Updated: August 2019) Data on file. 2019.

[20] National Institute for Health and Care Excellence. *Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab. NICE technology appraisal guidance 472 [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 5.1.19] Available from: https://www.nice.org.uk/guidance/ta472

[21] National Institute for Health and Care Excellence. *Rituximab for the first-line treatment of stage III-IV follicular lymphoma. NICE technology appraisal guidance 243 [Internet]*. London: National Institute for Health and Care Excellence, 2012 [accessed 26.4.19] Available from: https://www.nice.org.uk/guidance/ta243

[22] National Institute for Health and Care Excellence. *Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma. NICE technology appraisal guidance 226 [Internet]*. London: National Institute for Health and Care Excellence, 2011 [accessed 5.5.19] Available from: https://www.nice.org.uk/guidance/ta226

[23] National Institute for Health and Care Excellence. *Obinutuzumab for untreated advanced follicular lymphoma. NICE technology appraisal guidance 513 [Internet]*. London: National Institute for Health and Care Excellence, 2018 [accessed 5.1.19] Available from: https://www.nice.org.uk/guidance/ta513

[24] National Institute for Health and Care Excellence. NICE Pathways: Treating follicular lymphoma [Internet]. London: National Institute for Health and Care Excellence, Updated Jan 2019 [accessed 13.2.19]. Available from: https://pathways.nice.org.uk/pathways/non-hodgkins-lymphoma/treating-follicular-lymphoma#content=view-node%3Anodes-first-line-treatment-for-early-stage-disease

[25] National Institute for Health and Care Excellence. *Non-Hodgkin's lymphoma: diagnosis and management. NICE guideline 52 [Internet]*. London: National Institute for Health and Care Excellence, 2016 [accessed 25.4.19] Available from: https://www.nice.org.uk/guidance/ng52

[26] Celgene Ltd. *Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Response to request for clarification from the ERG. August 2019*: Celgene Ltd, 2019

[27] National Institute for Health and Care Excellence. Cancer Drugs Fund [Internet]. London: National Institute for Health and Care Excellence, [accessed 3.9.19]. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/cancer-drugs-fund

[28] National Institute for Health and Care Excellence. *Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma. Response to consultee and commentator comments on the draft remit and draft scope (pre-referral) [Internet]*. London: National Institute for Health and Care Excellence, 2019 [accessed 25.7.19] Available from:

https://www.nice.org.uk/guidance/gid-ta10323/documents/scope-consultation-comments-and-responses

[29] Dreyling M, Thieblemont C, Gallamini A, Arcaini L, Campo E, Hermine O, et al. ESMO Consensus conferences: guidelines on malignant lymphoma. Part 2: marginal zone lymphoma, mantle cell lymphoma, peripheral T-cell lymphoma. *Ann Oncol* 2013;24(4):857-77.

[30] Lymphoma Research Foundation. *Getting the facts. Marginal zone lymphoma [Internet]*: Lymphoma Research Foundation, 2018 [accessed 5.5.19] Available from: https://www.lymphoma.org/wp-content/uploads/2018/07/MarginalZoneLymphomaF_Fact-Sheet_2018.pdf

[31] Celgene Ltd. [redacted] clinical validation. 12 June 2019. Data on file. 2019.

[32] National Institute for Health and Care Excellence. *Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2019 [accessed 25.7.19] Available from: https://www.nice.org.uk/guidance/gid-ta10323/documents/final-scope

[33] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Document A. Company evidence submission summary for committee. Single technology appraisal (STA)*: Celgene Ltd, 2019

[34] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Document B. Appendices. Single technology appraisal (STA)*: Celgene Ltd, 2019

[35] 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.

[36] Andorsky D, Coleman M, Yacoub A, Melear JM, Brooks HD, Fanning SR, et al. Response rate to lenalidomide plus rituximab (R2) as independent of number of prior lines of therapy: interim analysis of initial phase of MAGNIFY phase IIIb study of R2 followed by maintenance in relapsed/refractory indolent NHL. *J Clin Oncol* 2018;36(15 Suppl):7516.

[37] Leonard JP, Jung SH, Johnson J, Pitcher BN, Bartlett NL, Blum KA, et al. Randomized trial of lenalidomide alone versus lenalidomide plus rituximab in patients with recurrent follicular lymphoma: CALGB 50401 (ALLIANCE). *J Clin Oncol* 2015;33(31):3635-40.

[38] Tuscano JM, Dutia M, Chee K, Brunson A, Reed-Pease C, Abedi M, et al. Lenalidomide plus rituximab can produce durable clinical responses in patients with relapsed or refractory, indolent non-Hodgkin lymphoma. *Br J Haematol* 2014;165(3):375-81.

[39] van Oers MH, 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.

[40] Sehn LH, Chua N, Mayer J, Dueck G, Trněný 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.

[41] Phillippo D, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. *NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE [Internet]*. Sheffield: NICE Decision Support Unit, 2016 [accessed 2.5.19] Available from: http://www.nicedsu.org.uk

[42] Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12:9.

[43] 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* 2018;21(10):960-7.

[44] 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.

[45] All Wales Therapeutics & Toxicology Centre. *AWMSG Secretariat Assessment Report. Idelalisib (Zydelig®) 100 mg and 150 mg film-coated tablets. Reference number: 2597 [Internet]*: All Wales Therapeutics & Toxicology Centre, 2017 [accessed 14.3.19] Available from: http://www.awmsg.org/awmsgonline/app/appraisalinfo/2597

[46] Scottish Medicines Consortium (SMC). *Idelalisib, 100mg and 150mg film-coated tablets (Zydelig®). SMC No. (1039/15) [Internet]*. Glasgow: Scottish Medicines Consortium (SMC), 2015 [accessed 14.3.19] Available from: https://www.scottishmedicines.org.uk/

[47] Blommestein HM, Issa DE, Pompen M, Ten Hoor G, Hogendoorn M, Joosten P, et al. Cost-effectiveness of rituximab as maintenance treatment for relapsed follicular lymphoma: results of a population-based study. *Eur J Haematol* 2014;92(5):398-406.

[48] Pink J, Lane S, Hughes DA. Mechanism-based approach to the economic evaluation of pharmaceuticals: pharmacokinetic/pharmacodynamic/pharmacoeconomic analysis of rituximab for follicular lymphoma. *Pharmacoeconomics* 2012;30(5):413-29.

[49] 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.

[50] Deconinck E, Miadi-Fargier H, Pen CL, Brice P. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: long-term economic evaluation. *Pharmacoeconomics* 2010;28(1):35-46.

[51] Papadakis K, Follows GA, Boyer J, Bashir Z, Ball P, Aultman R, et al. Cost effectiveness analysis of rituximab maintenance in patients with untreated high tumour burden follicular lymphoma after response to immunochemotherapy: a UK national healthcare services perspective. *Blood* 2010;116(21):3833.

[52] Follicular Lymphoma Pharmacoeconomics Group. [Rituximab cost analysis for maintenance treatment of patients with follicular lymphoma]. *Farm Hosp* 2008;32(1):25-34.

[53] Chiattone C, Saggia MG, Santos EV. Cost-utility analysis of maintenance therapy with rituximab for follicular lymphoma in Brazil according to local HTA guidelines. *Ann Oncol* 2008;19(Suppl 8):737P.

[54] Hayslip JW, Simpson KN. Cost-effectiveness of extended adjuvant rituximab for US patients aged 65-70 years with follicular lymphoma in second remission. *Clin Lymphoma Myeloma* 2008;8(3):166-70.

[55] Hornberger J, Reyes C, Lubeck D, Valente N. Economic evaluation of rituximab plus cyclophosphamide, vincristine and prednisolone for advanced follicular lymphoma. *Leuk Lymphoma* 2008;49(2):227-36.

[56] Kasteng F, Erlanson M, Hagberg H, Kimby E, Relander T, Lundkvist J. Cost-effectiveness of maintenance rituximab treatment after second line therapy in patients with follicular lymphoma in Sweden. *Acta Oncol* 2008;47(6):1029-36.

[57] Scottish Medicines Consortium (SMC). *Rituximab 10mg/ml concentrate for infusion (MabThera®) (No.330/06) [Internet]*. Glasgow: Scottish Medicines Consortium (SMC), 2006 [accessed 14.3.19] Available from: https://www.scottishmedicines.org.uk/

[58] Scottish Medicines Consortium (SMC). *Ibritumomab tiuxetan (Zevalin®) No. (171/05). Re-submission [Internet]*. Glasgow: Scottish Medicines Consortium (SMC), 2007 [accessed 14.3.19] Available from: https://www.scottishmedicines.org.uk/

[59] National Institute for Health and Care Excellence. *Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. NICE technology appraisal guidance 137 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.2.19] Available from: https://www.nice.org.uk/guidance/ta137

[60] Guzauskas GF, Masaquel A, Thuresson P-O, Li J, Reyes CM, Veenstra D. Is obinutuzumab plus chemotherapy followed by obinutuzumab monotherapy for previously untreated follicular lymphoma patients cost-effective in the US? *Blood* 2017;130:5603.

[61] Scottish Medicines Consortium (SMC). Obinutuzumab 1,000mg concentrate for solution for infusion (Gazyvaro®) [Internet]. Glasgow: Scottish Medicines Consortium (SMC), 2017 [accessed 14.3.19]. Available from: https://www.scottishmedicines.org.uk/

[62] Soini EJ, 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* 2011;22(5):1189-97.

[63] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU technical support document 19: partitioned survival analysis for decision modelling in health care: a critical review [Internet]*. Sheffield: NICE Decision Support Unit, 2017 [accessed 28.5.19] Available from: http://www.nicedsu.org.uk

[64] Leonard JP, Trneny M, Izutsu K, et al. AUGMENT phase III study: Lenalidomide/rituximab (R2) improved efficacy over rituximab/placebo in relapsed/refractory follicular patients irrespective of POD24 status. Paper submitted to 15-ICML International Conference on Malignant Lymphoma; 18-22 Jun 2019; Lugano: Switzerland. 2019.

[65] Haematological Malignancy Research Network (HMRN). Clinical management and outcome of follicular lymphoma (FL) with a focus on relapsed/refractory disease (draft report). 17 July 2019. Haematological Malignancy Research Network, 2019.

[66] Morschhauser F, Fowler NH, Feugier P, Bouabdallah R, Tilly H, Palomba ML, et al. Rituximab plus lenalidomide in advanced untreated follicular lymphoma. *N Engl J Med* 2018;379(10):934-47.

[67] 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.

[68] National Institute for Health and Care Excellence. *Non-Hodgkin's lymphoma: diagnosis and management. NICE guideline 52 [Internet]*. London: National Institute for Health and Care Excellence, 2016 [accessed 25.7.19] Available from: https://www.nice.org.uk/guidance/ng52

[69] Leger C, Sabloff M, McDiarmid S, Bence-Bruckler I, Atkins H, Bredeson C, et al. Outpatient autologous hematopoietic stem cell transplantation for patients with relapsed follicular lymphoma. *Ann Hematol* 2006;85(10):723-9.

[70] Wild D, Walker M, Pettengell R, Lewis G. Pcn62 utility elicitation in patients with follicular lymphoma. *Value Health* 2006;9(6):A294.

[71] Pereira C, Negreiro F, Silva C. PSY43 Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. *Value Health* 2010;13(7):A468.

[72] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[73] National Institute for Health and Care Excellence. *Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma. NICE technology appraisal guidance 306 [Internet]*. London: National Institute for Health and Care Excellence, 2014 [accessed 25.4.19] Available from: https://www.nice.org.uk/guidance/ta306

[74] Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin* 2010;26(5):1091-6.

[75] Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 2010;8:50.

[76] Stein EM, Yang M, Guerin A, Gao W, Galebach P, Xiang CQ, et al. Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health Qual Life Outcomes* 2018;16(1):193.

[77] Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ* 2013;14(5):749-59.

[78] National Institute for Health and Care Excellence. *Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens. NICE technology appraisal guidance 423 [Internet]*. London: National Institute for Health and Care Excellence, 2016 [accessed 25.4.19] Available from: https://www.nice.org.uk/guidance/ta423

[79] Hannouf MB, Sehgal C, Cao JQ, Mocanu JD, Winquist E, Zaric GS. Cost-effectiveness of adding cetuximab to platinum-based chemotherapy for first-line treatment of recurrent or metastatic head and neck cancer. *PLoS One* 2012;7(6):e38557.

[80] Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer* 2008;62(3):374-80.

[81] Kind P, Hardman G, Macran S. *UK population norms for EQ-5D. Discussion paper 172*. York: Centre for Health Economics, 1999 Available from: https://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20172.pdf

[82] NHS Improvement. *NHS reference costs 2017/18 [Internet]*. London: NHS Improvement, 2018 [accessed 26.4.19] Available from: https://improvement.nhs.uk/resources/reference-costs/

[83] Personal Social Services Research Unit. *Unit costs of health and social care 2018*. Canterbury: University of Kent, 2018 [accessed 25.7.19] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/

[84] MIMS. MabThera [Internet]. 2019 [accessed 27.2.19]. Available from: https://www.mims.co.uk/drugs/cancer/antineoplastics/mabthera

[85] MIMS. Revlimid [Internet]. 2019 [accessed 25.4.19]. Available from: https://www.mims.co.uk/drugs/cancer/antineoplastics/revlimid

[86] MIMS. Neupogen [Internet]. 2019 [accessed 31.5.19]. Available from: https://www.mims.co.uk/drugs/cardiovascular-system/anaemia-neutropenia/neupogen

[87] MIMS. Allopurinol [Internet]. 2019 [accessed 31.5.19]. Available from: https://www.mims.co.uk/drugs/musculoskeletal-disorders/gout/allopurinol

[88] MIMS. Gazyvaro [Internet]. 2019 [accessed 27.2.19]. Available from: https://www.mims.co.uk/drugs/cancer/antineoplastics/gazyvaro

[89] MIMS. Rixathon [Internet]. 2019 [accessed 27.2.19]. Available from: https://www.mims.co.uk/drugs/cancer/antineoplastics/rixathon

[90] MIMS. Zydelig [Internet]. 2019 [accessed 1.6.19]. Available from: https://www.mims.co.uk/drugs/cancer/antineoplastics/zydelig

[91] GOV.UK. Drugs and pharmaceutical electronic market information tool (eMIT) [Internet]. 2019 [accessed 25.4.19]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[92] Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: a modelling study. *Palliat Med* 2015;29(10):899-907.

[93] Vemer P, Corro Ramos I, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. *Pharmacoeconomics* 2016;34(4):349-61.

[94] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

[95] Pereira C, Negreiro F, Silva C. Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. PSY43. *Value Health* 2010;13(7):A468.

Appendix 1: Additional results from the AUGMENT trial

Table 4.5 in Section 4.2.5 of this report presents a summary of the main results from the AUGMENT trial for the full ITT population. Results for FL and MZL separately are reported in Tables A1.1 and A1.2 below.

Table A1.1: Summary of results from the AUGMENT trial: ITT population (FL).

Endpoint	FL	
	R^2 (n=147)	**R-mono (n=148)**
Median OS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted][c]	
Median PFS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted][c]	
Best response, n (%)		
ORR (CR+PR)	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
CR rate	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
PR	[redacted]	[redacted]
SD	[redacted]	[redacted]
PD/ death	[redacted]	[redacted]
No evidence of disease	[redacted]	[redacted]
Unknown/ND/Missing	[redacted]	[redacted]
Median TTNLT, months (95% CI)[a]	[redacted]	[redacted]
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	
Median EFS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	
Median TTNCT, months (95% CI)[a]	NR	NR
TTNCT rate at 2 years, % (95% CI)	NR	NR
Hazard ratio (95% CI)	NR	
p-value	NR	
RTNLT		
ORR, n (% [95% CI][d])	NR	NR
p-value	NR	
CR, n (% [95% CI][d])	NR	NR
p-value	NR	
DCRR, n/N (%)	[redacted]	[redacted]

95% CI[d]	[redacted]
p-value	[redacted]
N, Median DOR, months (95% CI)[a]	[redacted]
Hazard ratio (95% CI)[c]	[redacted]
p-value[e]	[redacted]
N, Median DOCR, months (95% CI)[a]	[redacted]
Hazard ratio (95% CI)[h]	[redacted]
p-value	[redacted]

Source: Response to CL, Table 5, pages 19 and 20.
CI = confidence interval; CR = complete response; DCRR = durable complete response rate, DOCR = duration of complete response; DOR = duration of response; EFS = event-free survival; FL = follicular lymphoma; IRC = Independent Review Committee; ITT = intent-to-treat; MZL = marginal zone lymphoma; ND = not done; NE = not estimable; ORR = overall response rate; OS = overall survival; PD = progressive disease; PR = partial response; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo; RTNLT = response rate to next anti-lymphoma treatment; SD = stable disease; TTNLT = time to next anti-lymphoma treatment; TTNCT = time to next anti-lymphoma chemotherapy treatment.
Notes: [a]) median estimate is from Kaplan–Meier analysis; [b]) from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last anti-lymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c]) from Cox proportional hazard model; [d]) exact confidence interval for binomial distribution; [e]) from CMH test adjusting for the three stratification factors; [f]) from Fisher-Exact test; [g]) from log-rank test adjusting for the three stratification factors; [h]) from log-rank test

Table A1.2: Summary of results from the AUGMENT trial: ITT population (MZL).

Endpoint	**MZL**	
	R^2 (n=31)	**R-mono (n=32)**
Median OS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
Median PFS, months (95% CI)[a]	24.9 ([redacted])	25.2 ([redacted])
Hazard ratio (95% CI)	[redacted]	
Best response, n (%)		
ORR (CR+PR)	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
CR rate	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
PR	[redacted]	[redacted]
SD	[redacted]	[redacted]
PD/ death	[redacted]	[redacted]
No evidence of disease	[redacted]	[redacted]
Unknown/ND/Missing	[redacted]	[redacted]
Median TTNLT, months (95% CI)[a]	[redacted]	[redacted]
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	

Endpoint	MZL	
	R^2 (n=31)	R-mono (n=32)
Median EFS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	
Median TTNCT, months (95% CI)[a]	NR	NR
TTNCT rate at 2 years, % (95% CI)	NR	NR
Hazard ratio (95% CI)	NR	
p-value	NR	
RTNLT		
ORR, n (% [95% CI][d])	NR	NR
p-value	NR	
CR, n (% [95% CI][d])	NR	NR
p-value	NR	
DCRR, n/N (%)	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
N, Median DOR, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)[c]	[redacted]	
p-value[e]	[redacted]	
N, Median DOCR, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)[h]	[redacted]	
p-value	[redacted]	

Source: Response to CL, Table 2, pages 20 and 21.
CI = confidence interval; CR = complete response; DCRR = durable complete response rate, DOCR = duration of complete response; DOR = duration of response; EFS = event-free survival; FL = follicular lymphoma; IRC = Independent Review Committee; ITT = intent-to-treat; MZL = marginal zone lymphoma; ND = not done; NE = not estimable; ORR = overall response rate; OS = overall survival; PD = progressive disease; PR = partial response; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo; RTNLT = response rate to next anti-lymphoma treatment; SD = stable disease; TTNLT = time to next anti-lymphoma treatment; TTNCT = time to next anti-lymphoma chemotherapy treatment.
Notes: [a]) median estimate is from Kaplan–Meier analysis; [b]) from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last anti-lymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c]) from Cox proportional hazard model; [d]) exact confidence interval for binomial distribution; [e]) from CMH test adjusting for the three stratification factors; [f]) from Fisher-Exact test; [g]) from log-rank test adjusting for the three stratification factors; [h]) from log-rank test

Summaries of the treatment-emergent adverse event (TEAEs) during AUGMENT for the FL and MZL populations separately are is presented in Tables A1.3 and A1.4, respectively.

Table A1.3: Summary of treatment-emergent adverse events in AUGMENT: FL Safety population

	FL	
	R^2 (n=146)	R-mono (n=148)
Number of patients (%)		
Any TEAE	[redacted]	[redacted]

	FL	
	R^2 (n=146)	**R-mono (n=148)**
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 3–4 TEAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 5 TEAE	[redacted]	[redacted]
Any SAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of R	[redacted]	[redacted]
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to discontinuation of R	[redacted]	[redacted]
Source: Clarification Letter, Table 6, page 21. FL = follicular lymphoma; Len = lenalidomide; Pbo = placebo; R = rituximab; R^2 = lenalidomide + rituximab; R mono= placebo, rituximab + placebo; SAE = serious adverse event; TEAE = treatment-emergent adverse event.		

Table A1.4: Summary of treatment-emergent adverse events in AUGMENT: MZL Safety population

	MZL	
	R^2 (n=30)	**R-mono (n=32)**
Number of patients (%)		
Any TEAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 3–4 TEAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 5 TEAE	[redacted]	[redacted]
Any SAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of R	[redacted]	[redacted]
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to discontinuation of R	[redacted]	[redacted]
Source: Clarification Letter, Table 6, page 21.		

Len = lenalidomide; MZL = marginal zone lymphoma; Pbo = placebo; R = rituximab; R2 = lenalidomide + rituximab; R mono = placebo, rituximab + placebo; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

The most common TEAEs, occurring in more than 10% of patients, are presented in Tables A1.5 and A1.6 for FL and MZL patients, respectively.

Table A1.5: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT: FL Safety population

	FL	
	R^2 (n=146)	**R-mono (n=148)**
Number of patients (%)		
Blood and lymphatic system disorders	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Constipation	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
URTI	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Investigations	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]

	FL	
	R^2 (n=146)	**R-mono (n=148)**
Decreased appetite	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]
Tumour flare	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]
Source: Clarification Letter, Table 7, pages 22-23. FL = follicular lymphoma; R^2 = lenalidomide + rituximab; R = placebo, rituximab + placebo; URTI = upper respiratory tract infection.		

Table A1.6: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT: MZL Safety population

	MZL	
	R^2 (n=30)	**R-mono (n=32)**
Number of patients (%)		
Blood and lymphatic system disorders	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Constipation	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
URTI	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]

	MZL	
	R^2 (n=30)	**R-mono (n=32)**
Pruritus	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Investigations	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]
Tumour flare	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]

Source: Clarification Letter, Table 7, pages 22-23.
MZL = marginal zone lymphoma; R^2 = lenalidomide + rituximab; R = placebo, rituximab + placebo; URTI = upper respiratory tract infection.

Appendix 2: MAIC reporting checklist

According to the NICE DSU, the following themes should be considered and addressed explicitly, when reporting population-adjusted analyses (See TSD 18, pages 64-65).[41]

Criteria	Addressed in CS (Y/N)	ERG Comments
1. The variables available in each study should be listed, along with their distributions (e.g. through box plots or histograms).	Y	The variables were listed along with summary statistics, although there were no plots of their distributions
2. Sufficient covariate overlap between the populations should be assessed: - for population reweighting methods (such as MAIC), the number of individuals assigned zero weight should be reported; - for outcome regression methods (such as STC), the amount of extrapolation required should be considered. - For anchored comparisons this applies only to effect modifiers (see point 2); - for unanchored comparisons all variables relevant to outcome should be presented.	Y	The CS used a MAIC and the details of the weighting, number of zero weights were provided.
3. Evidence for effect modifier status should be given, along with the proposed size of the interaction effect and the imbalance between the study populations.	N	No information about those variables considered to be effect modifiers and their interaction with the treatment effect.
4. The resulting potential bias reduction compared with a standard indirect comparison may be calculated by multiplying the interaction coefficient by the difference in means.	N	For some analyses there was also an unadjusted indirect comparison but there was no estimate of the bias reduction.
5. The distribution of weights should be presented for population weighting analyses, and used to highlight any issues with extreme or highly variable weights.	Y	Histograms showing the distribution of the weights were provided
6. Presentation of the effective sample size may also be useful.	Y	The ESS was reported for each matched analysis
7. ESS may be approximated using equation (7) – which is likely to be an underestimate – but provides clear warning where inferences are being made based on just a small number of individuals Measures of uncertainty, such as confidence intervals, should always be presented alongside any estimates. - Care should be taken that uncertainty is appropriately propagated through to the final estimates. - For outcome regression methods, uncertainty is fully propagated for predictions into the	Y	Confidence intervals were reported. Standard errors for the survival analyses (Cox and parametric models) were calculated using robust sandwich estimators. A further sensitivity analysis was performed which estimated standard errors using bootstrapping.

Criteria	Addressed in CS (Y/N)	ERG Comments
aggregate population by the outcome regression model. - For population reweighting methods, a robust sandwich estimator (as typical for MAIC) provides estimates of standard error which account for all sources of uncertainty. - Other techniques include bootstrapping and Bayesian methods.		
8. For an unanchored comparison, estimates of systematic error before and after population adjustment should be presented	N	No information
9. Present estimates for the appropriate target population using the shared effect modifier assumption if appropriate, or comment on the representativeness of the aggregate population to the true target population.	Y	It was reported that the MAIC results are only application to the population of the specific comparator trials (the trial providing the summary characteristics). Not all relevant covariates could be included in all analyses so "*the key assumption of the MAIC may not hold, and the results should be interpreted cautiously*"
10. In order to convey some clarity about the impact of any population adjustment, the standard indirect comparison estimate should be presented alongside the population-adjusted indirect comparison if an anchored comparison is formed; for an unanchored comparison, a crude unadjusted difference should be presented alongside the MAIC/STC estimate.	N	The statistical report stated that unanchored indirect comparisons were performed but the results were not reported, or presented in the company submission alongside the MAIC results.

Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Addendum for the amended follicular lymphoma only population

1. Purpose of the addendum

During the EMA regulatory process, the decision was made to withdraw the [redacted]
The expected license wording for R^2 is now: [redacted]

Given that the original submission focussed on a pooled FL+MZL population, on 4 October 2019, NICE requested the company provide a follicular lymphoma-only addendum to the original submission document.

Celgene has provided a FL-only analyses in response to ERG clarifications, however, a more detailed response for the FL-only population is provided here.

1.1. Summary

There exists a currently unmet need for new treatments in the setting of relapsed/refractory FL. The efficacy of lenalidomide + rituximab (R^2) has been demonstrated across the full spectrum of relapsed/refractory FL. In AUGMENT (FL-only), PFS, TTNLT, OS and ORR were significantly improved for R^2 versus R mono, while the combination demonstrated a predictable, manageable and acceptable safety profile. Results from the MAGNIFY study are supportive of the AUGMENT data and provide evidence of favourable efficacy of R^2 in the rituximab-refractory (R-refractory) population (a population excluded from the AUGMENT trial). Indirect comparisons of R^2 versus the current predominant treatment options in the UK for relapsed/refractory FL, R-CHOP/CVP, using data from HMRN still show a statistically significant clinical benefit in OS and TTNLT when considering the FL-only population. Results of the O-Benda comparison have not been discussed within the addendum as they were originally conducted for the FL-only population; therefore, results of this comparison presented in Document B are still relevant. Due to the limitations of the R-CHOP comparison using published evidence (Van Oers[1]) highlighted by the ERG, results for the FL-only populations have not been replicated here.

Overall, the available clinical evidence provides an appropriate base to inform the assessment of clinical effectiveness and cost-effectiveness of R^2 for patients with previously treated FL.

The deterministic base case results from the cost effectiveness model show that R^2 remains cost-effective in the FL-only population versus R-CHOP and R-CVP as well as versus the AUGMENT study comparator R mono at the £30,000 willingness-to-pay threshold. An array of sensitivity and scenario analyses were performed to test both parameter and structural uncertainty. The results from these analyses show that in the vast majority of cases R^2 remains a cost-effective use of NHS resources.

2. Clinical evidence for the FL population

2.1. Background

The decision problem, description of the technology and clinical care pathway can be found in Section B.1 of Doc B of the company submission (CS). This remains unchanged for the FL-only population, so we have not restated all that information in this addendum.

The description of the symptomatic burden, impact on carers and unmet need are provided in Section B.1.3 of Doc B CS, pages 15–17; this remains unchanged for the FL-only population. Considering its mechanism of action and the totality of the clinical data package[2-4], R^2 is anticipated to provide durable benefit for a wide range of relapsed/refractory FL patients, providing an alternative to current chemotherapy-based treatment options.

2.2. Summary of the FL trial data

Section B.2.6 of Doc B CS summarizes the trial data for the overall population. This section focusses specifically on the FL-only populations within AUGMENT and MAGNIFY.

Baseline patient characteristics

AUGMENT

For the FL-only patients in the intention-to-treat (ITT) population, baseline disease characteristics were similar to the overall population. Overall, 219 patients (74%) had Ann Arbor Stage III to IV disease; 100 patients (34%) had a FLIPI score ≥3; and 145 patients (49%) had high tumour burden per GELF criteria. The majority of enrolled FL patients were at second-line (53%) and the remainder were at third-line or greater (47%).

For FL patients, baseline demographics between the R^2 arm and the R mono arm were generally similar; more patients in the R^2 arm than in R mono arm were female (59% versus 46%), aged ≥65 years (42% versus 37%) had Ann Arbor Stage III to IV disease at enrolment (77% versus 72%), FLIPI score ≥3 (37% versus 31%), had an ECOG score of 1 or 2 (33% versus 29%). Baseline characteristics for FL patients in the ITT population are summarized in Table 1.

Table 1: Baseline demographic and disease characteristics in AUGMENT – FL ITT population

	FL ITT population (n=295)	
	R^2 (n=147)	**R mono (n=148)**
Male, n (%)	61 (41.5)	80 (54.1)
Median age, years (range)	62.0 (26.0–86.0)	61.0 (35.0–88.0)
Age distribution, n (%)		
<65	86 (58.5)	94 (63.5)
≥65	61 (41.5)	54 (36.5)
≥70	34 (23.1)	32 (21.6)
Race, white (%)	[redacted]	[redacted]
Histology (investigator review), n (%)		
FL	[redacted]	[redacted]
Grade 1	[redacted]	[redacted]
Grade 2	[redacted]	[redacted]
Grade 3a	[redacted]	[redacted]
MZL	N/A	N/A
MALT	N/A	N/A
Nodal	N/A	N/A
Splenic	N/A	N/A
Ann Arbor stage at enrolment, n (%)		
I	13 (8.8)	13 (8.8)
II	21 (14.3)	29 (19.6)
III	69 (46.9)	60 (40.5)
IV	44 (29.9)	46 (31.1)
FLIPI category (derived), n (%)		
Low (0,1)	[redacted]	[redacted]
Intermediate (2)	[redacted]	[redacted]
High (≥3)	[redacted]	[redacted]
Baseline ECOG score, n (%)		
0	99 (67.3)	105 (70.9)
1	47 (32.0)	42 (28.4)
2	[redacted]	[redacted]
LDH elevated, n (%)		
Yes	39 (26.5)	43 (29.1)
No	107 (72.8)	105 (70.9)
High tumour burden (GELF criteria)		
Yes	77 (52.4)	68 (45.9)
No	70 (47.6)	80 (54.1)

	FL ITT population (n=295)	
	R^2 (n=147)	**R mono (n=148)**
Prior anti-lymphoma regimens		
1	[redacted]	[redacted]
>1	[redacted]	[redacted]
Refractory to last prior regimen		
Yes	26 (17.7)	25 (16.9)
No	121 (82.3)	123 (83.1)
POD24[a], n (%)		
Yes	[redacted]	[redacted]
No	[redacted]	[redacted]

Key: ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; FLIPI, follicular lymphoma international prognostic index; GELF, Groupe d'Etude des Lymphomes Folliculaires; LDH, lactate dehydrogenase; MALT, mucosa associated lymphatic tissue; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo.
Notes: [a], POD24 is defined as disease progression or relapse within two years of initial diagnosis after the start of initial chemoimmunotherapy.
Source: Celgene, 2018.[5]

MAGNIFY

Table 2 presents the baseline demographic and disease characteristics for FL patients in the induction efficacy evaluable (IEE) population. For the FL IEE population, baseline disease characteristics were similar to the total IEE population; the majority ([redacted]%) were Stage IV at enrolment and there was approximately a 1:1 male to female ratio ([redacted]% male).

Overall, the baseline characteristics of FL patients in MAGNIFY represent a population with a poorer prognosis compared to the AUGMENT FL population, specifically for the following factors (MAGNIFY versus AUGMENT): median age ([redacted] versus [redacted] years), Ann Arbor Stage III–IV ([redacted]% versus 74%), ECOG 0 ([redacted]% versus 69%). Also note that the median number of previous treatments for FL patients in MAGNIFY was greater than those in AUGMENT (2 versus 1, respectively). Furthermore, [redacted]% of the enrolled FL patients in MAGNIFY were R-refractory; however, for AUGMENT, all patients were non-rituximab-refractory (non-R-refractory).

Table 2: Baseline demographic and disease characteristics in MAGNIFY – FL IEE population

	FL IEE population (n=247)
Male, n (%)	[redacted]
Median age, years (range)	[redacted]
Age distribution, n (%)	

	FL IEE population (n=247)
<65	[redacted]
≥65	[redacted]
Race, white, n (%)	[redacted]
Histology (investigator review), n (%)	
FL	[redacted]
Grade 1	[redacted]
Grade 2	[redacted]
Grade 3a	[redacted]
MZL	[redacted]
MALT (non-gastric)	[redacted]
Gastric MALT	[redacted]
Nodal	[redacted]
Splenic	[redacted]
Ann Arbor stage at enrolment, n (%)	
I	[redacted]
II	[redacted]
III	[redacted]
IV	[redacted]
ECOG score at enrolment, n (%)	
0	[redacted]
1	[redacted]
2	[redacted]
POD24[a], n (%)	
Yes	[redacted]
No	[redacted]
Prior rituximab-containing therapy, n (%)	
Yes	[redacted]
Prior rituximab-containing combination chemotherapy	[redacted]
Prior rituximab monotherapy	[redacted]
No	[redacted]
R-refractory, n (%)	
Yes	[redacted]
No	[redacted]
Chemoresistant, n (%)	
Yes	[redacted]
No	[redacted]
Chemotherapy eligible, n (%)	
Yes	[redacted]
No	[redacted]

	FL IEE population (n=247)
High tumour burden, n (%)	
Yes	[redacted]
No	[redacted]
Baseline bulky disease status[c], n (%)	
Yes	[redacted]
No	[redacted]
Missing	[redacted]

Key: ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; IEE, induction efficacy evaluable; MZL, marginal zone lymphoma; MALT, mucosa associated lymphatic tissue; R, rituximab.
Notes: [a], POD24 is defined as disease progression or relapse within two years of initial diagnosis after the start of initial chemoimmunotherapy; [b], for 29 of these patients, this was their only prior systemic regimen; [c], bulky disease is defined as a nodal or extranodal (except spleen) mass >7 cm in greater diameter or involvement of at least three nodal or extra-nodal sites (each with a diameter >3 cm).
Source: Celgene, 2018.[6]

Clinical effectiveness results

AUGMENT

Primary endpoint: Progression-free survival

For the total FL population, there were [redacted] IRC-assessed events, which remained unchanged following the application of EMA censoring. PFS improvements for FL patients were consistent with those of the total ITT population ([redacted]; Table 35). The median PFS was greater for FL patients in the R^2 arm ([redacted] months) compared with those in the R mono arm ([redacted] months). Furthermore, the PFS rate was greater for FL patients in the R^2 arm at both 1 year ([redacted] and 2 years ([redacted] compared with the R mono arm.

In a post hoc analysis, among patients randomized to the R^2 arm, the median PFS was longer than that observed following the previous anti-lymphoma treatment ([redacted] versus [redacted]), while the median PFS among patients in the R mono arm was shorter than that from the previous anti-lymphoma treatment ([redacted] versus [redacted]).[7] The observation in the control arm was consistent with established expectations that PFS will decrease with each successive treatment based on those currently available[8]; however, the longer PFS for patients treated with R^2 compared with that achieved in response to their previous treatment reverses that trend. This provides evidence for a clinically significant contribution from the novel mechanism of action of this combination that distinguishes it from currently available options.

Table 3 presents a summary of PFS in the FL patient population, and Figure 1 presents the Kaplan–Meier curve of PFS for the FL population.

Table 3: Progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in AUGMENT – FL population

	FL ITT population (n=295)	
	R^2 (n=147)	**R mono (n=148)**
Number of patients, n (%)		
With event		
Censored		
Median PFS[a] (95% CI) (months)		
PFS rate at 6 months (95% CI)		
PFS rate at 1 year (95% CI)		
PFS rate at 2 years (95% CI)		
p-value		
Hazard ratio (95% CI)		

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IRC, Independent Review Committee; ITT, intent-to-treat; IWGRC, International Working Group Response Criteria; NE, not estimable; PFS, progression-free survival; R^2, lenalidomide plus rituximab; R mono, rituximab + placebo.
Notes: [a], median estimate is from Kaplan–Meier analysis; [b], p-value from log-rank test; [c], from Cox proportional hazard model.
Source: Celgene, 2018.[9]

Figure 1: Kaplan–Meier curve of progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in AUGMENT – FL ITT population

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IRC, Independent Review Committee; IWGRC, International Working Group Criteria; ITT, intention-to-treat; KM, Kaplan–Meier; Len, lenalidomide; NE, not estimable; Pbo, placebo; Rit, rituximab.
Source: Celgene, 2018.[9]

Secondary endpoints

The secondary efficacy endpoints for AUGMENT included OS, response rates (ORR, CR rate, DOR, DOCR and DCRR), EFS and TTNLT. Table 4 presents a summary of the secondary endpoints for the FL population only. Table 36 presents a summary of the secondary endpoints for the FL population versus the total ITT population. Of note, the secondary endpoints for the FL only population were similar to those of the total ITT population.

Table 4: Summary of secondary endpoints in AUGMENT – FL ITT population

Endpoint	FL ITT population (n=295)	
	R^2 (n=147)	**R mono (n=148)**
Median OS, months (95% CI)[a]		
Hazard ratio (95% CI)		
ORR (CR+PR)		
95% CI[c]		
p-value		
CR rate		
95% CI[d]		

Endpoint	FL ITT population (n=295)	
	R^2 (n=147)	**R mono (n=148)**
p-value		
PR		
SD		
PD/ death		
No evidence of disease		
Unknown/ND/Missing		
Median TTNLT, months (95% CI)[a]		
TTNLT rate at 2 years, % (95% CI)		
Hazard ratio (95% CI)		
p-value		
Median EFS, months (95% CI)[a]		
Hazard ratio (95% CI)		
p-value		
DCRR, n (%)		
95% CI[i]		
p-value		
	R^2 (n=118)	**R mono (n=82)**
Median DOR, months (95% CI)[a]		
Hazard ratio (95% CI)[b]		
p-value[h]		
	R^2 (n=51)	**R mono (n=29)**
Median DOCR, months (95% CI)[a]		
Hazard ratio (95% CI)[e]		
p-value[b]		

Key: CI, confidence interval; CR, complete response; DCRR, durable complete response rate, DOCR, duration of complete response; DOR, duration of response; EFS, event-free survival; FL, follicular lymphoma; IRC, Independent Review Committee; ND, not done; NE, not estimable; ORR, overall response rate; OS, overall survival; PD, progressive disease; PR, partial response; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SD, stable disease; TTNLT, time to next anti-lymphoma treatment.
Notes: [a], median estimate is from Kaplan-Meier analysis; [b], from Cox proportional hazard model; [c], exact confidence interval for binomial distribution; [d], from Fisher-Exact test; [e], from log-rank test.
Source: Leonard, et al. 2019; Celgene, 2018.[2, 9]

Health-related quality of life

Health-related quality of life (HRQL) was also assessed as a secondary endpoint. The results are presented in Appendix P of the Doc B CS, where details of the primary HRQL endpoint for the FL only population are available. Of note, HRQL results were consistent between the ITT and FL population (see Doc B, Appendix P of CS).

Overall survival

Overall survival (OS) data is relatively immature, with few events at time of analysis. In the FL patient population, there were [redacted] deaths in the R^2 arm and [redacted] deaths in the R mono arm (median follow-up 28.3 months). Overall, there was a statistically significant relative reduction of 55% in the risk of death (HR [95% CI]: [redacted] [redacted]) for FL patients treated with R^2 compared with those treated with R mono. For both treatment arms, median OS was [redacted].

Figure 2 presents a KM curve of OS for the FL-only population of AUGMENT. Both KM curves overlapped up to 1-year post-randomization (Figure 2); and the OS rate at 1-year was similar between the two treatment arms ([redacted]% versus [redacted]%, respectively). At 2 years, however, the OS rate was greater in the R^2 arm ([redacted]%) compared with the R mono arm ([redacted]%).

Figure 2: Kaplan–Meier curve of overall survival in AUGMENT – FL ITT population

Key: CI, confidence interval; FL, follicular lymphoma; ITT, intention-to-treat; KM, Kaplan–Meier; Len, lenalidomide; NE, not estimable; Pbo, placebo; Rit, rituximab.
Source: Celgene, 2018.[9]

Response rates

For FL patients in the ITT population, ORR was significantly greater for those in the R^2 arm compared with those in the R mono arm ([redacted]% versus [redacted]%, respectively; p<[redacted]). The CR rate was also greater for FL patients in the R^2 arm compared with those in the R mono arm ([redacted] versus [redacted] respectively; p=[redacted]).

Time to next anti-lymphoma treatment

TTNLT offers a clinical endpoint with additional value beyond PFS when considering the applicability of clinical study data to actual clinical practice in this disease setting. PFS can be hard to interpret due to discrepant approaches to determining disease progression. Progression is proactively evaluated on-study through scheduled investigations, potentially prior to symptom development, while progression is more likely to be discovered reactively in real-world practice after a patient presents with symptoms. Subsequent treatments, however, are generally initiated only following symptomatic progression, making TTNLT an endpoint more readily comparable between clinical studies and real-world practice. The importance of the TTNLT endpoint in this indolent disease was highlighted by the committee for TA513.[10] This concept is discussed further in Section B.2.13 of Doc B CS.

In the FL ITT population, [redacted] patients ([redacted]%) in the R^2 arm and [redacted] patients ([redacted]%) in the R mono arm received subsequent anti-lymphoma treatment. For FL patients in the R^2 arm, median TTNLT was not estimable; however, for those in the R mono arm, the median TTNLT was [redacted] months (HR [95% CI]: [redacted] Furthermore, at 2 years, a greater proportion of FL patients in the R^2 arm had not received subsequent anti-lymphoma treatment compared to FL patients in the R mono arm ([redacted] versus [redacted] respectively). Figure 3 presents a summary of TTNLT for the FL population in AUGMENT.

An exploratory analysis of response rate to next anti-lymphoma treatment (RTNLT) is reported below.

Figure 3: Kaplan–Meier curve of time to next anti-lymphoma treatment in AUGMENT – FL ITT population

Key: CI, confidence interval; FL, follicular lymphoma; ITT, intention-to-treat; KM, Kaplan–Meier; Len, lenalidomide; NE, not estimable; Pbo, placebo; Rit, rituximab.

Source: Celgene, 2018.[9]

Event-free survival

For EFS, an event was defined as documented progression, relapse, initiation of new anti-lymphoma treatment or death. For the FL ITT population, patients in the R^2 arm were [redacted]% less likely to experience an event compared with those in the R mono arm (HR [95%]: [redacted]; p<[redacted]). Furthermore, at 1 year, the EFS rate was greater for FL patients in the R^2 arm ([redacted] compared with those in the R mono arm ([redacted]

Exploratory endpoints

A summary of the exploratory endpoints, including RTNLT, time to next chemotherapy treatment (TTNCT), CR rate (based on the 1999 WGRC), PFS on next anti-lymphoma treatment (PFS2) and patients with HT for the FL ITT population are presented in Table 5. Table 37 presents a summary of the exploratory endpoints for the FL population versus the total ITT population. Of note, the exploratory endpoints for the FL only population were similar to that of the total ITT population.

A greater proportion of FL patients receiving R^2 responded to subsequent treatment compared with patients receiving R mono (ORR: [redacted] CR: [redacted][11], lending weight to the hypothesis that the combination may re-sensitize patients to subsequent therapy through immune function enhancement.[3] Indeed, for FL patients, TTNCT was significantly improved with R^2 compared with R mono (p=[redacted]); however, median TTNCT was not estimable for either treatment arm. Furthermore, the TTNCT rate at 2 years was higher for FL patients in the R^2 arm ([redacted] than for the R mono arm ([redacted]

Table 5: Summary of exploratory endpoints in AUGMENT – FL ITT population

Endpoint	FL ITT population (n=295)	
	R^2 (n=147)	R mono (n=148)
Median TTNCT, months (95% CI)[a]	[redacted]	[redacted]
TTNCT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	
Best response by IRC per 1999 IWGRC		
ORR (CR/CRu + PR), n (% [95% CI][d])	[redacted]	[redacted]
p-value	[redacted]	
CR rate (CR/CRu), n (% [95% CI][d])	[redacted]	[redacted]
p-value	[redacted]	
CR, n (%)	[redacted]	[redacted]
CRu, n (%)	[redacted]	[redacted]
PR, n (%)	[redacted]	[redacted]
Median PFS2, months (95% CI)[a]	[redacted]	[redacted]

Endpoint	FL ITT population (n=295)	
	R^2 (n=147)	R mono (n=148)
Hazard ratio (95% CI)		
p-value		
HT, n (% [95% CI][f])		
	R^2 (n=37)	R mono (n=70)
RTNLT		
ORR, n (% [95% CI][d])		
p-value		
CR, n (% [95% CI][d])		
p-value		

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; HT, histological transformation; ITT, intention-to-treat; IWGRC, International Working Group Response Criteria; NE, not estimable; ORR, overall response rate; PFS2, PFS on next anti-lymphoma treatment; PR, partial response; R^2, rituximab plus lenalidomide; R mono, rituximab plus placebo; RTNLT, response rate to next anti-lymphoma treatment; TTNCT, time to next anti-lymphoma chemotherapy treatment.
Notes: [a], median estimate is from Kaplan–Meier analysis; [b], from Cox proportional hazard model. [c], p-value from log-rank test; [d], exact confidence interval for binomial distribution; [e], p-value obtained from Fisher-Exact test; [f], 95% CI is based on the Clopper-Pearson exact method.
Source: Leonard, et al. 2019[2]; Celgene, 2018.[9, 11]

MAGNIFY

The MAGNIFY study provides supportive data for the efficacy of R^2 in the R-refractory population, a patient population excluded from the pivotal trial, AUGMENT.

The data discussed in this section have been taken from the interim analysis (data cut-off 10 August 2018). Primary efficacy analyses were conducted in the IEE population. Efficacy analyses conducted in the induction intention-to-treat (IITT) population (secondary endpoints) were not considered of primary importance and are presented in Appendix O (for which details of FL-only are available). Exploratory efficacy analyses, such as PFS, were conducted in the induction safety population.

Primary endpoint (interim analysis): Overall response rate

Response rates were based on best response and assessed using the modified 1999 IWGRC. In the IEE population, the ORR for FL patients was 75%. The CR rate (secondary endpoint) for FL patients in the IEE population was 46%.

Table 6 presents a summary of response rates by best response for the IEE FL population in the induction phase of MAGNIFY. Table 38 presents a summary of the response rates by best response for the FL population versus the total IEE population. Of note, the response rates for the FL only population were similar to that of the total IEE population.

Table 6: Response rate by best response per 1999 IWGRC in the induction phase in MAGNIFY – FL IEE population

Best response in induction phase	FL IEE population (n=247)
Number of patients, n (%)	
ORR (CR+CRu+PR), n (% [95% CI[a]])	184 (74.5 [68.6, 79.8])
CR	[redacted]
CRu	[redacted]
CR rate (CR+CRu), n (% [95% CI][a])	114 (46.2 [39.8, 52.6])
PR	[redacted]
SD	[redacted]
PD	[redacted]
Death w/o tumour assessment	[redacted]

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; FL, follicular lymphoma; ORR, overall response rate; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: [a], 95% CI based on the Clopper–Pearson exact method.
Source: Celgene, 2018.[4]

Secondary endpoints (interim analysis)

A summary of the secondary efficacy analyses conducted in the IITT population for MAGNIFY, including CR rate, DOR, DOCR and TTR are presented in Appendix O2, where details for the FL only population are available.

Exploratory endpoint (interim analysis): Progression-free survival

PFS analyses were conducted in the induction safety population, and were based on investigator assessment, using the modified 1999 IWGRC. As with AUGMENT, EMA censoring rules were applied to the analyses.

In the FL population, there were a total of 86 investigator-assessed events. The PFS rate at 1 year was [redacted] For the R-refractory and non-R-refractory FL populations, the PFS rate at 1 year was [redacted]% (95% CI: [redacted]) and [redacted]% (95% CI: [redacted]), respectively, demonstrating favourable efficacy for R^2 in both the R-refractory and non-R-refractory populations. A summary of PFS for the FL population in total, and split by R-refractory status, is presented in Table 7 and the KM curve is presented in Figure 4. Table 40 presents a summary of the PFS for the FL population versus the total induction safety population. Of note, the PFS results

for the FL only population were consistent with that of the total induction safety population.

Table 7: Progression-free survival by investigator assessment per 1999 IWGRC with censoring rules based on EMA guidance in MAGNIFY split by R-refractory status – FL induction safety population

	R-refractory FL patients (n=97)	Non-R-refractory FL patients (n=189)	Total FL induction safety population (n=286)
Number of patients, n (%)			
With event, n (%)	[redacted]	[redacted]	[redacted]
Censored, n (%)	[redacted]	[redacted]	[redacted]
PFS rate at 1 year, % (95% CI)[a]	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IWGRC, International Working Group Response Criteria; R-refractory, rituximab-refractory; PFS, progression-free survival.
Notes: [a], Statistics obtained from Kaplan–Meier method. 95% CI is based on Greenwood formula.
Source: Celgene, 2018.[12]

Figure 4: Kaplan–Meier curve of progression-free survival based on EMA guidance in MAGNIFY split by R-refractory status – FL induction safety population

Key: EMA, European Medicines Agency; FL, follicular lymphoma; R-refractory, rituximab-refractory.
Source: Celgene, 2018.[13]

Safety

AUGMENT

Data from this section are taken from the 22 June 2018 database cut-off; safety analyses were conducted in the safety population.

The proportion of TEAEs and most common TEAEs observed in the FL safety population were consistent with those seen in the total safety population presented in Table 40 and 41.

Summary of treatment-emergent adverse events

For FL patients in the safety population, TEAEs were reported in [redacted] patients ([redacted]%) in the R^2 arm and [redacted]patients ([redacted]%) in the R mono arm. More FL patients in the R^2 arm ([redacted]%) experienced a Grade 3 or 4 TEAE compared with FL patients in the R mono arm ([redacted]%), and only [redacted] FL patient in each treatment arm reported a Grade 5 TEAE. Additionally, a greater proportion of FL patients reported serious adverse events in the R^2 arm ([redacted]%) compared with those in the R mono arm ([redacted]%).

A greater proportion of FL patients experienced TEAEs leading to dose reductions of lenalidomide or placebo in the R^2 arm than in the R mono arm ([redacted]% versus [redacted]%, respectively). TEAEs leading to dose interruptions of lenalidomide or placebo were also more frequent for FL patients in the R^2 arm compared with the R mono arm ([redacted]% versus [redacted]%, respectively). Similarly, a greater proportion of FL patients in the R^2 arm ([redacted]%) had at least one TEAE leading to dose interruption of rituximab compared with those in the R mono arm ([redacted]%). Furthermore, the proportion of FL patients who experienced at least one TEAE leading to discontinuation of therapy was slightly higher in the R^2 arm compared with the R mono arm for both lenalidomide or placebo therapy ([redacted]% versus [redacted]%), and for rituximab therapy ([redacted]% versus [redacted]%).

Table 8 presents a summary of the TEAEs for FL patients in the safety population during AUGMENT.

Table 8: Summary of treatment-emergent adverse events in AUGMENT – FL safety population

	FL safety population (n=294)	
	R^2 (n=146)	**R mono (n=148)**
Number of patients (%)		
Any TEAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 3–4 TEAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]

Grade 5 TEAE	[redacted]	[redacted]
Any SAE	[redacted]	[redacted]
Len/Pbo related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of R	[redacted]	[redacted]
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to discontinuation of R	[redacted]	[redacted]

Key: FL, follicular lymphoma; Len, lenalidomide; Pbo, placebo; R, rituximab; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[9]

Most common treatment-emergent adverse events

For FL patients in the safety population, TEAEs that occurred more frequently (≥10% difference) in the R^2 arm than the R mono arm included the following: neutropenia ([redacted]% versus [redacted]%), diarrhoea ([redacted]% versus [redacted]%), constipation ([redacted]% versus [redacted]%), cough ([redacted]% versus [redacted]%), upper respiratory tract infection ([redacted]% versus [redacted]%) and leukopenia ([redacted]% versus [redacted]%).

The difference in the number of Grade 3 or 4 TEAEs between treatment arms (shown in Table 8) was largely driven by Grade 3 or 4 events of neutropenia and leukopenia. Neutropenia occurred in [redacted] FL patients ([redacted]%) in the R^2 arm compared with [redacted] FL patients ([redacted]%) in the R mono arm, and leukopenia occurred in [redacted] FL patients ([redacted]%) in the R^2 arm compared with [redacted]patients ([redacted]%) in the R mono arm.

Table 9 presents the most common TEAEs, occurring in more than 10% of FL patients.

Table 9: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT – FL safety population

	FL safety population (n=294)	
	R^2 (n=146)	**R mono (n=148)**
Number of patients (%)		
Blood and lymphatic system disorders	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]

	FL safety population (n=294)	
	R^2 (n=146)	R mono (n=148)
Thrombocytopenia	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Constipation	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
URTI	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Investigations	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]
Tumour flare	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]

	FL safety population (n=294)	
	R^2 (n=146)	**R mono (n=148)**
Cardiac disorders	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]

Key: FL, follicular lymphoma; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; URTI, upper respiratory tract infection.
Source: Celgene, 2018.[9]

MAGNIFY

All AEs were assessed starting after the patient signed the informed consent and until 28 days after they discontinued taking the study drug. AEs that lead to patients discontinuing the study were followed until the problem was resolved or stabilized.

Summary of treatment-emergent adverse events

Overall, TEAEs were reported in [redacted] FL patients ([redacted]%) in the induction safety population. A total of [redacted] FL patients ([redacted]%) reported a Grade 3 or 4 TEAE and only [redacted] FL patients ([redacted]%) reported a Grade 5 TEAE. Serious TEAEs were reported in [redacted] FL patients ([redacted]%).

In total, [redacted] FL patients ([redacted]%) reported a TEAE leading to a dose reduction of lenalidomide. A greater proportion of FL patients experienced a TEAE leading to a dose interruption of lenalidomide compared with those leading to a dose interruption of rituximab ([redacted]% versus [redacted]%, respectively). Similarly, more FL patients reported a TEAE leading to early discontinuation of lenalidomide compared with early discontinuation of rituximab ([redacted]% versus [redacted]%, respectively).

Table 10 presents a summary of the TEAEs for FL patients in the induction safety population during MAGNIFY. A comparison between the FL induction safety population and the total induction safety population is presented in Table 42. Of note, the proportion of TEAEs observed in the FL induction safety population were consistent with those seen in the total safety population.

Table 10: Summary of treatment-emergent adverse events in MAGNIFY – FL induction safety population

	FL induction safety population (n=286)
Number of patients (%)	
Any TEAE	[redacted]
Len related	[redacted]
R related	[redacted]
Grade 3–4 TEAE	[redacted]
Len related	[redacted]
R related	[redacted]
Grade 5 TEAE	[redacted]

	FL induction safety population (n=286)
Any SAE	[redacted]
Len related	[redacted]
R related	[redacted]
Any TEAE leading to dose reduction of Len	[redacted]
Any TEAE leading to dose interruption of Len	[redacted]
Any TEAE leading to dose interruption of R	[redacted]
Any TEAE leading to discontinuation of Len	[redacted]
Any TEAE leading to discontinuation of R	[redacted]

Key: FL, follicular lymphoma; Len, lenalidomide; R, rituximab; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[9]

Most common treatment-emergent adverse events

In the induction safety population, the most commonly reported TEAEs (≥10% of patients) included: fatigue ([redacted]%), neutropenia ([redacted]%), diarrhoea ([redacted]%), nausea ([redacted]%), constipation ([redacted]%), pruritus ([redacted]%) and cough ([redacted]%).

Table 11 presents the most common TEAEs, occurring in more than 10% of FL patients. A comparison between the FL induction safety population and the total induction safety population is presented in Table 43. Of note, the most common TEAEs observed in the FL induction safety population were consistent with those seen in the total induction safety population.

Table 11: Most common treatment-emergent adverse events reported in ≥10% of patients by system organ class in MAGNIFY – FL induction safety population

	FL induction safety population (n=286)
Gastrointestinal disorders	[redacted]
Diarrhoea	[redacted]
Nausea	[redacted]
Constipation	[redacted]
Abdominal pain	[redacted]
Vomiting	[redacted]
General disorders and administration site conditions	[redacted]
Fatigue	[redacted]
Oedema peripheral	[redacted]

	FL induction safety population (n=286)
Pyrexia	[redacted]
Skin and subcutaneous tissue disorders	[redacted]
Pruritus	[redacted]
Rash maculo-papular	[redacted]
Rash	[redacted]
Blood and lymphatic system disorders	[redacted]
Neutropenia	[redacted]
Anaemia	[redacted]
Thrombocytopenia	[redacted]
Leukopenia	[redacted]
Infections and infestations	[redacted]
URTI	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]
Muscle spasms	[redacted]
Arthralgia	[redacted]
Back pain	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]
Cough	[redacted]
Dyspnoea	[redacted]
Nervous system disorders	[redacted]
Headache	[redacted]
Dizziness	[redacted]
Metabolism and nutrition disorders	[redacted]
Decreased appetite	[redacted]
Hypokalaemia	[redacted]
Injury, poisoning and procedural complications	[redacted]
Infusion-related reaction	[redacted]

Key: URTI, upper respiratory tract infection.
Source: Celgene, 2018.[14]

2.3. MAIC using the FL data

No head-to-head data are available for R^2 versus any of the comparators of interest to this submission; only R mono was compared with R^2 within the AUGMENT RCT (Sections B.2.2 – B.2.7). As such, for all relevant comparators, indirect treatment comparisons (ITCs) were attempted using:

- Published evidence identified from the SLR (Section B.2.1) for comparisons to:
  - R-CHOP

  - O-Benda
- Evidence from a UK real-world evidence (RWE) registry, the HMRN database, for comparison to:
  - R-CHOP/R-CVP

ITCs with data from published evidence

Non-rituximab refractory population versus R-CHOP

The ITC comparing R^2 to R-CHOP using evidence from AUGMENT for R^2 (both FL and MZL subjects) and from the Van Oers 2011 study for R-CHOP were presented in Section B.2.9 of the original submission under the heading *ITCs with data from published evidence*. A key limitation to the analyses, highlighted in the submission, was:

(1) the incomparability of populations in terms of prior rituximab treatment – a key treatment effect modifier
    a. The R-CHOP Van Oers publication was 100% rituximab naïve, a population not considered relevant to UK clinical practice

This comparison was presented as a scenario in the original submission, with HMRN data being used in the base-case to overcome these limitations. The ERG has discounted the comparison with the Van Oers publication as a scenario due to the population not being relevant to UK clinical practice, as such we have not conducted the additional work required to provide the FL-only analyses for this addendum.

Rituximab-refractory population versus O-Benda

The analyses of R^2 versus O-Benda was already conducted in a FL-only population for the original submission. O-Benda is the only indicated and NICE-recommended treatment recommended for R-refractory FL patients; therefore, no additional analyses were required for this addendum. Methods and results for this comparison are presented in B.2.9 of the original submission under the heading *ITCs with data from published evidence.*

The ERG discounted O-Benda as a relevant comparator. However, given that O-Benda is the only NICE-recommended therapy for this population and clinical experts indicate that it is the predominant treatment option for R-refractory FL patients, we still believe it is a valid comparator for the R-refractory population despite being on the CDF.

ITCs with data from HMRN

HMRN data were used to help address the limitations associated with the ITCs using published evidence, as described above.

This section presents the updated results of the ITC of R^2 (informed by evidence from AUGMENT) and R-CHOP/R-CVP (informed by evidence from HMRN) where the R^2 evidence is now informed by the subgroup of AUGMENT subjects with FL (subjects with MZL have been removed from the data). Of the 178 subjects randomized to the AUGMENT R^2 arm, 147 had FL. All subjects in the HMRN data had FL.

Methods

The methodology used for these ITCs of R^2 and R-CHOP/R-CVP in the FL population was consistent with the ITC performed in the FL and MZL population which is detailed in Section B.2.9 of the original submission under the heading *ITCs with data from HMRN*. MAICs were performed using the methodology as described in Signorovitch et al., 2010[15], 2012[16] and referenced in the NICE Decision Support Unit (DSU) and Technical Support Document (TSD) 18.[17] The outcomes of interest were OS, TTNLT and PFS.

Matching variables

The baseline characteristics that were commonly collected by the HMRN and the AUGMENT study are presented in Table 12.

Table 12: Covariates collected commonly across AUGMENT and HMRN datasets

Data source	**HMRN**	**AUGMENT (FL subgroup)**
Treatment	**R-CVP/R-CHOP (2L+ population)**	**R^2**
N	[redacted]	147
Age (years):		
Median	[redacted]	62
Range	[redacted]	26–86
n (%) Age >=60yrs	[redacted]	86 (58.5%)
n (%) Age >=65yrs	[redacted]	61 (41.5%)
Sex, n, %		
n (%) Males	[redacted]	61 (41.5)
n (%) Females	[redacted]	86 (58.5)
Number of prior systemic anti-lymphoma regimens:		
n (%) 1	[redacted]	[redacted]
n (%) 2	[redacted]	[redacted]
n (%) ≥ 3	[redacted]	[redacted]
Prior rituximab treatment, n (%)	[redacted]	125 (85.0)
POD24[a], n (%)	[redacted]	[redacted]

Data source	HMRN	AUGMENT (FL subgroup)
Fully staged, n (%)	[redacted]	NA
Bone marrow involved, n (%)	[redacted]	20 (13.6)
Nodal sites		
n (%) ≤4	[redacted]	91 (61.9)
n (%) >4	[redacted]	56 (38.1)
Bulky disease	[redacted]	39 (26.5)
Stage		
n (%) I	[redacted]	13 (8.8)
n (%) II	[redacted]	21 (14.3)
n (%) III	[redacted]	69 (46.9)
n (%) IV	[redacted]	44 (29.9)

Key: HMRN, Haematological Malignancy Research Network; NA, not applicable; R^2, rituximab plus lenalidomide; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.
Notes: [a], POD24 is defined as disease progression or relapse within two years of initial diagnosis after the start of initial chemoimmunotherapy.

The list of potential effect modifiers/prognostic variables discussed in the original submission in the context of the ITC with published data (Section B.2.9), was used to identify the matching variables. A key treatment effect modifier/prognostic factor that was not collected by the HMRN was the FLIPI risk category; however, three of the four FLIPI components were collected at the relapsed/refractory baseline (only LDH was not collected). Matching was therefore performed for the following variables:

- Age ≥60 years (FLIPI component)
- Ann Arbor Stage III–IV (FLIPI component)
- Nodal sites >4 (FLIPI component)
- Prior rituximab treatment
- Prior lines of therapy (1 versus 2 versus >2)
- POD24 status
- Bone marrow involvement

Only 34 of the 63 R-CVP/R-CHOP-treated patients were fully staged, thereby providing information on nodal sites and Ann Arbour stage. For matching, it was therefore assumed that the missing data was equally distributed across the categories of these two factors.

Results

Following the matching procedure, the weighted baseline characteristics for the subgroup of AUGMENT FL patients were compared with the R-CVP/R-CHOP HMRN population. The MAIC has led to re-weighted AUGMENT covariates that are the same as in the R-CVP/R-CHOP HMRN population (matching results are presented in the Appendix, Section 5.3).

Table 13 provides a summary of the number of patients and events in the MAIC analyses for OS, TTNLT and PFS.

Table 13: Number of patients and events in the MAIC analyses for the R^2 (AUGMENT FL subjects) versus R-CVP/R-CHOP (HMRN) comparison

Treatment	N	Events	Median survival
OS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
PFS			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]
TTNLT			
R^2	[redacted]	[redacted]	[redacted]
R-CVP/R-CHOP	[redacted]	[redacted]	[redacted]

Key: HMRN, Haematological Malignancy Research Network; MAIC, matching-adjusted indirect comparison; NA, not applicable; OS, overall survival; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; TTNLT, time-to-next anti-lymphoma therapy.

The KM curves comparing R^2 (weighted and unweighted) to R-CVP/R-CHOP for OS, PFS and TTNLT are presented in Figure 5, Figure 6 and Figure 7, respectively. For all three endpoints, the weighting does not considerably alter the R^2 KM curves. The curves suggest R^2 has survival benefit compared to R-CVP/R-CHOP and a benefit for TTNLT; however, PFS looks similar for the two treatments. This latter observation may be a function of the PFS assessments for the AUGMENT and HMRN populations being conducted differently (proactively on-study and reactively in the real-world setting) as described below and in section B.2.13 of the original submission. The more like-for-like TTNLT comparison provides supporting evidence for the OS benefit observed with R^2.

Figure 5: Kaplan–Meier curve for the MAIC analysis of OS comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population

Key: MAIC, matching-adjusted indirect comparison; OS, overall survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Figure 6: Kaplan–Meier curve for the MAIC analysis of PFS comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population

Key: MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

Figure 7: Kaplan–Meier curve for the MAIC analysis of TTNLT comparing R^2 with R-CVP/R-CHOP in the non-R-refractory population

Key: MAIC, matching-adjusted indirect comparison; R, rituximab; R^2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; TTNLT, time-to-next anti-lymphoma treatment.

HRs from the Cox Proportional-Hazard models comparing R^2 and R-CVP/R-CHOP are presented in Table 14. Log-cumulative hazard plots for these three outcomes (see Section 3.1) suggest that the proportional hazards assumption may not hold. However, the HRs and corresponding CIs support the interpretation of the KM curves: R^2 has survival benefit compared to R-CVP/R-CHOP and a benefit for TTNLT, with modest PFS improvement.

Table 14: Results from Cox Proportional Hazard models comparing R² and R-CVP/R-CHOP

Outcome	R², adjusted N	R-CVP/R-CHOP N	HR (95% CI)[a]
OS	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]
TTNLT/death	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival; PFS, progression-free survival; R², rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone TTNLT, time to next anti-lymphoma treatment.
Notes: [a], bootstrapped CI.

2.4. Summary of the FL clinical evidence

Existing treatment options leave a substantial unmet need in the management of relapsed/refractory FL. Patients with FL who are non-R-refractory are generally offered R-CHOP and R-CVP, whilst those who are R-refractory have more limited options but most likely receive O-Benda, via the CDF. Existing chemotherapy-based treatment options do not fully address the immune dysfunction associated with FL and are associated with well-documented toxicities (e.g., vomiting, peripheral neuropathy, alopecia).[18] There is a need for novel treatment approaches, such as R², that target the immune microenvironment and offer alternatives to repeated exposure to chemotherapy-based interventions, with benefit derived from changes in both mechanism of action and toxicity profile.[19]

Overall, the clinical evidence available provides an appropriate base to inform the assessment of clinical effectiveness and cost-effectiveness of R² for the treatment of previously treated FL. These are largely in line with the presentation in the original CS Doc B. A full interpretation of the clinical evidence for R² from the AUGMENT and MAGNIFY studies is presented in Section B.2.13 of Doc B of the CS.

In AUGMENT (non-R-refractory FL patients only), PFS was significantly improved for R² versus R mono ([redacted] versus [redacted] months; HR: [redacted] [95% CI: [redacted]]; p<[redacted]), an outcome consistently observed across a wide range of pre-defined subgroups. TTNLT, an endpoint with greater clinical significance to patients than PFS, and one more readily applicable to comparison with real-world data, was significantly improved in FL patients treated with R² compared to those treated with R mono (HR: [redacted] [95% CI [redacted]] p<[redacted]). Furthermore, OS showed a statistically significant relative reduction of [redacted]% in the risk of death (HR [redacted] [95% CI [redacted]]) for FL patients treated with R² versus R mono.

For FL patients, ORR was significantly improved with R² versus R mono ([redacted]% versus [redacted]%; p<[redacted]). Additionally, FL patients in the R² arm were more likely to respond to subsequent anti-lymphoma treatment than those in the R mono arm

(ORR: [redacted]% versus [redacted]%; CR: [redacted]% versus [redacted]%)[11], lending weight to the hypothesis that R^2 may re-sensitize patients to subsequent treatment.[3]

Results from interim analysis of the MAGNIFY study are supportive of the AUGMENT data. The ORR for the FL IEE population was [redacted]% (95% CI: [redacted]), and the PFS rate at 1 year for FL patients in the induction safety population was [redacted]% (95% CI: [redacted]). Results from MAGNIFY also provide additional data for the R-refractory population, a population excluded from AUGMENT. MAGNIFY demonstrated favourable outcomes for R^2 in both R-refractory and non-R-refractory FL patients.

Overall, across both the AUGMENT and MAGNIFY studies, R^2 demonstrated a predictable, manageable and acceptable safety profile, consistent with those of the individual components of R^2 (lenalidomide and rituximab) with no new safety signals observed for R^2. Neutropenia, a known adverse event associated with lenalidomide, was well managed for patients receiving R^2 through dose modifications, with few FL patients requiring treatment discontinuation ([redacted]% discontinuing lenalidomide, and [redacted]% discontinuing rituximab). Further supportive safety data are available from a head-to-head study comparing R-chemo and R^2 in untreated FL patients, in which R^2 was associated with a lower incidence of TEAEs, manageable AEs, and no unexpected toxicities.[18] When AE profiles are compared between R^2 and components of existing treatment regimens for FL, R^2 exhibits a different AE profile including lower risk of Grade 3-4 neutropenia and febrile neutropenia.[18]

In the AUGMENT study, R^2 was found to have no detrimental impact on HRQL in patients with previously-treated FL despite the temporary impact of symptoms (e.g. fatigue, constipation, appetite loss, and diarrhoea) during the treatment period. In the context of the significant extension of PFS provided by R^2 compared with rituximab alone, the AE profile for R^2 appears to be outweighed by its clinical benefit.

In the absence of head-to-head comparisons of R^2 with relevant comparators, indirect comparisons were conducted. For the FL-only addendum, only the comparison of R^2 versus R-CHOP/R-CVP using RWE from the UK HMRN database has been presented. The comparison for R^2 versus O-Benda was already conducted in an FL-only population, and these results are provided in Section B.2.9, pages 79-82, of Doc B CS. The scenario analyses of R^2 versus R-CHOP using the Van Oers publication was not conducted for this addendum due to the limitations previously highlighted.

Similar to the overall population, results for the FL-only population derived from the Cox proportional hazard models using the HMRN data reported benefits in OS ([redacted]) and TTNLT ([redacted] for R^2 compared to R-CVP/R-CHOP. The benefit for PFS was more modest ([redacted]

The availability of TTNLT data from HMRN is important, as it provides reassurance regarding the OS benefit described for R^2 compared with R-CHOP/R-CVP given the more modest improvement in PFS. In contrast to PFS, TTNLT represents an

outcome triggered in a similar fashion in both clinical studies and real-world data, thus providing a more like-for-like comparison than PFS. TTNLT has also been recognized as a clinical endpoint that is more relevant to patients in clinical practice than PFS (TA513). This is discussed further in Section B.2.13 of Doc B of the CS.

3. Economic evidence for the FL only population

3.1. Model parameters in the FL only population

Clinical parameters for the FL model

For R^2 in comparison with R-CHOP/R-CVP, data for R^2 from AUGMENT in the FL only population were matched to pooled data from HMRN for R-CHOP and R-CVP. The matching methodology used was equivalent to that previously described in Section B.2.9 of the company submission for the FL + MZL population. Parametric curves were then fitted to the weighted R^2 KM's for OS, PFS, TTNLT and ToT.

As per the request in clarification question B3, the FL only analysis has also been conducted for the R^2 versus R-mono comparison based on the head-to-head AUGMENT trial. However, it is worth noting that R-mono is not considered a relevant comparator because it is rarely used in the relapsed/refractory setting in the UK. R-mono is therefore only presented as it provides a within trial comparison using head-to-head trial data.

As noted above the comparison of R^2 to O-Benda already uses the FL only population. For the R-refractory population all economic evidence presented in Section B.3 of Document B for the R^2 versus O-Benda comparison is still relevant for this submission. It has therefore not been included within this addendum.

Treatment effect

The treatment waning effect assumes that any treatment effect of R^2 only last up to 5 years. This is consistent with the FL+MZL population and details of the justifications for this time point are described in Document B Section B.3.3.

Overall survival

<u>R^2 versus R-CHOP/R-CVP</u>

For R^2 in comparison to R-CHOP/CVP, data from AUGMENT (FL only) were matched to pooled data from HMRN (see Section 2.3). The OS KMs for the R^2 (re-weighted) and R-CHOP/CVP are shown in Figure 8.

Figure 8: OS – R^2 (FL only – re-weighted) versus R-CHOP/CVP

Figure 9 presents the corresponding log-cumulative hazard versus time plot. The lines meet at approximately 300 days and then diverge thereafter, suggesting that the proportional hazards assumption is not appropriate. Additionally, as described in the NICE DSU TSD 14, it is unnecessary to rely upon the proportional hazards assumption when IPD are available.[20] As such, stratified statistical models have been used.

Figure 9: OS: log-cumulative hazard versus time plot – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Table 15 presents the AIC and BIC fit statistics for each distribution. The AIC/BIC for R-CHOP/R-CVP OS were also previously reported in Table 28 of Document B (Section B.3.3. page 138). For R^2, exponential or log-normal are statistically the best fitting; for R-CHOP/R-CVP, these are exponential or log-logistic with the difference between the fit statistics being minimal.

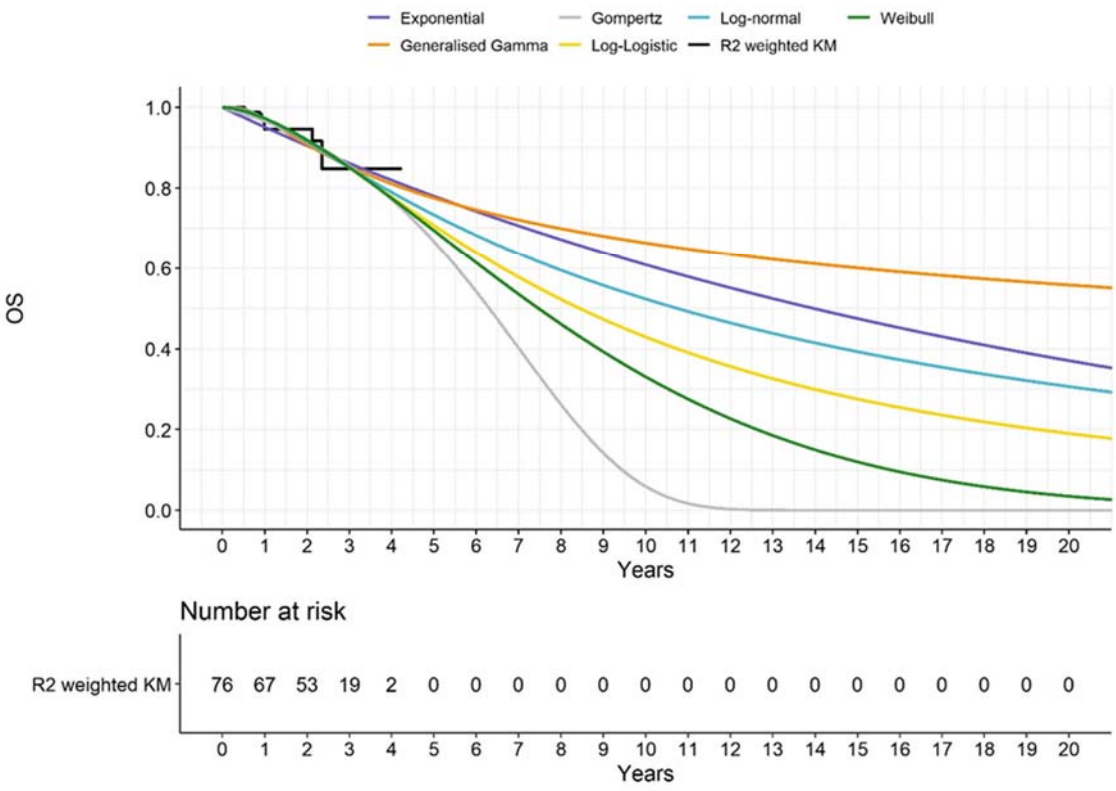
Table 15: OS: AIC and BIC – R^2 (FL only) versus R-CHOP/R-CVP

Distribution	R^2 (FL only, weighted)		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	177.67	**180.66**	598.43	**600.58**
Generalized gamma	178.71	187.68	600.05	606.48
Gompertz	178.68	184.66	598.05	602.34
Log-logistic	177.52	183.50	**598.03**	602.32
Log-normal	**177.03**	183.01	598.18	602.47
Weibull	**177.65**	**183.63**	**599.59**	**603.88**

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: Bold = statistically best fit; **Purple** = selected for base case.

Figure 10 presents the parametric distributions fit to the R^2 KM data (FL only) for OS. Figure 11 presents the OS parametric distributions fit to R-CHOP/R-CVP, as also previously shown in Figure 25 of Document B (Section B.3.3. page 140).

Figure 10: OS parametric curves for R^2 (FL only, weighted) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 11: OS parametric curves for R-CHOP/R-CVP in the non-rituximab refractory population (Document B Section B.3.3. Figure 25)

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Clinical expert opinion was sought to judge the appropriateness of model extrapolations for R-CHOP/CVP OS; the bottom 3 curves (exponential, Weibull and log-logistic) were considered more plausible than the top 3 (generalized gamma, Gompertz, and log-normal). All three were thought to be plausible with a preference expressed for the middle curve (Weibull), as the 20-year OS estimate seems reasonable and unlikely to attract challenge. In addition, the AIC/BIC for Weibull suggests it is a reasonable fit, and this distribution was considered appropriate for the relapsed/refractory setting in TA137 (the same population as this appraisal).

Given Weibull demonstrated preferred clinical plausibility, reasonable statistical fit and consistency with a previous NICE appraisal in the same population (TA137), this distribution was used for the base case for R-CHOP/CVP. The same distribution has also been selected for R^2 (FL only) to estimate OS at base case. Nonetheless, as clinical expert opinion also considered the exponential and log-logistic curves for R-CHOP/CVP OS to be plausible, these curve options for both treatment arms are

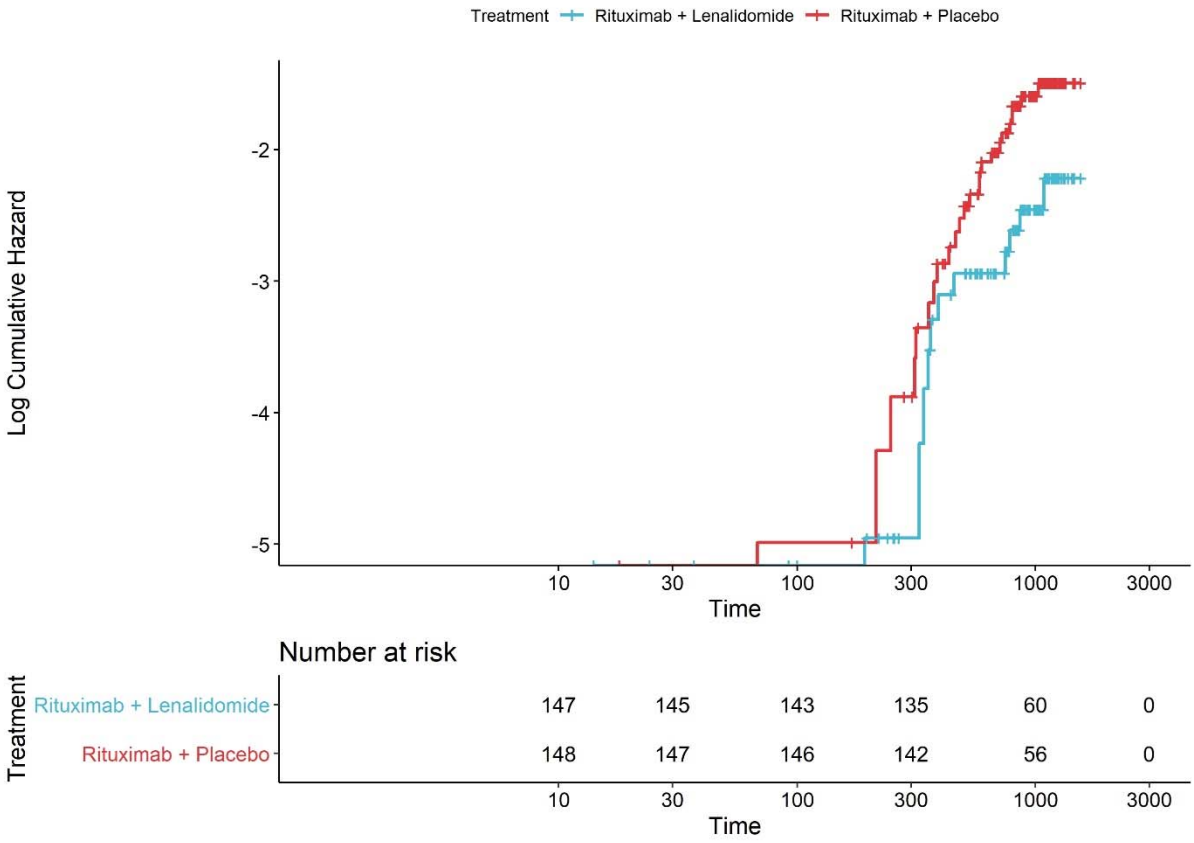
further explored in scenario analysis. The log-normal distribution for R^2 (FL only) OS was also explored in scenario analysis as it has the best statistical fit according to the AIC statistic.

The curves are adjusted to use general population mortality as described in Document B, Section B.3.3.

R^2 versus R-mono

For R^2 versus R-mono, parametric curves were fit directly to the AUGMENT FL only trial data. Figure 12 presents the log-cumulative hazard plot for both R^2 and R-mono. This shows that the curves come together at around 300 days, but proportional hazards could be considered reasonable. However, based on the availability of IPD stratified models have been used as suggested by NICE DSU TSD 14.[20]

Figure 12: OS: log-cumulative hazard versus time plot – R^2 versus R-mono

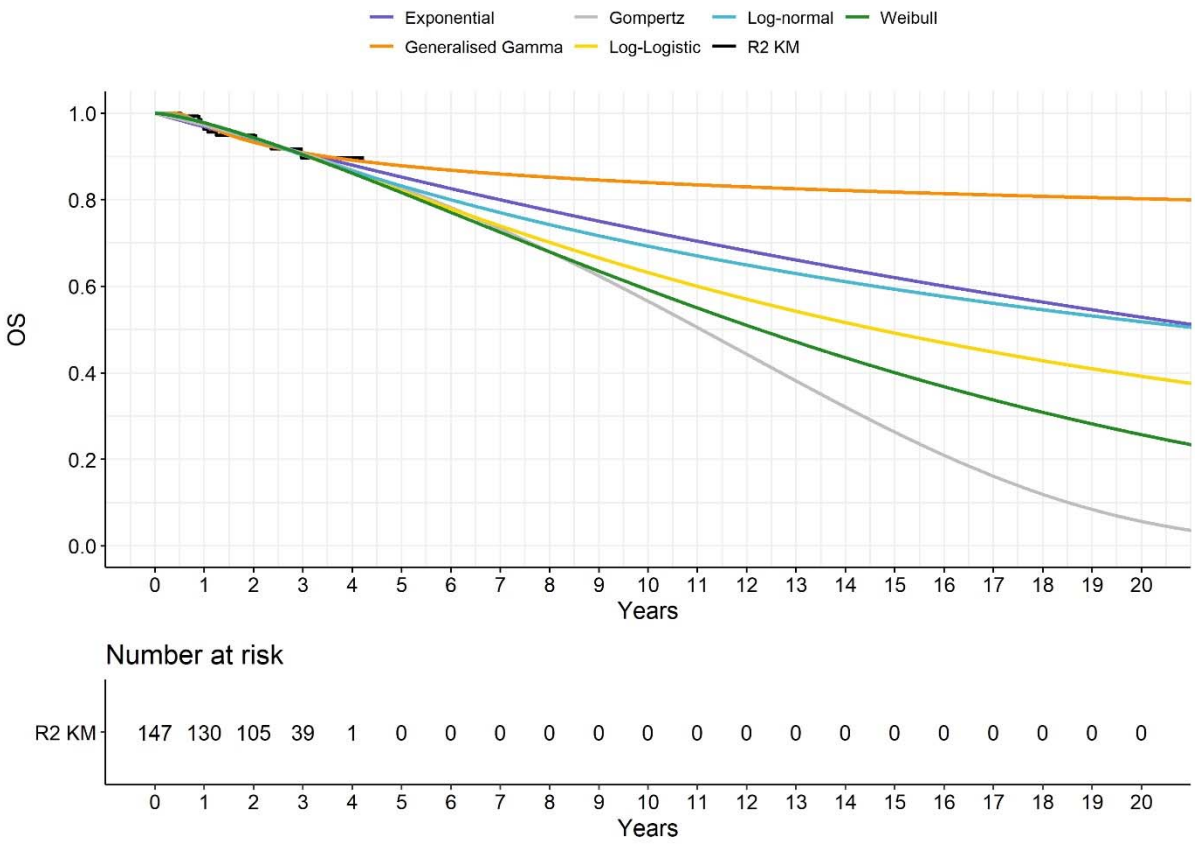


Key: KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

The parametric distributions fit to the R^2 and R-mono Kaplan–Meier data (FL only) are presented in Figure 13 and Figure 14, respectively. Table 16 presents the AIC and BIC fit statistics for each distribution. According to AIC and BIC, generalized gamma and exponential are the best fitting for R^2 and log-normal and exponential are best fitting for R-mono. At the advisory board, clinicians were shown the

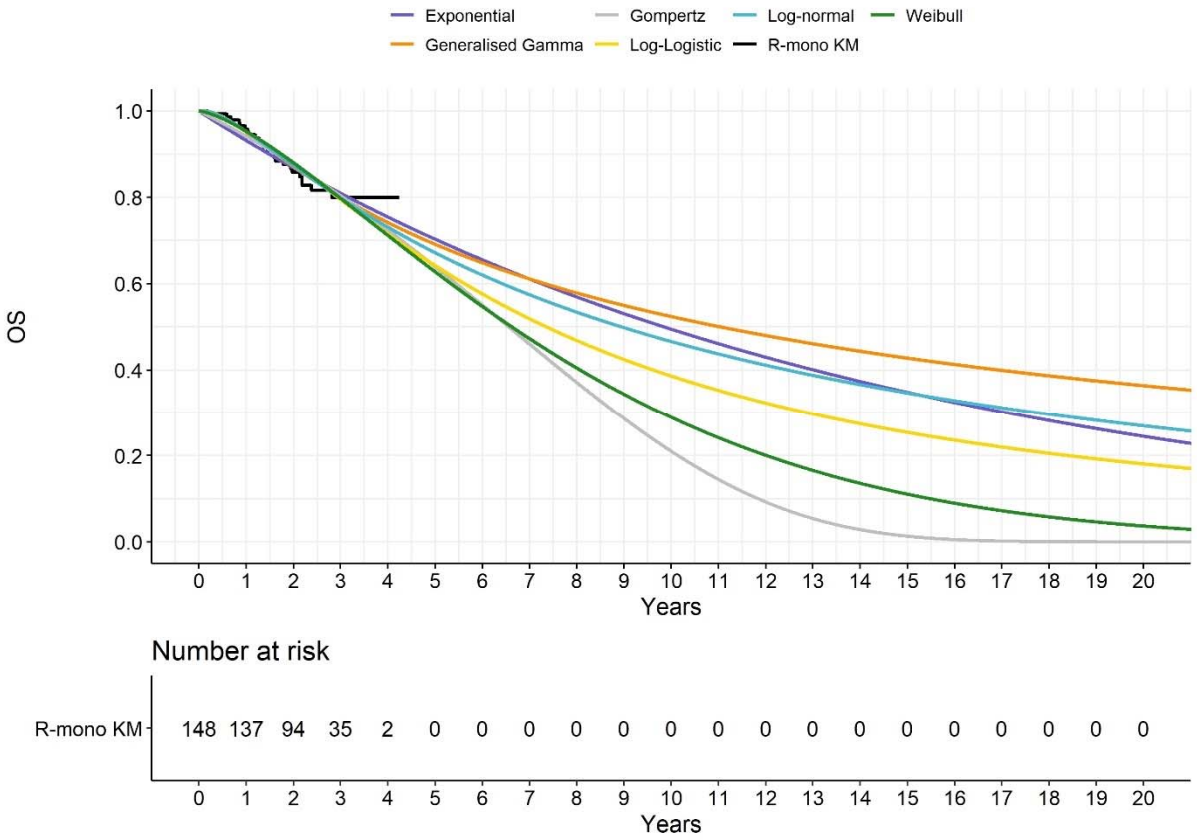
extrapolations from R-mono for the FL+MZL population. They judged the lower curve distributions with approximately 40–50% survival at 10-years to be the most plausible. Therefore, the Weibull distribution was selected for that population for both treatment arms. In response to clarification questions, Weibull was kept as the distribution for the FL only populations to be consistent. However, based on clinical expert feedback on survival at 10-years, either log-logistic or log-normal could also be plausible distributions. These are presented within scenario analysis (Table 34).

Figure 13: OS parametric curves for R^2 (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 14: OS parametric curves for R-mono (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; mono, monotherapy; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.

Table 16: OS AIC and BIC – R^2 versus R-mono (FL only)

Distribution	R^2		R-mono	
	AIC	BIC	AIC	BIC
Exponential	229.63	**232.62**	460.59	**463.59**
Generalized gamma	**228.48**	237.45	459.42	468.41
Gompertz	231.45	237.44	461.95	467.94
Log-logistic	230.34	236.32	458.89	464.88
Log-normal	229.30	235.28	**457.63**	463.62
Weibull	**230.48**	**236.46**	**459.56**	**465.56**

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; mono, monotherapy; OS, overall survival; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: Bold = statistically best fit; **Purple** = selected for base case.

The curves are adjusted to use general population mortality as described in Document B, Section B.3.3.

Progression-free survival

R^2 versus R-CHOP/R-CVP

When comparing the associated KM curves for R^2 and R-CHOP/R-CVP, these appear to be diverged from initiation and then converge at approximately 800 days (see Figure 15) and then diverge again, suggesting the relative treatment effect of R^2 versus R-CHOP/R-CVP is non-constant. This is supported by the log-cumulative hazard plot which is non-parallel (Figure 16). Therefore, in the base case analysis, PFS for R^2 was modelled using the KM data until the maximum follow-up (46.7 months), beyond which the comparator hazard was applied to extrapolate (see below). This approach ensures the relative treatment effect of R^2 versus R-CHOP/R-CVP based on the MAIC is accurately reflected. Moreover, the extrapolation method utilizes the longer follow-up from the HMRN dataset (11.6 years; additional 7.7 years versus AUGMENT) to inform the hazard, which is conservatively applied to both arms over lifetime.

Figure 15: PFS – R^2 (FL only, re-weighted) versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 16: PFS: log-cumulative hazard versus time plot – R^2 (FL only, re-weighted) versus R-CHOP/R-CVP

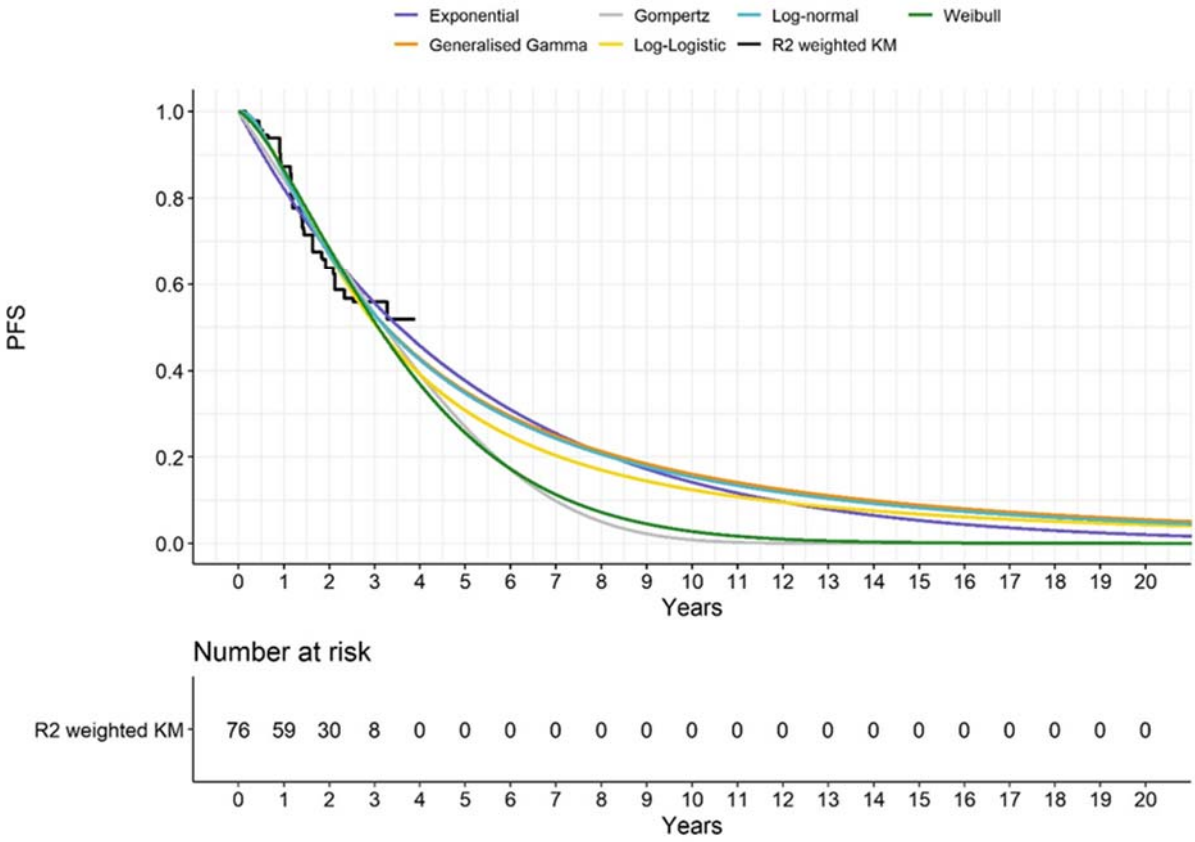
Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

The parametric distributions fit to the R-CHOP/R-CVP KM data and, for completeness and to inform scenario analyses, R^2 (FL only) are presented in Figure 17 (also presented in Document B Section B3.3 Figure 34) and Figure 18, respectively.

Figure 17: PFS parametric curves for R-CHOP/R-CVP in the non-rituximab refractory population (Document B Section B.3.3. Figure 34)

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 18: PFS parametric curves for R^2 (FL only, weighted) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Note: For the base case, R^2 is assumed to assumed to follow the KM and then use the hazard of R-CHOP/R-CVP.

Table 17 presents the AIC and BIC fit statistics for each distribution. The AIC/BIC for R-CHOP/R-CVP PFS were also previously reported in Table 30 of Document B (Section B.3.3. page 148). Log-normal is the best statistical fit for R^2 (FL only) whilst Weibull is the best statistical fit for R-CHOP/R-CVP.

Table 17: PFS: AIC and BIC – R^2 (FL only) versus R-CHOP/R-CVP

Distribution	R^2 (FL only, weighted)		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	457.55	460.54	696.16	698.30
Generalized gamma	456.00	464.97	**671.80**	**678.23**
Gompertz	458.82	464.80	684.73	689.01
Log-logistic	454.41	460.39	673.73	678.02
Log-normal	**454.00**	**459.99**	677.98	682.27
Weibull	456.14	462.12	**670.65**	**674.94**

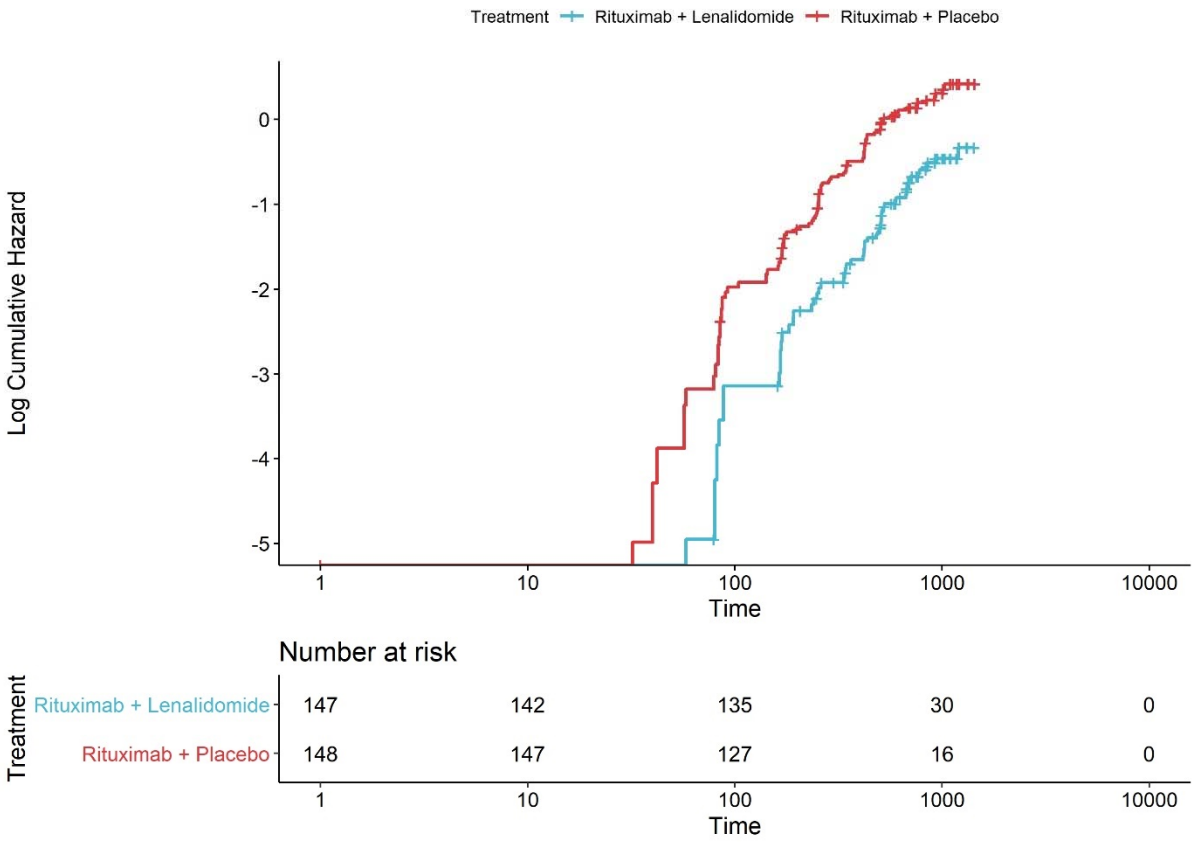
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Bold = statistically best fit; **Purple** = selected for base case.
Note: For the base case, R^2 is assumed to follow KM data and then to match the comparator effectiveness.

The Weibull and generalized gamma curves both appear to fit the R-CHOP/R-CVP data best in comparison to the other distributions which either slightly under- or over-estimate the actual observed data (Figure 17). The Weibull PFS curve crossed the TTNLT curve at approximately 8 years for R-CHOP/R-CVP whereas, in clinical practice, it is unlikely that patients would start their next treatment prior to progressing and so this was considered implausible. Therefore, the second-best fitting curve based on AIC/BIC (generalized gamma) was selected as the base case PFS distribution for R-CHOP/R-CVP. Even with its limitation in crossing the TTNLT curve, Weibull was explored in scenario analysis.

R^2 versus R-mono

For R^2 versus R-mono, parametric curves were fit onto the PFS data from AUGMENT (FL only) trial. Figure 19 presents the log-cumulative hazard plot that shows that assuming proportional hazard is reasonable. However, based on the availability of IPD and to be consistent with OS, stratified models have been used.

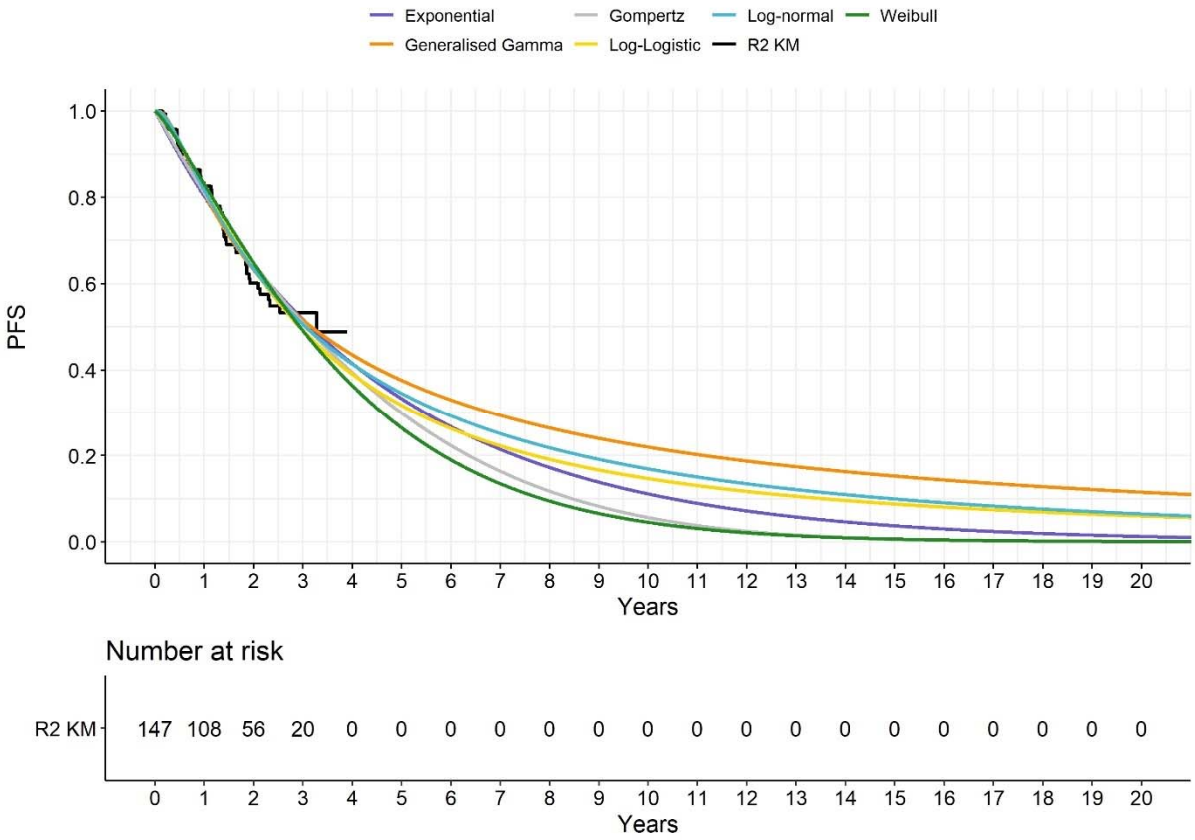
Figure 19: PFS: log-cumulative hazard versus time plot – R^2 versus R-mono



Key: KM, Kaplan–Meier; progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

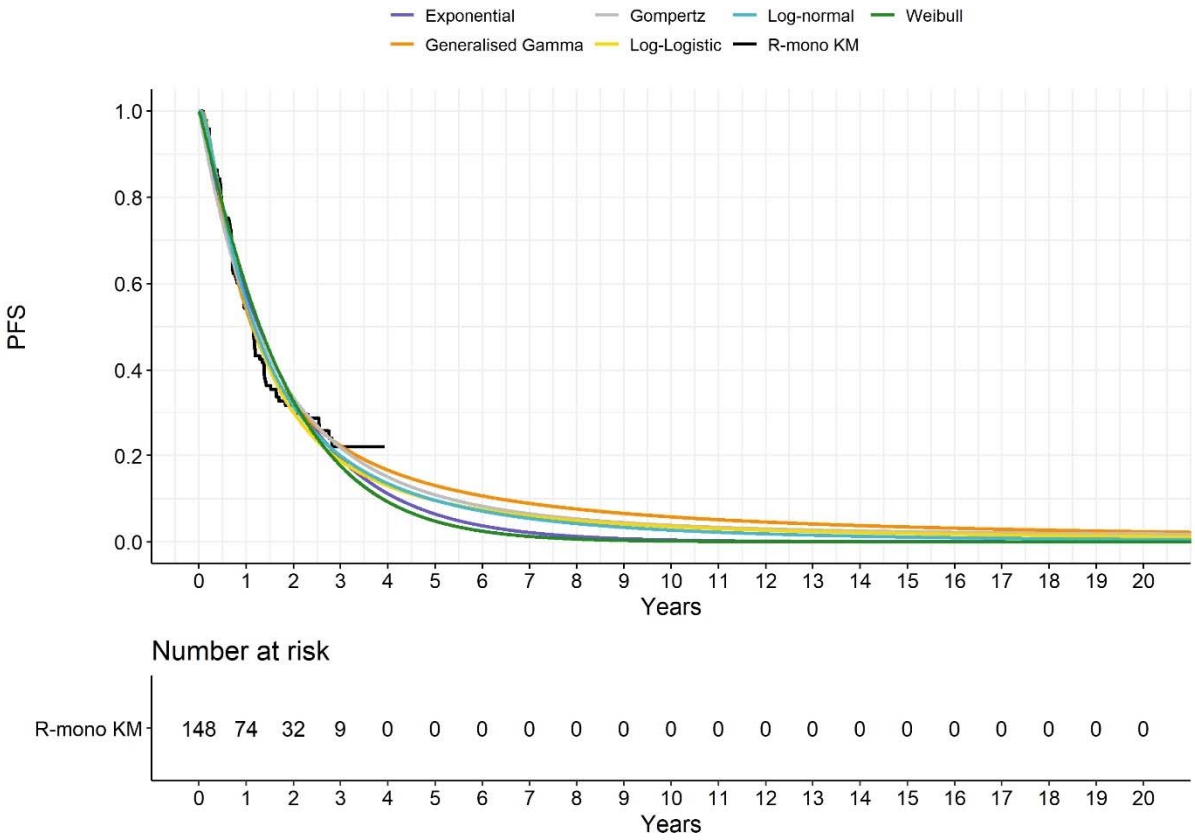
The parametric distributions fit to the R^2 and R-mono KM data (FL only) are presented in Figure 20 and Figure 21, respectively. Table 18 presents the AIC and BIC fit statistics for each distribution. According to AIC and BIC, log-normal is the best fitting for R^2 and generalized gamma or log-normal for R-mono. Log-normal could have been chosen as the base case distribution. However, log-logistic was selected as base case to be consistent with what was chosen for the FL+MZL population. The AIC/BIC statistics show minimal differences between log-logistic and log-normal, additionally, the log-logistic extrapolations are very similar to log-normal so the effect on the ICER is minimal.

Figure 20: PFS parametric curves for R^2 (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 21: PFS parametric curves for R-mono (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; mono, monotherapy; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.

Table 18: PFS AIC and BIC – R^2 versus R-mono (FL only)

Distribution	R^2		R-mono	
	AIC	BIC	AIC	BIC
Exponential	944.49	947.48	1517.31	1520.31
Generalized gamma	940.79	949.76	**1502.76**	1511.75
Gompertz	946.33	952.31	1517.96	1523.95
Log-logistic	**941.10**	**947.08**	**1505.14**	**1511.13**
Log-normal	**939.19**	**945.17**	1502.77	**1508.76**
Weibull	943.81	949.80	1518.19	1524.18

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; mono, monotherapy; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: Bold = statistically best fit; **Purple** = selected for base case.

Time to next anti-lymphoma treatment

R^2 versus R-CHOP/R-CVP

As for OS and PFS, TTNLT data for R^2 from AUGMENT (FL only) were matched to the R-CHOP/CVP cohort from HMRN (see Section 2.3).

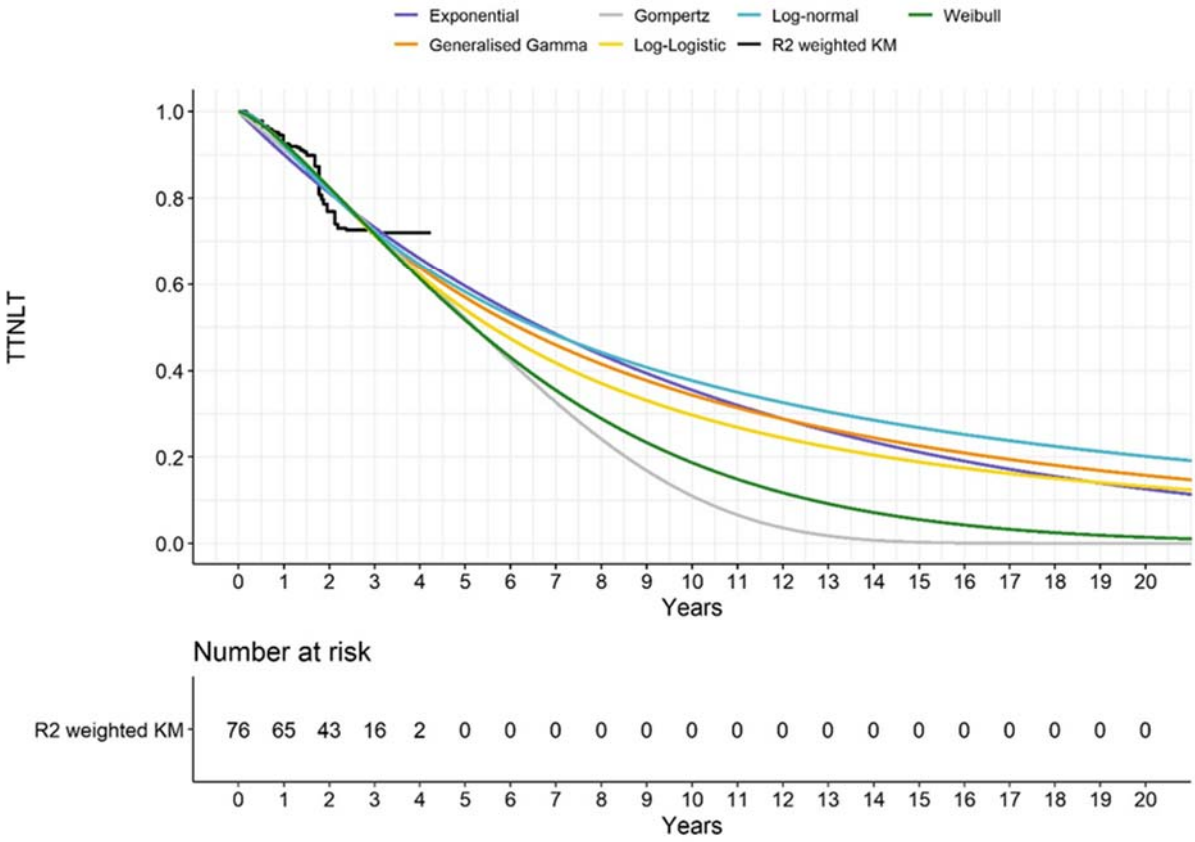
Figure 22 presents the log-cumulative hazard versus time plot between R^2 and R-CHOP/R-CVP from AUGMENT and HMRN. The lines of the log cumulative hazards show that it could be argued that the proportional hazards assumption is not unreasonable. However, to be consistent with the OS and PFS data, stratified statistical models have been used.

Figure 22: TTNLT: log-cumulative hazard versus time plot – R^2 versus R-CHOP/R-CVP

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

Figure 23 presents the parametric distributions fit to the R^2 KM data (FL only) for TTNLT. Figure 24 presents the TTNLT parametric distributions fit to R-CHOP/R-CVP, as also previously shown in Figure 43 of Document B (Section B.3.3. page 160).

Figure 23: TTNLT parametric curves for R^2 (FL only, weighted) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

Figure 24: TTNLT parametric curves for R- CHOP/R-CVP in the non-rituximab refractory population (Document B Section B.3.3. Figure 43)

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

Table 19 presents the AIC and BIC fit statistics for each distribution. The AIC/BIC for R-CHOP/R-CVP TTNLT were also previously reported in Table 32 of Document B (Section B.3.3. page 159). For R^2, log-normal and exponential are statistically the best fitting, whereas for R-CHOP/R-CVP log-normal is the statistically best fitting.

Table 19: TTNLT: AIC and BIC – R^2 (FL only) versus R-CHOP/R-CVP

Distribution	R^2 (FL only, weighted)		R-CHOP/R-CVP	
	AIC	BIC	AIC	BIC
Exponential	316.51	**319.50**	652.37	654.51
Generalized gamma	318.29	327.26	648.65	655.08
Gompertz	318.08	324.06	647.18	651.47
Log-logistic	316.38	322.36	647.77	652.06
Log-normal	**316.32**	**322.30**	**646.97**	**651.26**
Weibull	316.86	322.84	651.70	655.99

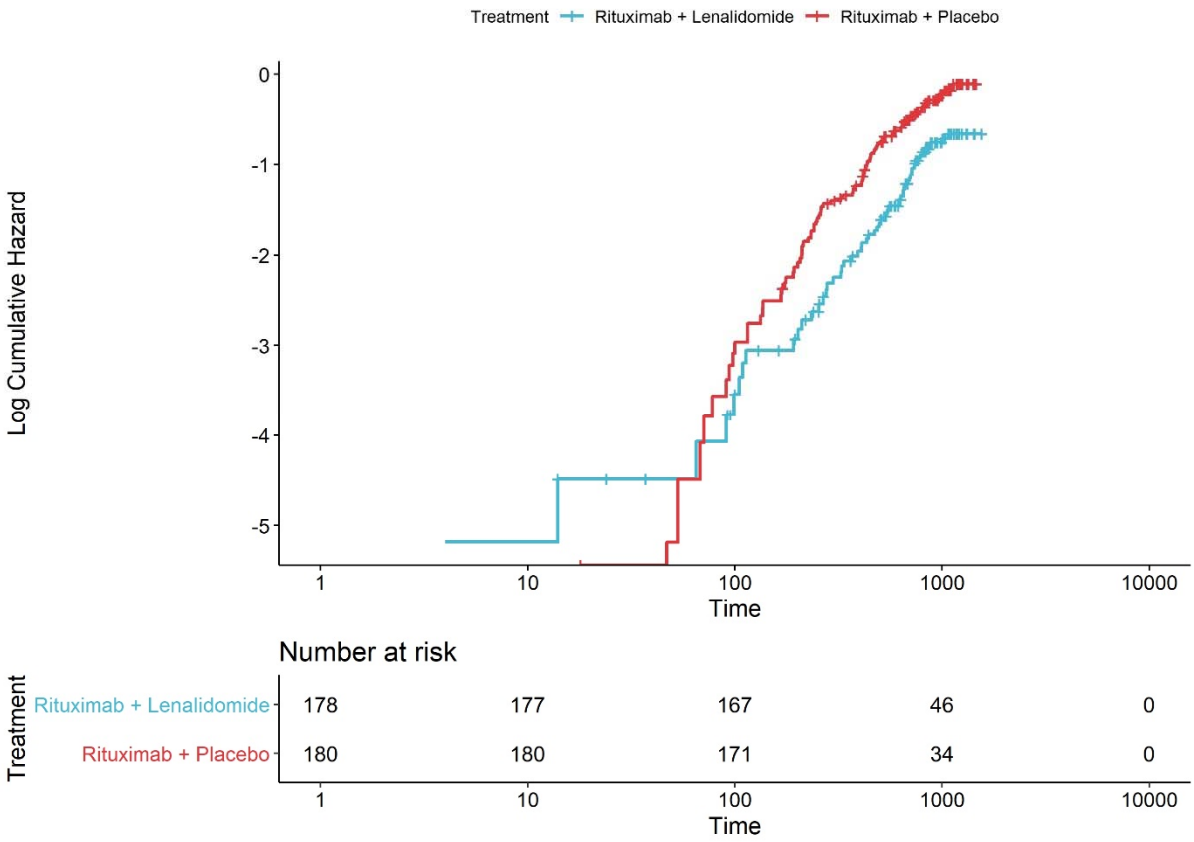
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.
Bold = statistically best fit; **Purple** = selected for base case.

Given that log-normal is the statistically best fitting extrapolation for R-CHOP/R-CVP and R^2 (FL only) according to the AIC statistic, and provides a reasonable fit, this has been selected as the base case curve.

R^2 versus R-mono

As for OS and PFS, parametric curves were fit directly to the AUGMENT (FL only) trial data. Figure 25 presents the log-cumulative hazard plot that shows that the curves meet just under 100 days, but stay reasonably parallel thereafter. However, as with the other end points, stratified models have been selected.

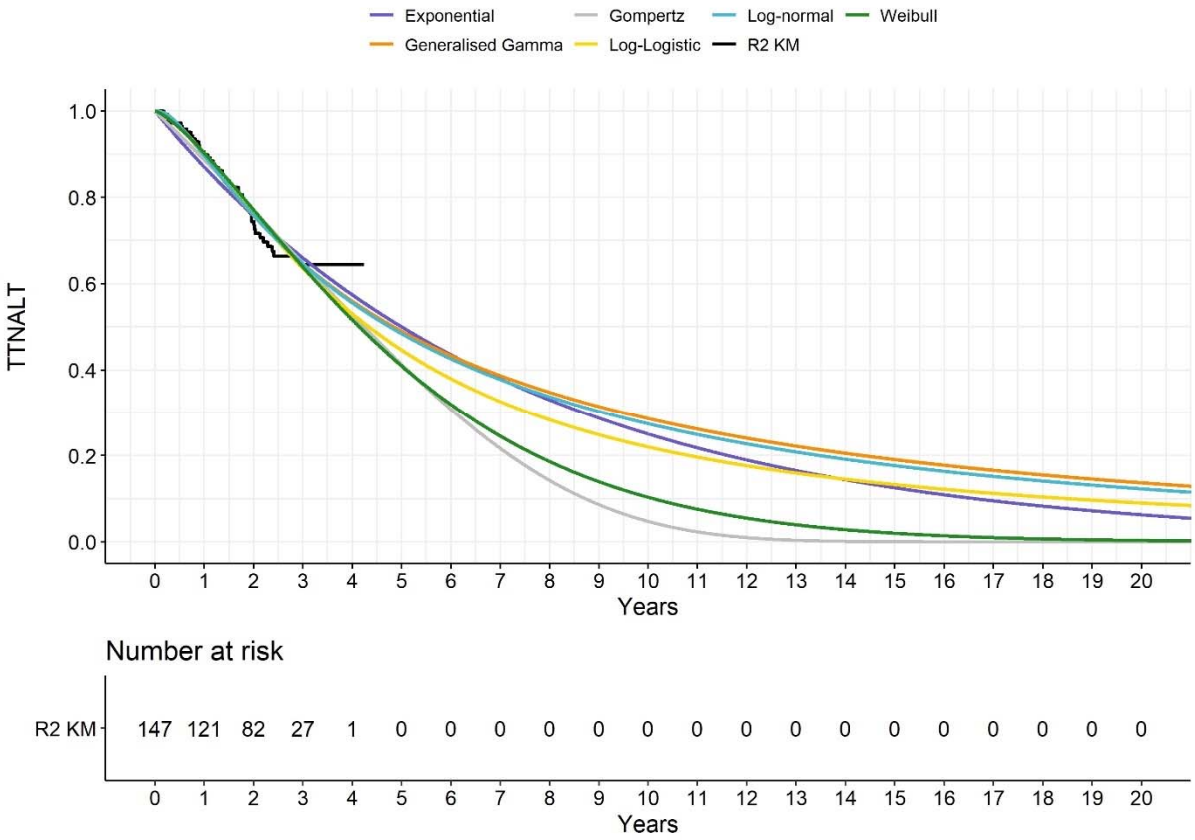
Figure 25: TTNLT: log-cumulative hazard versus time plot – R^2 versus R-mono



Key: KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

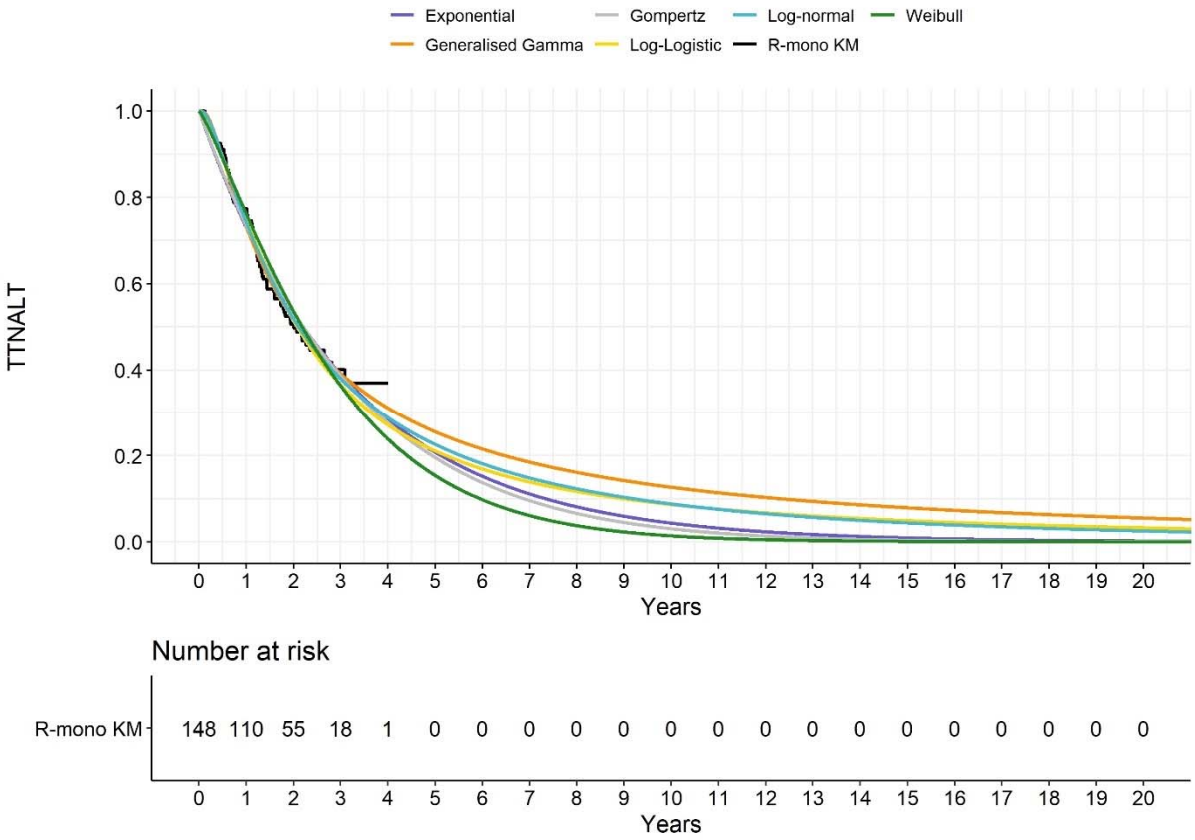
The parametric distributions fit to the R^2 and R-mono KM data (FL only) are presented in Figure 26 and Figure 27, respectively. Table 20 presents the AIC and BIC fit statistics for each distribution that shows that log-normal has the best statistical fit. While acknowledging the uncertainty in the selection of distributions, the log-normal has been selected for the base case. This is due to having the lowest AIC/BIC value for both treatment arms and this maintains consistency with the distribution selected for the R-CHOP/R-CVP comparison.

Figure 26: TTNLT parametric curves for R^2 (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.

Figure 27: TTNLT parametric curves for R-mono (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; mono, monotherapy; R, rituximab; TTNLT, time to next anti-lymphoma treatment.

Table 20: TTNLT: AIC and BIC – R^2 (FL only) versus R-mono

Distribution	R^2 (FL only)		R-mono	
	AIC	BIC	AIC	BIC
Exponential	747.65	750.64	1275.31	1278.30
Generalized gamma	745.17	754.14	1267.31	1276.30
Gompertz	748.59	754.57	1277.27	1283.26
Log-logistic	744.05	750.03	1268.45	1274.44
Log-normal	**743.18**	**749.16**	**1265.99**	**1271.99**
Weibull	745.52	751.50	1274.11	1280.11

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; FL, follicular lymphoma; mono, monotherapy; R, rituximab; R^2, lenalidomide plus rituximab; TTNLT, time to next anti-lymphoma treatment.
Bold = statistically best fit; **Purple** = selected for base case.

Time on treatment

Data for ToT were used to establish the proportion of patients on treatment per cycle to calculate the overall drug costs. ToT KM data were used and extrapolated using parametric distributions. As all induction treatments and maintenance treatments have a maximum duration, all extrapolations were capped at this time.

R^2 versus R-CHOP/R-CVP

ToT data for R^2 versus R-CHOP/CVP were taken from the same efficacy source as the OS, PFS and TTNLT. The proportion of patients who were still on treatment over time was extracted and fitted with parametric curves. Due to the shape of the ToT KMs, the parametric curves produced poor fits to the data, largely over-estimating or under-estimating the actual proportion of patients on treatment (see Figure 28 and Figure 29). Consequently, and since induction and maintenance treatment is for a fixed period, the KM data were used directly in the model to inform the proportion of patients on treatment.

The maximum treatment durations used to cap the ToT curves were the same as the FL+MZL analysis and are described in Section B.3.3 of Document B.

Figure 28: ToT parametric curves for R^2 (FL only, weighted) in the non-rituximab refractory population

Key: FL, follicular lymphoma; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration.

Figure 29: ToT parametric curves for R- CHOP/R-CVP in the non-rituximab refractory population (Document B Section B.3.3. Figure 48)

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; KM, Kaplan–Meier; R, rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration. R-CHOP and R-CVP ToT is capped at 990 days.

Table 21 presents the AIC and BIC fit statistics for the parametric distributions for completeness. The AIC/BIC for R-CHOP/R-CVP ToT were previously reported in Table 34 of Document B (Section B.3.3. page 165).

Table 21: ToT: AIC and BIC – R^2 (FL only) versus R-CHOP/R-CVP

Distribution	R^2 (FL only, weighted)		R-CHOP/R-CVP	
	AIC	**BIC**	**AIC**	**BIC**
Exponential	1028.70	1031.69	**882.61**	**884.76**
Generalized gamma	**828.23**	837.18	884.96	891.39
Gompertz	830.22	**836.19**	883.73	888.02
Log-logistic	931.13	937.10	894.76	899.05
Log-normal	974.55	980.52	901.89	906.18
Weibull	886.16	892.12	884.61	888.90

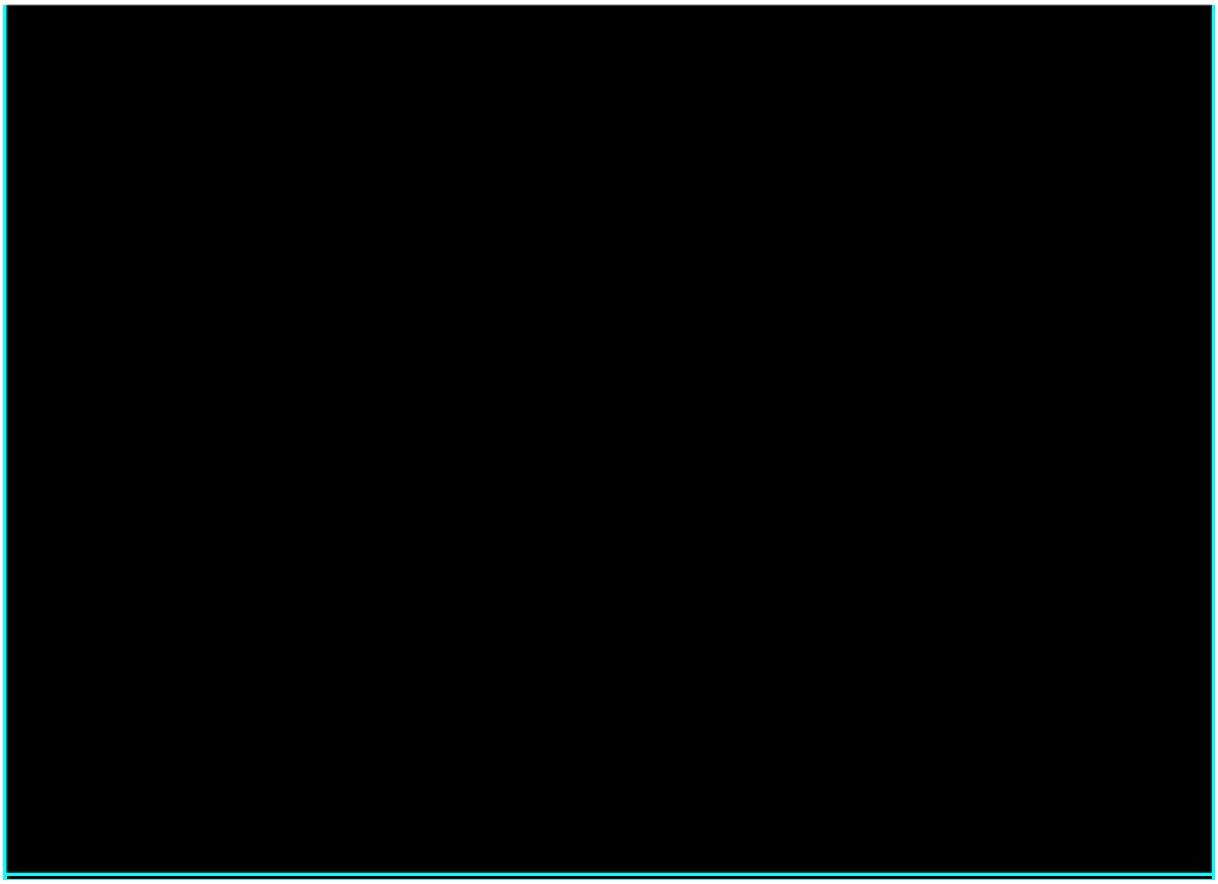
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; KM, Kaplan–Meier; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: KM data used directly for the base case.
Bold = statistically best fit

R^2 versus R-mono

As per the other endpoints, parametric curves were fit directly onto the ToT AUGMENT (FL only) trial data. Similarly for the R^2 versus R-CHOP/CVP analysis, the shape of the KMs meant that the curve produced poor fits (see Figure 30 and Figure 31). Therefore, the KM was used directly for the base case. ToT data was capped at 336 days as this was the maximum treatment duration in the AUGMENT trial and no maintenance treatment was given to either treatment arm.

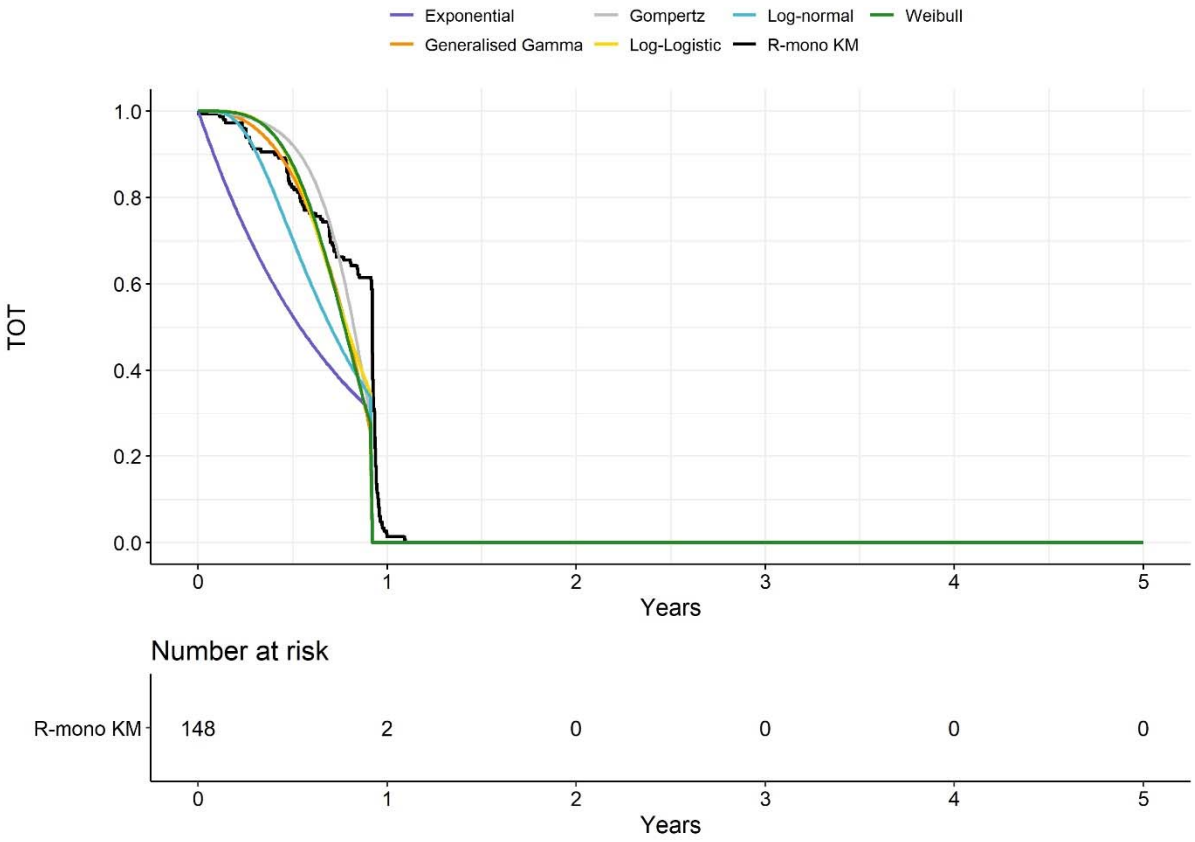
For completeness, the parametric distributions fit to the R^2 and R-mono Kaplan-Meier data (FL only) are presented in Figure 30 and Figure 31, respectively. Table 22 presents the AIC and BIC fit statistics for each distribution.

Figure 30: ToT parametric curves for R^2 (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration.

Figure 31: ToT parametric curves for R-mono (FL only) in the non-rituximab refractory population



Key: FL, follicular lymphoma; KM, Kaplan–Meier; mono, monotherapy; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: Treatment periods are fixed in the model, and therefore any distributions are capped at the maximum treatment duration.

Table 22: ToT: AIC and BIC – R^2 (FL only) versus R-mono

Distribution	R^2 (FL only)		R-mono	
	AIC	BIC	AIC	BIC
Exponential	1964.40	1967.38	1970.05	1973.05
Generalized gamma	1653.88	1662.83	**1681.20**	**1690.19**
Gompertz	**1633.83**	**1639.80**	1687.24	1693.23
Log-logistic	1779.98	1785.95	1838.60	1844.59
Log-normal	1848.60	1854.56	1930.14	1936.13
Weibull	1704.84	1710.80	1766.22	1772.22

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; FL, follicular lymphoma; KM, Kaplan–Meier; mono, monotherapy; R, rituximab; R^2, lenalidomide plus rituximab; ToT, time on treatment.
Note: KM data used directly for the base case.
Bold = statistically best fit

Safety

As described in Document B Section B.3.3 of the original submission, Grade 3/4 adverse events of greater than 2% in either treatment arm were considered. For the FL only analysis in the cost-effectiveness model, only the AEs from the FL population in AUGMENT were used to calculate the incidence and resulting costs of adverse events. As per the methodology used for the FL and MZL pooled population, the adverse events from RELEVANCE were adjusted to the AUGMENT data and used to inform the adverse event incidence for the R-CHOP and R-CVP arms. The same costs as those presented in Section B.3.5 were applied to these adverse event rates to calculate the total adverse event cost for each treatment.

Table 23 presents the adverse event incidences and total costs per treatment arm.

Table 23: Grade 3/4 AE incidence: non-rituximab (FL only) refractory population

AE	R^2 (n=146)		R-CHOP (%)	R-CVP (%)	R-mono (n=148)	
	n	%			n	%
Neutropenia	74	50.7%	91.5%	86.5%	18	12.2%
Leukopenia	9	6.2%	27.8%	15.0%	3	2.0%
Anaemia	6	4.1%	4.1%	4.1%	1	0.7%
Pneumonia	5	3.4%	NR	NR	0	0.0%
Lymphocyte count decreased	5	3.4%	NR	NR	2	1.4%
Lymphopenia	5	3.4%	NR	NR	2	1.4%
Febrile neutropenia	4	2.7%	9.0%	4.9%	1	0.7%
White blood cell count decreased	5	3.4%	NR	NR	2	1.4%
Diarrhoea	5	3.4%	1.9%	6.7%	0	0.0%
Thrombocytopenia	2	1.4%	1.2%	0.0%	0	0.0%
Hypokalaemia	4	2.7%	NR	NR	0	0.0%
Pulmonary embolism	4	2.7%	NR	NR	1	0.7%
Infusion related reaction	1	0.7%	0.1%	0.0%	0	0.0%
Nausea and emesis	0	0.0%	1.4%	3.8%	1	0.7%
Allergic reaction	1	0.7%	NR	NR	0	0.0%
Hypotension	1	0.7%	NR	NR	0	0.0%
Fatigue	2	1.4%	3.8%	0.0%	1	0.7%
Alopecia	NR	NR	0.8%	0.0%	NR	NR
Abdominal pain	1	0.7%	0.7%	0.0%	0	0.0%
Acute kidney injury	2	1.4%	NR	NR	0	0.0%
Total cost	**£1,796**		**£3,471**	**£2,714**	**£393**	
Source	AUGMENT		RELEVANCE (adjusted)[a]		AUGMENT	

Key: AE, adverse event; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; mono, monotherapy; NR, not reported; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a] Comparator AE incidence = (AECOMPARATOR incidence in RELEVANCE/AER^2 incidence in RELEVANCE) x AER^2 incidence in AUGMENT.

HRQL inputs for the FL model

The EQ-5D meta-regression model was re-ran from AUGMENT using only the FL population according to the same methodology previously described in Section B.3.4 of the original submission. The parameters of the parsimonious mixed effects model and resulting final utility values used to inform the FL only population analyses in the cost-effectiveness model are displayed in Table 24 and Table 25, respectively.

Table 24: Parsimonious mixed effects FL only model

Parameters	Estimate	SE	p-value
Intercept	0.341	0.037	<0.001
Health state: Pre-progression	0.026	0.010	0.016
Health state: Progressed on treatment	-0.035	0.014	0.015
Randomized treatment arm: R^2	0.002	0.015	0.918
Baseline utility	0.548	0.039	<0.001
Previous rituximab exposure: no	0.037	0.021	0.078
Refractory to last therapy: yes	-0.046	0.021	0.031
Number of prior therapies: 1	0.039	0.015	0.012
Key: R^2, lenalidomide plus rituximab; SE, standard error.			

Table 25: Final utility values from the FL only model

Health state	Utility (R^2 vs R-CHOP/CVP)	Utility (R^2 vs R-mono)
Progression-free	0.867	0.846
Progressed (off treatment)	0.841	0.820
Progressed (on treatment)	0.806	0.785
Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; mono, monotherapy; R^2, lenalidomide plus rituximab.		

Cost inputs for the FL model

Any cost inputs for the FL population that are different from the costs outlined in the original submission Document B has been detailed below. All other costs not discussed are relevant for the submission and are presented in Document B.

Treatment costs

The same costs for treatments and methodology as described in Document B, Section B.3.5 have been used to inform the treatment costs. The dose reductions for lenalidomide data have been taken directly from the AUGMENT trial using only the FL patients for the analysis (as described in Section B.3.5 in the original submission

document). These are then used to calculate the cost per cycle applied to the proportion of patients on treatment.

Subsequent treatments

Subsequent treatments usage from the FL only population in AUGMENT were used to inform the costs of subsequent treatments for the R^2 arm in this analysis. The durations of the subsequent treatments and cost per treatment remain the same as per Section B.3.5 in the original submission document. Table 26 presents the FL only subsequent treatment usage from AUGMENT and the resulting subsequent treatment costs for the R^2 and R-mono arm.

Table 26: Subsequent treatments and costs

Subsequent treatment	**R^2 (FL only AUGMENT) n (%)[a] n=147**	**R-mono (FL only AUGMENT) n (%)[a]** n=148
R-mono	[redacted]	[redacted]
R-Benda	[redacted]	[redacted]
R-CHOP	[redacted]	[redacted]
R-CVP	[redacted]	[redacted]
Other R-chemo	[redacted]	[redacted]
O-Benda	[redacted]	[redacted]
Bendamustine	[redacted]	[redacted]
Other chemotherapy	[redacted]	[redacted]
Targeted therapies	[redacted]	[redacted]
Radiotherapy	[redacted]	[redacted]
Other	[redacted]	[redacted]
ASCT	[redacted]	[redacted]
Total weighted treatment cost (R^2)	**£3,128**	**£6,126**
Total weighted administration cost (R^2)	**£349**	**£524**

Key: ASCT, autologous stem cell transplant; Benda, bendamustine; chemo, chemotherapy; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; mono, monotherapy; FL, follicular lymphoma; R, rituximab; R^2, lenalidomide plus rituximab.
Notes: [a]Percentages include multiple lines therefore total may be over 100%.

3.2. Economic results

Base case results for the FL population

Deterministic

Deterministic base case results are presented in Table 27. As per the request in clarification question B3, the FL only analysis has also been conducted for the R^2 versus R-mono comparison. The results show that R^2 remains cost-effective in the FL-only population versus R-CHOP and R-CVP as well as being cost-effective versus the AUGMENT study comparator R-mono at the £30,000 willingness-to-pay threshold.

Table 27: Deterministic pairwise results (based on PAS price) – FL only

Technology	Total			Incremental			ICER (£/QALYs)
	Costs	LYG	QALYs	Costs	LYG	QALYs	
R^2 versus R-CHOP							
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
R2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R^2 versus R-CVP							
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R^2 versus R-mono							
R-mono	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£20,274

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; LYG, life years gained; mono, monotherapy; PAS, patient access scheme; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab.

Table 28 presents the fully incremental analysis for the non-R-refractory population. R-CHOP is strictly dominated by R-CVP, given that R-CVP has smaller AE costs and no patients who went onto ASCT.

Table 28: Deterministic fully incremental analysis – non-rituximab refractory population R² versus R-CHOP and R-CVP (with PAS)

Treatment	Total		Incremental		ICER (strict dominance)	Incremental ICER (extended dominance)
	Costs	QALYs	Costs	QALYs		
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]		
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Strictly Dominated
R²	[redacted]	[redacted]	[redacted]	[redacted]	£23,746	£23,746

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; PAS, patient access scheme; QALY, quality-adjusted life-year; R, rituximab; R², lenalidomide plus rituximab.

Although Table 27 presents the deterministic results derived from base case curve selections for R-CHOP and R-CVP, the impact of selecting other OS and PFS distributions which also show clinical and statistical plausibility is explored below in scenario analysis (Table 30 and Table 31).

Base case results for R² versus O-Benda in the FL only population are as previously shown in Table 64 of Document B (Section B.3.7. page 210).

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed within the cost-effectiveness model for 1,000 iterations. The mean incremental costs and QALYs from R² versus the comparators are displayed in Table 29. The visual results of the PSA runs are displayed in Figure 32–Figure 34.

There is more variation between the results of the probabilistic analysis and the deterministic analysis for the R-CHOP/CVP comparison in the FL only population than was seen in the FL and MZL population of the original submission. This is driven by differences in probabilistic total QALYs versus deterministic total QALYs. Focussing on a smaller population (FL only) leads to increased uncertainty in the R² OS extrapolations. Although these differences in probabilistic results between the FL only and FL+MZL populations increase the probabilistic ICERs of R² versus comparators, the probabilistic uncertainty is highly influenced by the survival distributions selections. Therefore, both the probabilistic and deterministic results for other combinations of plausible survival curve selections are also worth considering, as displayed in Table 30 and Table 31 below.

The results of the probabilistic analysis for R² versus R-mono are similar to those of the deterministic analysis, with R² remaining cost-effective at the £30,000 willingness-to-pay (WTP) threshold.

Table 29: Mean results of PSA (1,000 runs) and comparison with deterministic results – FL only

Technology	Total costs (£)		Total QALYs		ICER (£) versus baseline (QALYs)	
	PSA	Deterministic	PSA	Deterministic	PSA	Deterministic
R^2 versus R-CHOP						
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]		
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,768	£15,909
R^2 versus R-CVP						
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]		
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£41,602	£23,746
R^2 versus R-mono						
R-mono	[redacted]	[redacted]	[redacted]	[redacted]		
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,412	£20,274

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CI, confidence interval; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; mono, monotherapy; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab.

Figure 32: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus R-CHOP – FL only

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; FL, follicular lymphoma; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 33: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus R-CVP – FL only

Key: CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 34: Cost-effectiveness plane (1,000 PSA runs) – R^2 versus R-mono – FL only

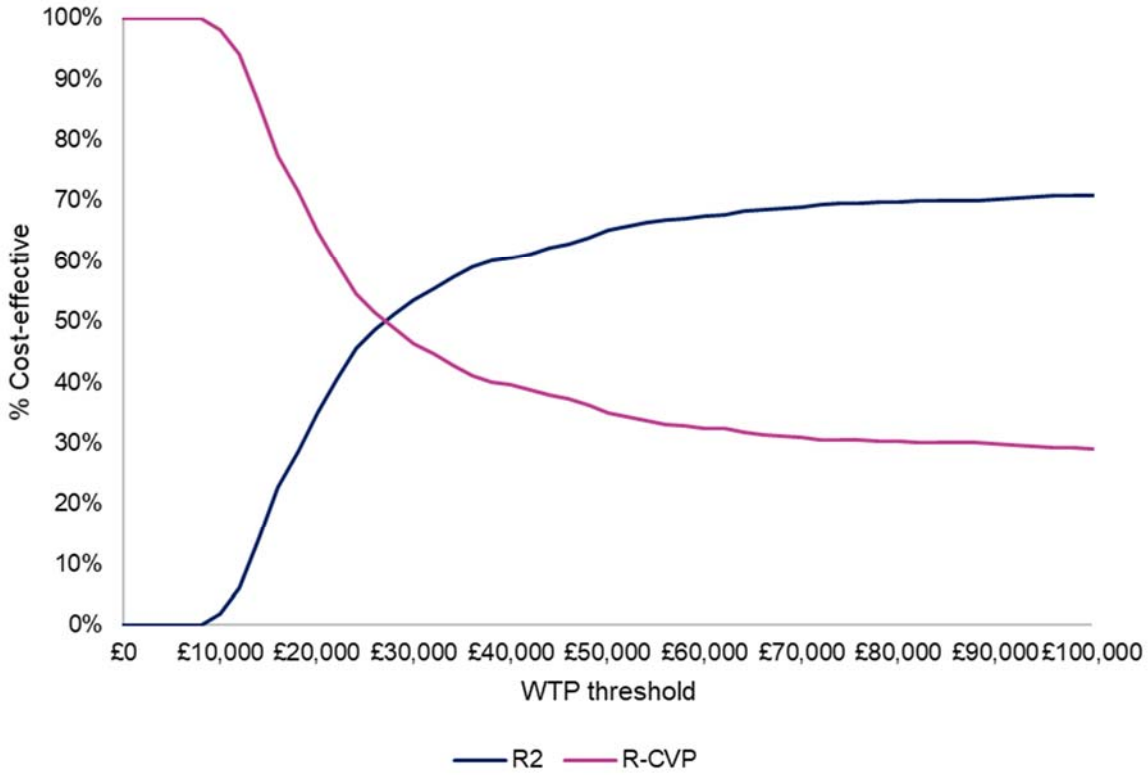
Key: FL, follicular lymphoma; mono, monotherapy; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 35–Figure 37 presents the cost-effectiveness acceptability curves for R^2 versus comparators in the FL only population based on the 1,000 PSA iterations at different willingness-to-pay (WTP) thresholds. At the £30,000 WTP threshold, the probability of R^2 being cost-effective is 63.3%, 53.6% and 77.3% compared with R-CHOP, R-CVP and R-mono, respectively.

Figure 35: Cost-effectiveness acceptability curve – R^2 versus R-CHOP - FL only

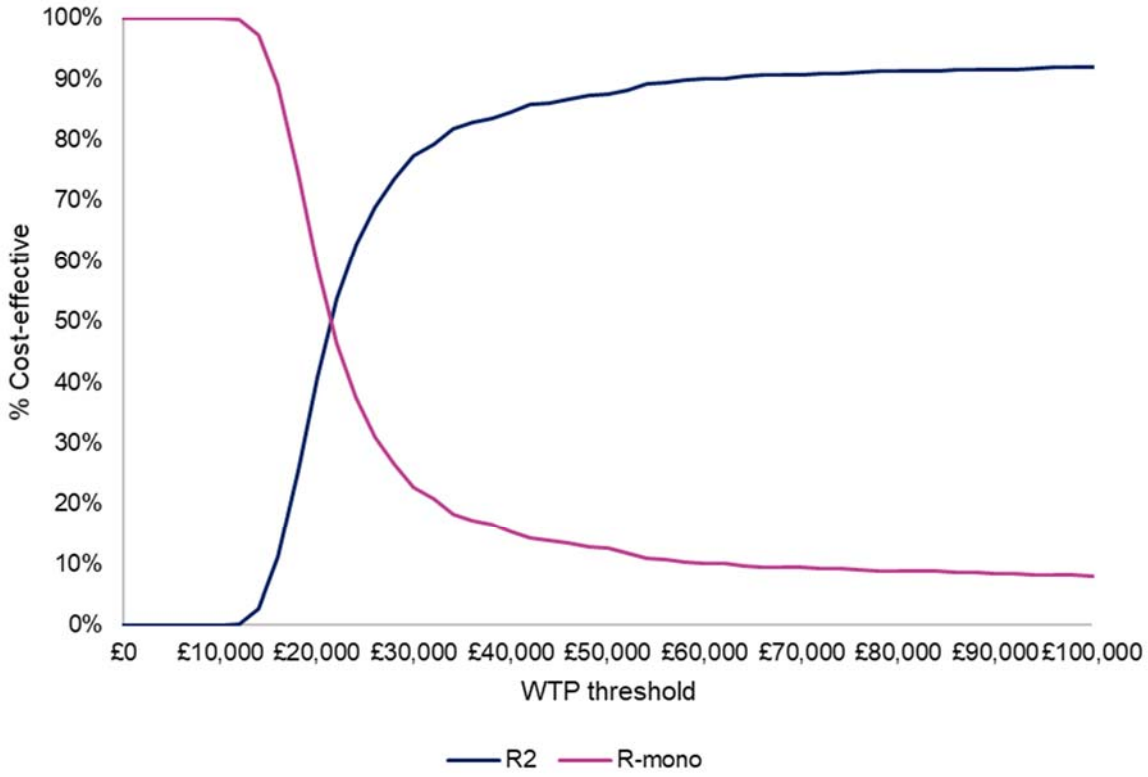
Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; FL, follicular lymphoma; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 36: Cost-effectiveness acceptability curve – R^2 versus R-CVP - FL only



Key: CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Figure 37: Cost-effectiveness acceptability curve – R^2 versus R-mono - FL only



Key: FL, follicular lymphoma; mono, monotherapy; R, rituximab; R^2, lenalidomide plus rituximab; WTP, willingness to pay.

Probabilistic results for R^2 versus O-Benda in the FL only population are as previously shown in Tables 66–67, Figure 53 and Figure 56 of Document B (Section B.3.8.).

Key sensitivity and scenario analyses

Curve fit scenario analyses

In addition to the PFS and OS curve selections applied at base case, several alternative distributions for OS and PFS were identified in Section 3.1 as being of interest for exploration in scenario analysis for the R-CHOP and R-CVP comparisons:

- Weibull is the best statistical fit for R-CHOP/R-CVP PFS and also gives good visual fit to the observed Kaplan Meier data. However, the clinical plausibility of this curve may be somewhat limited given it crosses the R-CHOP/R-CVP TTNLT curve at approximately 8 years.

- In addition to the Weibull base case OS curve, the exponential and log-logistic curves were also considered plausible for R-CHOP/CVP OS
- The log-normal distribution for R^2 (FL only) OS has the best statistical fit according to the AIC statistic

Table 30 and Table 31 present the deterministic and probabilistic cost effectiveness results for the R-CHOP and R-CVP comparisons, respectively, across the combinations of key survival curve scenarios which show plausibility in addition to those selected at base case. The majority of curve fits explored result in deterministic and probabilistic ICERs below the £30,000 threshold for the comparison with R-CHOP and R-CVP.

Table 30: Results of key curve fit scenario analyses (with PAS) - R^2 versus R-CHOP – FL only

Curve fit selection			ICER (£) versus baseline (QALYs)	
OS		PFS		
R^2	**R-CHOP/CVP**	**R-CHOP/CVP**	**PSA**	**Deterministic**
Base case				
Weibull	Weibull	Generalized gamma	£27,768	£15,909
Plausible scenarios				
Weibull	Weibull	Weibull[a]	£26,827	£15,105
Exponential[b]	Exponential[b]	Generalized gamma	£12,953	£12,651
Exponential[b]	Exponential[b]	Weibull[a]	£12,002	£12,061
Log-logistic[b]	Log-logistic[b]	Generalized gamma	£15,925	£12,955
Log-logistic[b]	Log-logistic[b]	Weibull[a]	£14,800	£12,322
Log-normal[c]	Weibull	Generalized gamma	£15,894	£13,176
Log-normal[c]	Weibull	Weibull[a]	£14,835	£12,541

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab.
Note: [a], Weibull based on best PFS AIC/BIC for R-CHOP/CVP and fit the data well; [b], based on other clinically plausible OS extrapolations from clinical expert opinion; [c], Based on best OS AIC/BIC for R^2.

Table 31: Results of key curve fit scenario analyses (with PAS) - R^2 versus R-CVP – FL only

Curve fit selection			ICER (£) versus baseline (QALYs)	
OS		PFS		
R^2	R-CHOP/CVP	R-CHOP/CVP	PSA	Deterministic
Base case				
Weibull	Weibull	Generalized gamma	£41,602	£23,746
Scenarios				
Weibull	Weibull	Weibull[a]	£40,630	£22,886
Exponential[b]	Exponential[b]	Generalized gamma	£18,555	£18,394
Exponential[b]	Exponential[b]	Weibull[a]	£18,514	£17,774
Log-logistic[b]	Log-logistic[b]	Generalized gamma	£26,740	£18,886
Log-logistic[b]	Log-logistic[b]	Weibull[a]	£25,172	£18,218
Log-normal[c]	Weibull	Generalized gamma	£23,321	£19,253
Log-normal[c]	Weibull	Weibull	£22,089	£18,584

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; ICER, incremental cost-effectiveness ratio; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; PSA, probabilistic sensitivity analysis; QALYs, quality-adjusted life years; R, rituximab; R^2, lenalidomide plus rituximab.
Note: [a], Weibull based on best PFS AIC/BIC for R-CHOP/CVP and fit the data well; [b], Based on other clinically plausible OS extrapolations from clinical expert opinion; [c], Based on best OS AIC/BIC for R^2.

Additional scenario analyses

Table 32–Table 34 present the scenario analysis performed to assess structural uncertainty within the model for the FL only population. The majority of the scenarios tested resulted in ICERs below the £30,000 threshold for the comparisons with the following exceptions:

- a 5-year time horizon for all comparators
- a 10-year treatment effect cap for R-CHOP/R-CVP comparisons
- a 10-year time horizon or Weibull PFS distribution for the R-CVP comparison
- a 3-year treatment effect cap for the R-mono comparison

Given the indolent nature of lymphoma, 5- and 10-year time horizons are considered too short to capture the life-time benefit. Applying the treatment effect at 10-years causes R^2 to be dominated by R-CHOP and R-CVP as the base case curves for OS cross before this timepoint. It is highly implausible that the survival of R^2 would diminish and become worse than R-CHOP and R-CVP after a significant initial gain. Other plausible curves tested for R^2 and R-CHOP/CVP do not result in dominance

when applying the treatment effect at 10-years with most ICERs improving and all remaining under £30,000 per QALY.

Scenario analysis results for R^2 versus O-Benda in the FL only population are as previously shown in Table 70 of Document B (Section B.3.8.).

Table 32: Results of scenario analysis (with PAS) – R^2 versus R-CHOP – FL only

Parameter	Base case	Scenario	ICER (£/QALY)
Base case			*£15,909*
Time horizon	40 years	5 years	£33,331
		10 years	£21,900
		20 years	£17,067
No half cycle correction	Applied	Not applied	£15,819
Discount rate for costs	3.50%	0.0%	£15,969
		6.0%	£15,814
Discount rate for QALYs	3.50%	0.0%	£11,913
		6.0%	£18,873
Source of adverse event frequencies: literature	1L trial	Literature	£19,250
Apply comparator hazard to R^2 arms: after	5 years	3 years	£14,213
		10 years	Dominated[a]
Use lenalidomide trial RDI	IPD	RDI	£12,224
Use rituximab trial RDI (R^2)	IPD	RDI	£15,328
Use comparator arm RDIs	IPD	RDI	£15,777
Vial costing option	Method of moments	Dose banding	£15,872
		No wastage	£15,980
Exclude wastage of prednisone	Include wastage	Exclude wastage	£15,910
Use of rituximab biosimilar in clinical practice	100%	50.0%	£15,950
		75.0%	£15,929
Use of subcutaneous rituximab during maintenance	100%	90.0%	£15,943
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied	£20,729

Parameter	Base case	Scenario	ICER (£/QALY)
Source for subsequent treatment mean durations	HMRN	AUGMENT	£17,084
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18	£20,048
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.	£16,255
Do not apply age-adjusted utility values	Applied	Not applied	£15,115
Do not apply adverse event disutility values	Applied	Not applied	£15,949
Cap utilities at general population equivalent utility	Not applied	Applied	£17,475
Distribution for R^2 ToT	KM	Exponential	£4,598
		GenGamma	£15,097
		Gompertz	£16,691
		Log-logistic	£15,547
		Log-normal	£12,597
		Weibull	£15,467
Distribution for R-CHOP ToT	KM	Exponential	£16,585
		GenGamma	£15,729
		Gompertz	£16,103
		Log-logistic	£16,682
		Log-normal	£17,320
		Weibull	£16,572

Parameter	Base case	Scenario	ICER (£/QALY)
Distribution for R^2 PFS	KM+comparator hazard	Exponential	£18,079
		GenGamma	£18,919
		Gompertz	£21,312
		Log-logistic	£20,317
		Log-normal	£19,047
		Weibull	£21,782
Distribution for R-CHOP PFS	GenGamma	Exponential	£17,560
		GenGamma	£15,909
		Gompertz	£13,740
		Log-logistic	£13,953
		Log-normal	£13,803
		Weibull	£15,105
Distribution for R^2 OS	Weibull	Exponential	£11,108
		GenGamma	£11,340
		Gompertz	£19,117
		Log-logistic	£14,917
		Log-normal	£13,176
		Weibull	£15,909
Distribution for R-CHOP OS	Weibull	Exponential	£19,719
		GenGamma	£13,581
		Gompertz	£12,842
		Log-logistic	£13,664
		Log-normal	£13,094
		Weibull	£15,909

Parameter	Base case	Scenario	ICER (£/QALY)
Distribution for R^2 TTNLT	Log-normal	Exponential	£15,870
		GenGamma	£15,989
		Gompertz	£16,307
		Log-logistic	£16,175
		Log-normal	£15,909
		Weibull	£16,321
Distribution for R-CHOP TTNLT	Log-normal	Exponential	£15,825
		GenGamma	£15,900
		Gompertz	£15,942
		Log-logistic	£15,678
		Log-normal	£15,909
		Weibull	£15,466

Key: CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; NHS, National Health Service; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; ToT, time on treatment; TTNLT, time to next anti-lymphoma treatment.
Notes: [a]Applying the treatment effect at 10-years causes R^2 to be dominated by R-CHOP as the base case curves for OS cross before this timepoint. It is highly implausible that the survival of R^2 would diminish and become worse than R-CHOP after a significant initial gain.

Table 33: Results of scenario analysis (with PAS) – R^2 versus R-CVP – FL only

Parameter	Base case	Scenario	ICER (£/QALY)
Base case			*£23,746*
Time horizon	40 years	5 years	£50,127
		10 years	£32,837
		20 years	£25,434
No half cycle correction	Applied	Not applied	£23,644
Discount rate for costs	3.50%	0.0%	£23,807
		6.0%	£23,650
Discount rate for QALYs	3.50%	0.0%	£17,774
		6.0%	£28,180
Source of adverse event frequencies: literature	1L trial	Literature	£26,067
Apply comparator hazard to R^2 arms: after	5 years	3 years	£20,816
		10 years	Dominated[a]
Use lenalidomide trial RDI	IPD	RDI	£20,054
Use rituximab trial RDI (R^2)	IPD	RDI	£23,164
Use comparator arm RDIs	IPD	RDI	£23,615
Vial costing option	Method of moments	Dose banding	£23,709
		No wastage	£23,743
Exclude wastage of prednisone	Include wastage	Exclude wastage	£23,747
Use of rituximab biosimilar in clinical practice	100%	50.0%	£23,787
		75.0%	£23,766
Use of subcutaneous rituximab during maintenance	100%	90.0%	£23,780
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied	£28,575

Parameter	**Base case**	**Scenario**	**ICER (£/QALY)**
Source for subsequent treatment mean durations	HMRN	AUGMENT	£24,923
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18	£25,193
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.	£24,264
Do not apply age-adjusted utility values	Applied	Not applied	£22,558
Do not apply adverse event disutility values	Applied	Not applied	£23,760
Cap utilities at general population equivalent utility	Not applied	Applied	£26,088
Distribution for R^2 ToT	KM	Exponential	£12,413
		GenGamma	£22,933
		Gompertz	£24,530
		Log-logistic	£23,384
		Log-normal	£20,428
		Weibull	£23,303
Distribution for R-CVP ToT	KM	Exponential	£24,423
		GenGamma	£23,565
		Gompertz	£23,943
		Log-logistic	£24,520
		Log-normal	£25,153
		Weibull	£24,410
Distribution for R^2 PFS	KM+comparator hazard	Exponential	£26,084
		GenGamma	£26,973
		Gompertz	£29,541
		Log-logistic	£28,472
		Log-normal	£27,111
		Weibull	£30,043

Parameter	**Base case**	**Scenario**	**ICER (£/QALY)**
Distribution for R-CVP PFS	GenGamma	Exponential	£25,546
		GenGamma	£23,746
		Gompertz	£21,448
		Log-logistic	£21,657
		Log-normal	£21,458
		Weibull	£22,886
Distribution for R^2 OS	Weibull	Exponential	£15,854
		GenGamma	£16,236
		Gompertz	£29,021
		Log-logistic	£22,116
		Log-normal	£19,253
		Weibull	£23,746
Distribution for R-CVP OS	Weibull	Exponential	£29,994
		GenGamma	£19,915
		Gompertz	£18,698
		Log-logistic	£20,052
		Log-normal	£19,114
		Weibull	£23,746
Distribution for R^2 TTNLT	Log-normal	Exponential	£23,682
		GenGamma	£23,859
		Gompertz	£24,301
		Log-logistic	£24,120
		Log-normal	£23,746
		Weibull	£24,324

Parameter	Base case	Scenario	ICER (£/QALY)
Distribution for R-CVP TTNLT	Log-normal	Exponential	£23,865
		GenGamma	£23,718
		Gompertz	£23,758
		Log-logistic	£23,482
		Log-normal	£23,746
		Weibull	£23,375
Key: CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; NHS, National Health Service; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; ToT, time on treatment; TTNLT, time to next anti-lymphoma treatment. **Notes**: [a]Applying the treatment effect at 10-years causes R^2 to be dominated by R-CVP as the base case curves for OS cross before this timepoint. It is highly implausible that the survival of R^2 would diminish and become worse than R-CVP after a significant initial gain.			

Table 34: Results of scenario analysis (with PAS) – R^2 versus R-mono – FL only

Parameter	Base case	Scenario	ICER (£/QALY)
Base case			*£20,274*
Time horizon	40 years	5 years	£52,723
		10 years	£25,188
		20 years	£20,592
No half cycle correction	Applied	Not applied	£20,237
Discount rate for costs	3.50%	0.0%	£20,683
		6.0%	£20,058
Discount rate for QALYs	3.50%	0.0%	£15,638

Parameter	Base case	Scenario	ICER (£/QALY)
		6.0%	£23,953
Apply comparator hazard to R^2 arms: after	5 years	3 years	£33,912
		10 years	£12,334
Use lenalidomide trial RDI	IPD	RDI	£17,550
Use rituximab trial RDI (R^2)	IPD	RDI	£19,838
Use comparator arm RDIs	IPD	RDI	£20,492
Vial costing option	Method of moments	Dose banding	£19,822
		No wastage	£19,657
Use of rituximab biosimilar in clinical practice	100%	50.0%	£20,281
		75.0%	£20,278
Subsequent treatment costs for R^2 equal the comparator	Individual	Same costs applied	£22,893
Source for subsequent treatment mean durations	HMRN	AUGMENT	£19,404
Source for Stem cell transplant cost: NHS ref 17/18	NG52	NHS ref 17/18	£20,597
Source for utility values: Wild et al. 2006	AUGMENT	Wild et al.	£19,234
Do not apply age-adjusted utility values	Applied	Not applied	£19,428
Do not apply adverse event disutility values	Applied	Not applied	£20,192
Cap utilities at general population equivalent utility	Not applied	Applied	£21,341
Distribution for R^2 ToT	KM	Exponential	£12,085
		GenGamma	£19,530
		Gompertz	£20,370
		Log-logistic	£19,826
		Log-normal	£18,029
		Weibull	£19,754
Distribution for R-mono ToT	KM	Exponential	£21,670

Parameter	Base case	Scenario	ICER (£/QALY)
		GenGamma	£20,389
		Gompertz	£20,344
		Log-logistic	£20,365
		Log-normal	£20,474
		Weibull	£20,390
Distribution for R^2 PFS	Log-logistic	Exponential	£19,795
		GenGamma	£19,477
		Gompertz	£20,281
		Log-logistic	£20,274
		Log-normal	£19,872
		Weibull	£20,891
Distribution for R-mono PFS	Log-logistic	Exponential	£21,959
		GenGamma	£20,587
		Gompertz	£21,003
		Log-logistic	£20,274
		Log-normal	£20,823
		Weibull	£21,897
Distribution for R^2 OS	Weibull	Exponential	£18,437
		GenGamma	£17,417
		Gompertz	£19,691
		Log-logistic	£20,074
		Log-normal	£19,470
		Weibull	£20,274
Distribution for R-mono OS	Weibull	Exponential	£19,938

Parameter	Base case	Scenario	ICER (£/QALY)
		GenGamma	£17,479
		Gompertz	£24,830
		Log-logistic	£16,911
		Log-normal	£17,191
		Weibull	£20,274
Distribution for R^2 TTNLT	Log-normal	Exponential	£20,232
		GenGamma	£20,254
		Gompertz	£20,537
		Log-logistic	£20,420
		Log-normal	£20,274
		Weibull	£20,545
Distribution for R-mono TTNLT	Log-normal	Exponential	£21,111
		GenGamma	£20,331
		Gompertz	£21,375
		Log-logistic	£20,164
		Log-normal	£20,274
		Weibull	£21,605

Key: FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; ICER, incremental cost-effectiveness ratio; IPD, individual patient data; KM, Kaplan–Meier; mono, monotherapy; NHS, National Health Service; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life-year; R, rituximab; R^2, lenalidomide plus rituximab; RDI, relative dose intensity; SmPC, summary of product characteristics; ToT, time on treatment; TTNLT, time to next anti-lymphoma treatment.

Deterministic sensitivity analysis

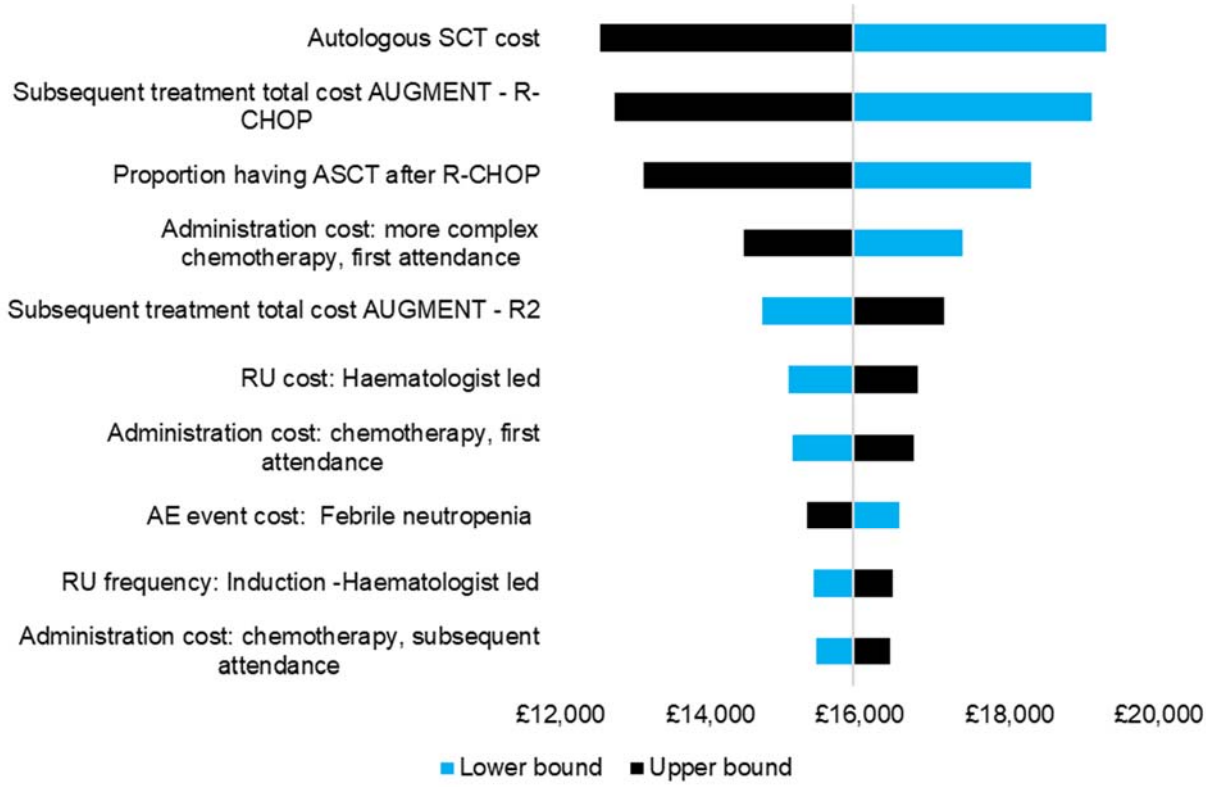
Figure 38–Figure 40 present the tornado diagrams showing the parameters with the greatest impact on the results with descending sensitivity from one-way sensitivity analysis (OWSA).

The parameters that had the largest impact on the ICER for R^2 versus R-CHOP was the cost of SCT, subsequent treatment costs for R-CHOP and proportion of patients who receive SCT. The ICER ranged from £12,522 to £19,295, demonstrating that values tested at their upper and lower bounds still produced ICERs below £30,000. Similarly, with the R-CVP comparison, the parameters that had the largest impact were subsequent treatment costs (including ASCT cost, which is modelled as a subsequent therapy), administration costs and resource use, with all ICERs remaining under £30,000. Results from OWSA from the other plausible parametric distributions (shown in Table 30 and Table 31) are presented in the Appendix (Section 5.1) and show that all ICERs remain under £30,000.

Similar parameters had the largest effect on the ICER for R^2 versus R-mono, but all parameters tested resulted in R^2 being more cost effective than R-mono at the £30,000 willingness to pay threshold.

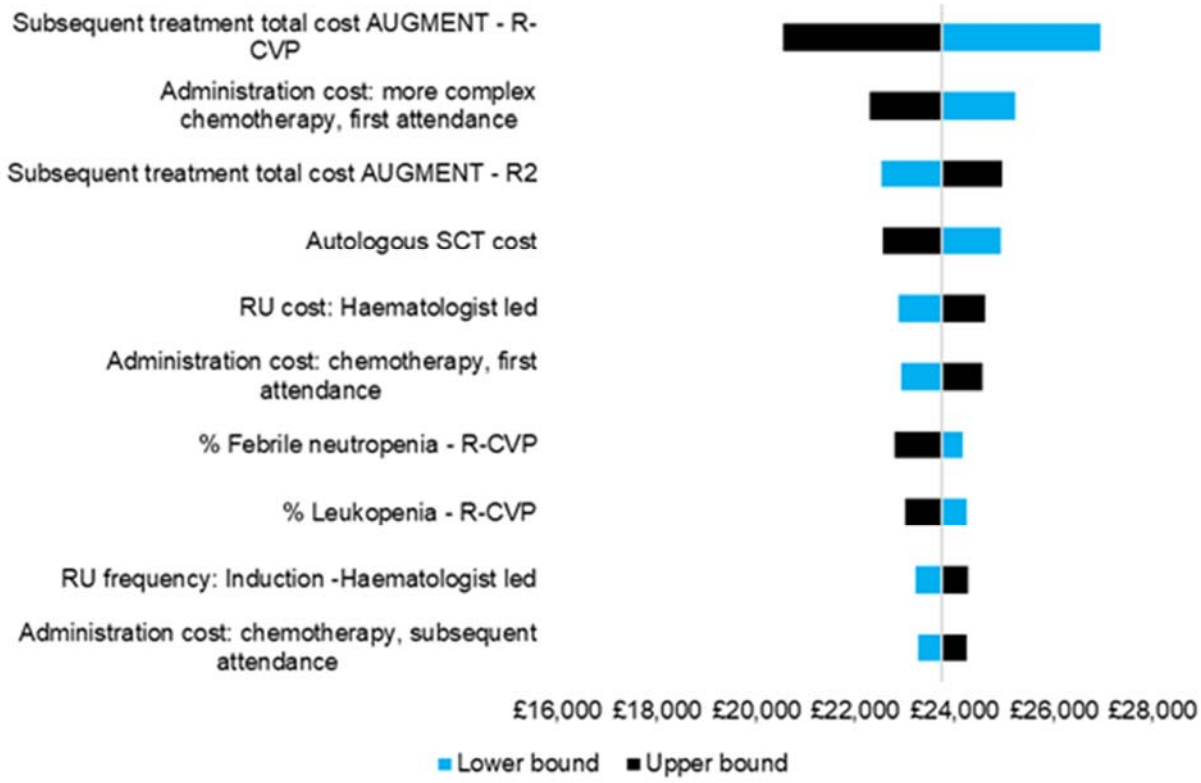
OWSA results for R^2 versus O-Benda in the FL only population are as previously shown in Figure 59 of Document B (Section B.3.8.).

Figure 38: Tornado diagram showing OWSA results on ICER – R^2 versus R-CHOP – FL only



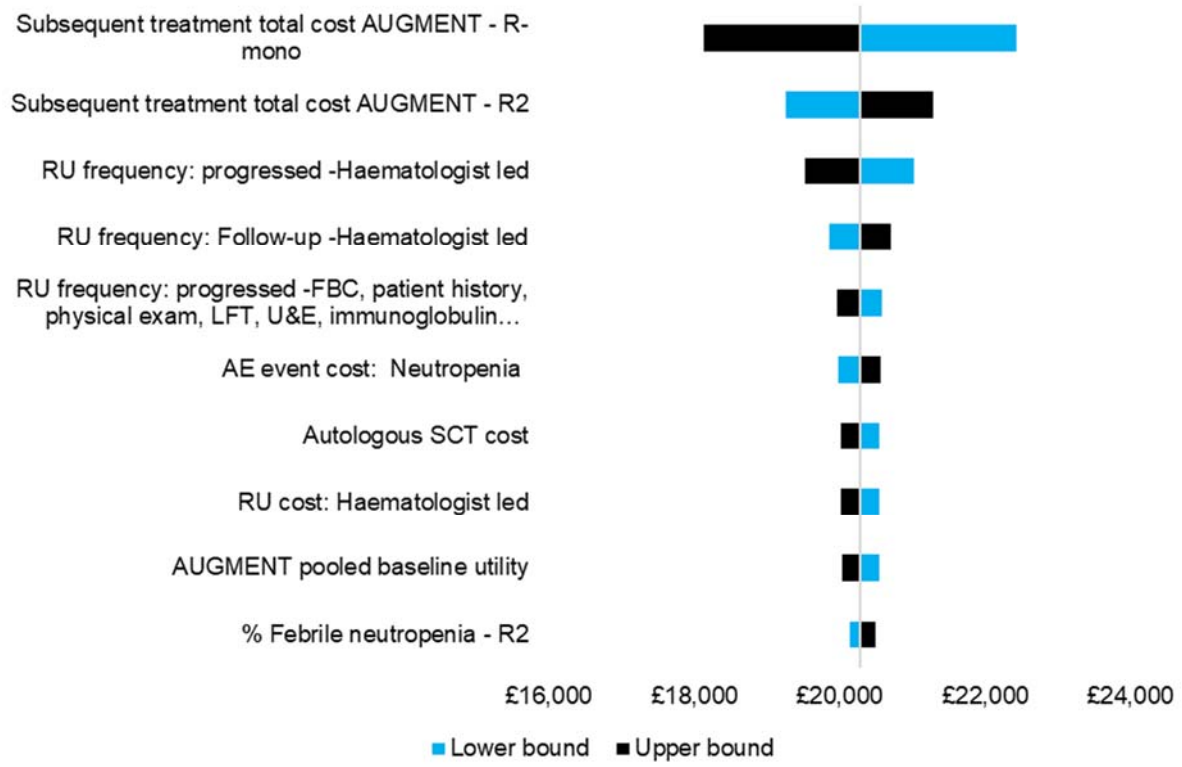
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use.

Figure 39: Tornado diagram showing OWSA results on ICER – R^2 versus R-CVP - FL only



Key: CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 40: Tornado diagram showing OWSA results on ICER – R^2 versus R-mono - FL only



Key: AE, adverse event; FBC, full blood count; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; mono, monotherapy; LFT, liver function tests; OWSA, one-way sensitivity analysis; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant; U&E, urea and electrolytes.

3.3. Interpretation and conclusion of the economic evidence

The deterministic results demonstrate that R^2 is cost-effective versus R-CHOP, R-CVP and R-mono in the FL only population with deterministic ICERs all under the £30,000 willingness to pay threshold. A summary of the results of R^2 versus O-Benda are presented in Section B.3.11 in Document B.

Results of the cost-effectiveness were thoroughly tested in sensitivity analysis testing both parameter and structural uncertainty. All the results from OWSA and the majority of scenario analyses remained under the £30,000 willingness to pay threshold. The majority of the scenarios where the ICER exceeded this were mainly implausible or not appropriate for the decision problem.

Survival extrapolations remains the largest area of uncertainty within the cost-effectiveness results; however, curves were selected based on the most plausible choices and expert opinion where possible. Other plausible curves were tested in scenario analysis resulting in ICERs ranging from £12,061 - £15,150 per QALY

versus R-CHOP and £17,774 - £22,886 per QALY versus R-CVP. The probabilistic results show that R^2 is likely to be cost-effective compared to R-CHOP, R-CVP and R-mono at the £30,000 willingness to pay threshold.

4. References

1. van Oers MH, Klasa R, Marcus RE, 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.
2. 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. *J Clin Oncol*. 2019; 37(14):1188-99.
3. Chiu H, Trisal P, Bjorklund C, et al. Combination lenalidomide-rituximab immunotherapy activates anti-tumour immunity and induces tumour cell death by complementary mechanisms of action in follicular lymphoma. *Br J Haematol*. 2019; 185(2):240-53.
4. Celgene. MAGNIFY - updated response rate tables. 2018. Data on File.
5. Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007) 2018.
6. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – Demographic and Baseline Characteristics for IEE population. 04 June 2019 2018. Data on File.
7. Celgene. CC-5013-NHL-007 (22 Jun 2018 data cut) – PFS post-hoc analysis. 10 October 2019 2018. Data on File.
8. 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*. 2019; 184(4):660-3.
9. Celgene. A Phase 3, double-blind randomized study to compare the efficacy and safety of rituximab plus lenalidomide (CC-5013) versus rituximab plus placebo in subjects with relapsed/refractory indolent lymphoma. (AUGMENT primary analysis (Data cutoff 22 Jun 2018): CC-5013-NHL-007 (updated tables)) 2018. Data on File.
10. National Institute for Health and Care Excellence (NICE) TA513 Obinutuzumab for untreated advanced follicular lymphoma. 2018. Available at: https://www.nice.org.uk/guidance/ta513. Accessed: 05 January 2019.
11. Celgene. CC-5013-NHL-007 (22 Jun 2018 data cut) – FL only RTNLT. Run date: 11 October 2019 2018. Data on File.
12. Celgene. MAGNIFY - updated PFS tables (FL patients, R-refractory split) 10 Aug 2018 2018. Data on File.
13. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – FL only (R-R split) PFS KM plot. 8 October 2019 2018.
14. Celgene. CC-5013-NHL-008 (10 Aug 2018 data cut) – Most common TEAEs. 13 June 2019 2018. Data on file.
15. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics*. 2010; 28(10):935-45.
16. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health*. 2012; 15(6):940-7.

17. Phillippo D, Ades A, Dias S, et al. NICE DSU Technical Support Document 18: Methods for Population-adjusted Indirect Comparisons in Submissions to NICE. 2016. Available at: https://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf. Accessed: 02 May 2019.
18. Morschhauser F, Fowler NH, Feugier P, et al. Rituximab plus Lenalidomide in Advanced Untreated Follicular Lymphoma. *NEJM*. 2018; 379(10):934-47.
19. Fowler N. Frontline strategy for follicular lymphoma: are we ready to abandon chemotherapy? *Hematology Am Soc Hematol Educ Program*. 2016; 2016(1):277-83.
20. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluaions Alongside Clinical Trials - Extrapolation With Patient-Level Data. 2011. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed: 10 May 2019.

5. Appendix

5.1. Efficacy comparison

AUGMENT efficacy comparison

Table 35: Progression-free survival by IRC assessment per 2007 IWGRC with censoring rules based on EMA guidance in AUGMENT – FL versus ITT population

	FL ITT population (n=295)		ITT population (n=358)	
	R^2 (n=147)	**R mono (n=148)**	**R^2 (n=178)**	**R mono (n=180)**
Number of patients, n (%)				
With event				
Censored				
Median PFS (95% CI) (months)[a]				
PFS rate at 6 months (95% CI)				
PFS rate at 1 year (95% CI)				
PFS rate at 2 years (95% CI)				
p-value				
Hazard ratio (95% CI)				

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IRC, Independent Review Committee; ITT, intent-to-treat; IWGRC, International Working Group Response Criteria; NE, not estimable; PFS, progression-free survival; R^2, lenalidomide plus rituximab; R mono, rituximab plus placebo.
Notes: [a], median estimate is from Kaplan–Meier analysis; [b], p-value from log-rank test stratified by three factors: previous rituximab treatment (yes, no), time since last anti-lymphoma therapy (≤2, >2 years), and disease histology (FL, MZL); [c], from Cox proportional hazard model adjusting for the three stratification factors; [d], p-value from log-rank test; [e], from Cox proportional hazard model.
Source: Celgene, 2018.[9]

Table 36: Summary of secondary endpoints in AUGMENT – FL versus ITT population

Endpoint	FL ITT population (n=295)		ITT population (n=358)	
	R^2 (n=147)	**R mono (n=148)**	**R^2 (n=178)**	**R mono (n=180)**
Median OS, months (95% CI)[a]	[redacted]	[redacted]	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	[redacted]		0.61 (0.33, 1.13)[b]	
ORR (CR+PR)	[redacted]	[redacted]	138 (77.5)	96 (53.3)
95% CI[d]	[redacted]	[redacted]	70.7, 83.4	45.8, 60.8
p-value	[redacted]		<0.0001[e]	
CR rate	[redacted]	[redacted]	60 (33.7)	33 (18.3)
95% CI[d]	[redacted]	[redacted]	26.8, 41.2	13.0, 24.8
p-value	[redacted]		0.001[e]	
PR	[redacted]	[redacted]	78 (43.8)	63 (35.0)
SD	[redacted]	[redacted]	20 (11.2)	55 (30.6)
PD/ death	[redacted]	[redacted]	7 (3.9)	23 (12.8)
No evidence of disease	[redacted]	[redacted]	3 (1.7)	4 (2.2)
Unknown/ND/Missing	[redacted]	[redacted]	10 (5.6)	2 (1.1)
Median TTNLT, months (95% CI)[a]	[redacted]	[redacted]	NE (NE, NE)	32.2 (23.2, NE)
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]	73.6 (65.6, 80.1)	57.3 (49.3, 64.5)
Hazard ratio (95% CI)	[redacted]		0.54 (0.38, 0.78)[b]	
p-value	[redacted]		0.0007[g]	
Median EFS, months (95% CI)[a]	[redacted]	[redacted]	27.6 (22.1, NE)	13.9 (11.4, 16.7)
Hazard ratio (95% CI)	[redacted]		0.51 (0.38 to 0.67)[b]	
p-value	[redacted]		<0.0001[g]	
DCRR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
95% CI[i]	[redacted]	[redacted]	[redacted]	[redacted]
p-value	[redacted]		[redacted]	
	R^2 (n=118)	**R mono (n=82)**	**R^2 (n=138)**	**R mono (n=96)**
Median DOR, months (95% CI)[a]	[redacted]	[redacted]	36.6 (22.9, NE)	21.7 (12.8, 27.6)
Hazard ratio (95% CI)[c]	[redacted]		0.53 (0.36 to 0.79)	
p-value[h]	[redacted]		0.0015	
	R^2 (n=51)	**R mono (n=29)**	**R^2 (n=60)**	**R mono (n=33)**

Endpoint	FL ITT population (n=295)		ITT population (n=358)	
	R^2 (n=147)	**R mono (n=148)**	**R^2 (n=178)**	**R mono (n=180)**
Median DOCR, months (95% CI)[a]	[redacted]	[redacted]	[redacted]	[redacted]
Hazard ratio (95% CI)[h]	[redacted]		[redacted]	
p-value[c]	[redacted]		[redacted]	

Key: CI, confidence interval; CR, complete response; DCRR, durable complete response rate, DOCR, duration of complete response; DOR, duration of response; EFS, event-free survival; FL, follicular lymphoma; IRC, Independent Review Committee; ND, not done; NE, not estimable; ORR, overall response rate; OS, overall survival; PD, progressive disease; PR, partial response; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SD, stable disease; TTNLT, time to next anti-lymphoma treatment.
Notes: [a], median estimate is from Kaplan-Meier analysis; [b], from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last anti-lymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c], from Cox proportional hazard model; [d], exact confidence interval for binomial distribution; [e], from CMH test adjusting for the three stratification factors; [f], from Fisher-Exact test; [g], from log-rank test adjusting for the three stratification factors; [h], from log-rank test; [i], exact confidence interval for binomial distribution.
Source: Leonard, et al. 2019; Celgene, 2018.[2, 9]

Table 37: Summary of exploratory endpoints in AUGMENT – FL versus ITT population

Endpoint	FL ITT population (n=295)		ITT population (n=358)	
	R^2 (n=147)	**R^2 (n=178)**	**R^2 (n=178)**	**R mono (n=180)**
Median TTNCT, months (95% CI)[a]	[redacted]	[redacted]	NE (NE, NE)	NE (NE, NE)
TTNCT rate at 2 years, % (95% CI)	[redacted]	[redacted]	84.1 (77.1, 89.1)	67.5 (59.7, 74.1)
Hazard ratio (95% CI)	[redacted]		0.50[g] (0.32, 0.78)	
p-value	[redacted]		0.0017[h]	
Best response by IRC per 1999 IWGRC				
ORR (CR/CRu + PR), n (% [95% CI][d])	[redacted]	[redacted]	138 (77.5 [70.7, 83.4])	96 (53.3 [45.8, 60.8])
p-value	[redacted]		<0.0001[i]	
CR rate (CR/CRu), n (% [95% CI][d])	[redacted]	[redacted]	73 (41.0 [33.7, 48.6])	40 (22.2 [16.4, 29.0])
p-value	[redacted]		0.0002[i]	
CR, n (%)	[redacted]	[redacted]	60 (33.7)	32 (17.8)
CRu, n (%)	[redacted]	[redacted]	13 (7.3)	8 (4.4)
PR, n (%)	[redacted]	[redacted]	65 (36.5)	56 (31.1)

Endpoint	FL ITT population (n=295)		ITT population (n=358)	
	R^2 (n=147)	**R^2 (n=178)**	**R^2 (n=178)**	**R mono (n=180)**
Median PFS2, months (95% CI)[a]	[redacted]	[redacted]	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	[redacted]		0.52[g] (0.32, 0.82)	
p-value	[redacted]		0.0046[h]	
HT, n (% [95% CI][f])	[redacted]	[redacted]	2 (1.1 [0.1, 4.0])	10 (5.6 [2.7, 10.0])
	R2 (n=37)	**R mono (n=70)**	**R2 (n=49)**	**R mono (n=80)**
RTNLT				
ORR, n (% [95% CI][d])	[redacted]	[redacted]	28 (57.1 [42.2, 71.2])	29 (36.3 [25.8, 47.8])
p-value	[redacted]		0.0282[e]	
CR, n (% [95% CI][d])	[redacted]	[redacted]	15 (30.6 [18.3, 45.4])	13 (16.3 [8.9, 26.2])
p-value	[redacted]		0.0775[e]	

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; HT, histological transformation; ITT, intention-to-treat; IWGRC, International Working Group Response Criteria; NE, not estimable; ORR, overall response rate; PFS2, PFS on next anti-lymphoma treatment; PR, partial response; R^2, rituximab plus lenalidomide; R mono, rituximab plus placebo; RTNLT, response rate to next anti-lymphoma treatment; TTNCT, time to next anti-lymphoma chemotherapy treatment.
Notes: [a], median estimate is from Kaplan–Meier analysis; [b], from Cox proportional hazard model. [c], p-value from log-rank test; [d], exact confidence interval for binomial distribution; [e], p-value obtained from Fisher-Exact test; [f], 95% CI is based on the Clopper-Pearson exact method; [g], from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes, no), time since last anti-lymphoma therapy (≤2, >2 year), and disease histology (FL, MZL); [h], p-value from log-rank test adjusted by the three stratification factors; [i], p-value obtained from CMH test adjusting for three stratification factors.
Source: Leonard, et al. 2019[2]; Celgene, 2018.[9, 11]

MAGNIFY efficacy comparison

Table 38: Response rate by best response per 1999 IWGRC in the induction phase in MAGNIFY – FL versus IEE population

Best response in induction phase	FL IEE population (n=247)	IEE population (n=310)
Number of patients (%)		
ORR (CR+CRu+PR), n (% [95% CI[a]])	184 (74.5 [68.6, 79.8])	225 (72.6 [67.3, 77.5])
CR	[redacted]	[redacted]
CRu	[redacted]	[redacted]
CR rate (CR+CRu), n (% [95% CI][a])	114 (46.2 [39.8, 52.6])	138 (44.5 [38.9, 50.2])
PR	[redacted]	[redacted]
SD	[redacted]	[redacted]
PD	[redacted]	[redacted]
Death w/o tumour assessment	[redacted]	[redacted]

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; FL, follicular lymphoma; IEE, induction efficacy evaluable; ORR, overall response rate; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: [a], 95% CI based on the Clopper–Pearson exact method.
Source: Celgene, 2018.[4]

Table 39: Progression-free survival by investigator assessment per 1999 IWGRC with censoring rules based on EMA guidance in MAGNIFY split by R-refractory status – FL induction safety population

	FL induction safety population (n=286)	Induction safety population (n=359)
Number of patients, n (%)		
With event, n (%)	[redacted]	[redacted]
Censored, n (%)	[redacted]	[redacted]
PFS rate at 1 year, % (95% CI)[a]	[redacted]	[redacted]

Key: CI, confidence interval; EMA, European Medicines Agency; FL, follicular lymphoma; IWGRC, International Working Group Response Criteria; R-refractory, rituximab-refractory; PFS, progression-free survival.
Notes: [a], Statistics obtained from Kaplan–Meier method. 95% CI is based on Greenwood formula.
Source: Celgene, 2018.[12]

5.2. Adverse reactions comparison

AUGMENT adverse reactions comparison

Table 40: Summary of treatment-emergent adverse events in AUGMENT – FL versus safety population

	FL safety population		Total safety population	
	R^2 (n=146)	**R mono (n=148)**	**R^2 (n=176)**	**R mono (n=180)**
Number of patients (%)				
Any TEAE	[redacted]	[redacted]	174 (98.9)	173 (96.1)
Len/Pbo related	[redacted]	[redacted]	159 (90.3)	118 (65.6)
R related	[redacted]	[redacted]	132 (75.0)	105 (58.3)
Grade 3–4 TEAE	[redacted]	[redacted]	121 (68.8)	58 (32.2)
Len/Pbo related	[redacted]	[redacted]	101 (57.4)	38 (21.1)
R related	[redacted]	[redacted]	57 (32.4)	19 (10.6)
Grade 5 TEAE	[redacted]	[redacted]	2 (1.1)	2 (1.1)
Any SAE	[redacted]	[redacted]	45 (25.6)	25 (13.9)
Len/Pbo related	[redacted]	[redacted]	23 (13.1)	8 (4.4)
R related	[redacted]	[redacted]	13 (7.4)	3 (1.7)
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]	46 (26.1)	6 (3.3)
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]	112 (63.6)	47 (26.1)
Any TEAE leading to dose interruption of R	[redacted]	[redacted]	60 (34.1)	37 (20.6)
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]	15 (8.5)	9 (5.0)
Any TEAE leading to discontinuation of R	[redacted]	[redacted]	6 (3.4)	2 (1.1)

Key: FL, follicular lymphoma; Len, lenalidomide; Pbo, placebo; R, rituximab; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[9]

Table 41: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT – FL versus safety population

	FL safety population (n=294)		Total safety population (n=356)	
	R^2 (n=146)	R mono (n=148)	R^2 (n=176)	R mono (n=180)
Number of patients (%)				
Blood and lymphatic system disorders	[redacted]	[redacted]	118 (67.0)	58 (32.2)
Neutropenia	[redacted]	[redacted]	102 (58.0)	40 (22.2)
Leukopenia	[redacted]	[redacted]	36 (20.5)	17 (9.4)
Anaemia	[redacted]	[redacted]	28 (15.9)	8 (4.4)
Thrombocytopenia	[redacted]	[redacted]	26 (14.8)	8 (4.4)
Gastrointestinal disorders	[redacted]	[redacted]	115 (65.3)	88 (48.9)
Diarrhoea	[redacted]	[redacted]	55 (31.3)	41 (22.8)
Constipation	[redacted]	[redacted]	46 (26.1)	25 (13.9)
Abdominal pain	[redacted]	[redacted]	22 (12.5)	16 (8.9)
Nausea	[redacted]	[redacted]	20 (11.4)	23 (12.8)
Infections and infestations	[redacted]	[redacted]	110 (62.5)	88 (48.9)
URTI	[redacted]	[redacted]	32 (18.2)	23 (12.8)
Nasopharyngitis	[redacted]	[redacted]	13 (7.4)	18 (10.0)
General disorders and administration site conditions	[redacted]	[redacted]	98 (55.7)	89 (49.4)
Fatigue	[redacted]	[redacted]	38 (21.6)	33 (18.3)
Pyrexia	[redacted]	[redacted]	37 (21.0)	27 (15.0)
Asthenia	[redacted]	[redacted]	24 (13.6)	19 (10.6)
Oedema peripheral	[redacted]	[redacted]	23 (13.1)	16 (8.9)
Skin and subcutaneous tissue disorders	[redacted]	[redacted]	89 (50.6)	43 (23.9)
Pruritus	[redacted]	[redacted]	21 (11.9)	7 (3.9)
Rash	[redacted]	[redacted]	19 (10.8)	7 (3.9)
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	73 (41.5)	58 (32.2)
Muscle spasms	[redacted]	[redacted]	23 (13.1)	9 (5.0)
Back pain	[redacted]	[redacted]	14 (8.0)	18 (10.0)
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]	73 (41.5)	65 (36.1)
Cough	[redacted]	[redacted]	40 (22.7)	31 (17.2)
Dyspnoea	[redacted]	[redacted]	19 (10.8)	8 (4.4)
Investigations	[redacted]	[redacted]	60 (34.1)	50 (27.8)
Alanine aminotransferase increased	[redacted]	[redacted]	18 (10.2)	15 (8.3)

	FL safety population (n=294)		Total safety population (n=356)	
	R^2 (n=146)	**R mono (n=148)**	**R^2 (n=176)**	**R mono (n=180)**
Metabolism and nutrition disorders	[redacted]	[redacted]	58 (33.0)	40 (22.2)
Decreased appetite	[redacted]	[redacted]	23 (13.1)	11 (6.1)
Nervous system disorders	[redacted]	[redacted]	58 (33.0)	39 (21.7)
Headache	[redacted]	[redacted]	26 (14.8)	17 (9.4)
Injury, poisoning and procedural complications	[redacted]	[redacted]	42 (23.9)	40 (22.2)
Infusion related reaction	[redacted]	[redacted]	26 (14.8)	24 (13.3)
Eye disorders	[redacted]	[redacted]	28 (15.9)	14 (7.8)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]	26 (14.8)	9 (5.0)
Tumour flare	[redacted]	[redacted]	19 (10.8)	1 (0.6)
Psychiatric disorders	[redacted]	[redacted]	24 (13.6)	20 (11.1)
Cardiac disorders	[redacted]	[redacted]	21 (11.9)	17 (9.4)
Vascular disorders	[redacted]	[redacted]	21 (11.9)	22 (12.2)

Key: FL, follicular lymphoma; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; URTI, upper respiratory tract infection.
Source: Celgene, 2018.[9]

MAGNIFY adverse reactions comparison

Table 42: Summary of treatment-emergent adverse events in MAGNIFY – FL versus induction safety population

	FL induction safety population (n=286)	Total induction safety population (n=359)
Number of patients (%)		
Any TEAE		
Len related		
R related		
Grade 3–4 TEAE		223 (62.1)
Len related		
R related		
Grade 5 TEAE		
Any SAE		
Len related		
R related		
Any TEAE leading to dose reduction of Len		
Any TEAE leading to dose interruption of Len		
Any TEAE leading to dose interruption of R		
Any TEAE leading to discontinuation of Len		
Any TEAE leading to discontinuation of R		

Key: FL, follicular lymphoma; Len, lenalidomide; R, rituximab; R^2, lenalidomide + rituximab; R mono, rituximab + placebo; SAE, serious adverse event; TEAE, treatment-emergent adverse event.
Source: Celgene, 2018.[9]

Table 43: Most common treatment-emergent adverse events reported in ≥10% of patients by system organ class in MAGNIFY – FL versus induction safety population

	FL induction safety population (n=286)	Total induction safety population (n=359)
Gastrointestinal disorders		
Diarrhoea		126 (35.1)
Nausea		107 (29.8)
Constipation		103 (28.7)
Abdominal pain		

	FL induction safety population (n=286)	Total induction safety population (n=359)
Vomiting	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Fatigue	[redacted]	172 (47.9)
Oedema peripheral	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]
Rash maculo-papular	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Blood and lymphatic system disorders	[redacted]	[redacted]
Neutropenia	[redacted]	143 (39.8)
Anaemia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
URTI	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion-related reaction	[redacted]	[redacted]

Key: FL, follicular lymphoma; URTI, upper respiratory tract infection.
Source: Celgene, 2018.[14]

5.3. Patient characteristics for the AUGMENT versus HMRN comparison

Table 44: Non-R-refractory patient characteristics, observed and match-adjusted, comparing R^2 (AUGMENT) and R-CVP/R-CHOP (HMRN)

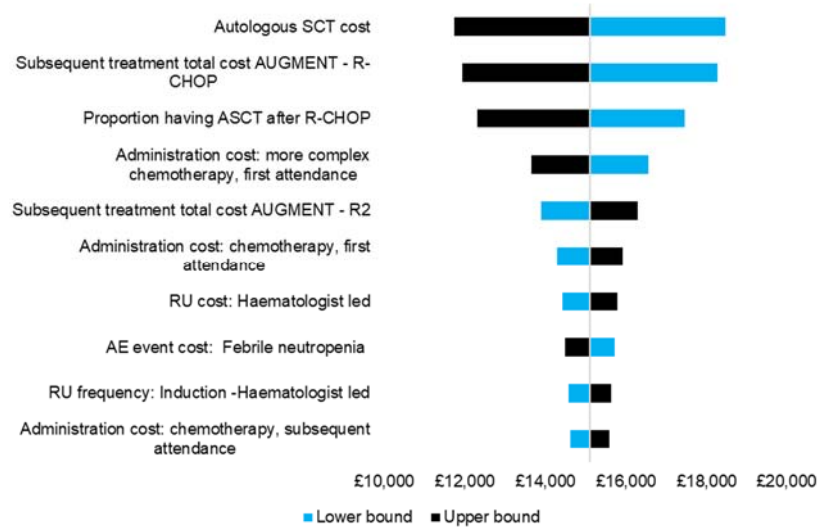
	R^2, observed	R^2, adjusted	R-CVP/R-CHOP
N			
% Prior rituximab			
% Age ≥60yrs			
% Ann Arbor Stage I-II			
% Nodal sites ≤4			
% 1 prior line of therapy			
% 2 prior lines of therapy			
% Early relapse			
% Bone marrow involved			
Weight summaries			
ESS			
Sum (weights)			
Mean (weights)			
Range (weights)			
Number of zero weights			
Number of near zero (<0.01) weights			

Key: ESS, effective sample size; HMRN, Haematological Malignancy Research Network; NA, not applicable; R, rituximab; R2, rituximab plus lenalidomide; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

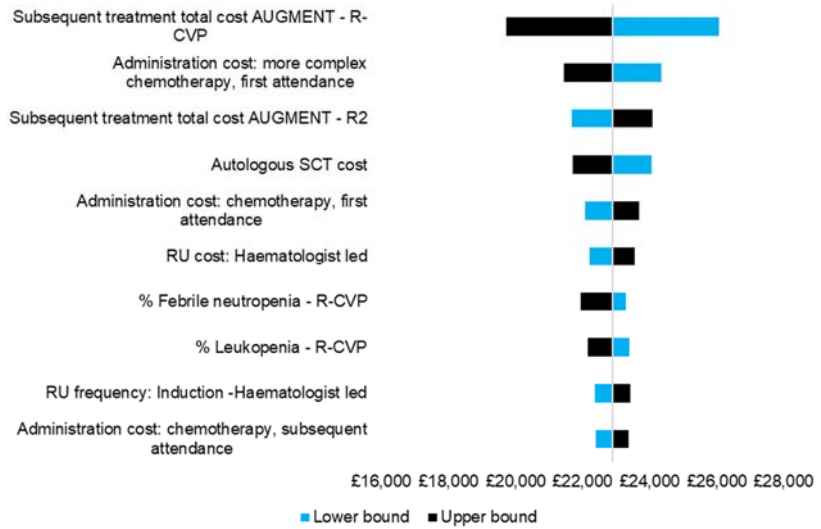
5.4. OWSA for other parametric distribution scenarios

Figure 41: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: Weibull OS R^2, Weibull OS R-CHOP/CVP, Weibull PFS R-CHOP/CVP

R^2 versus R-CHOP



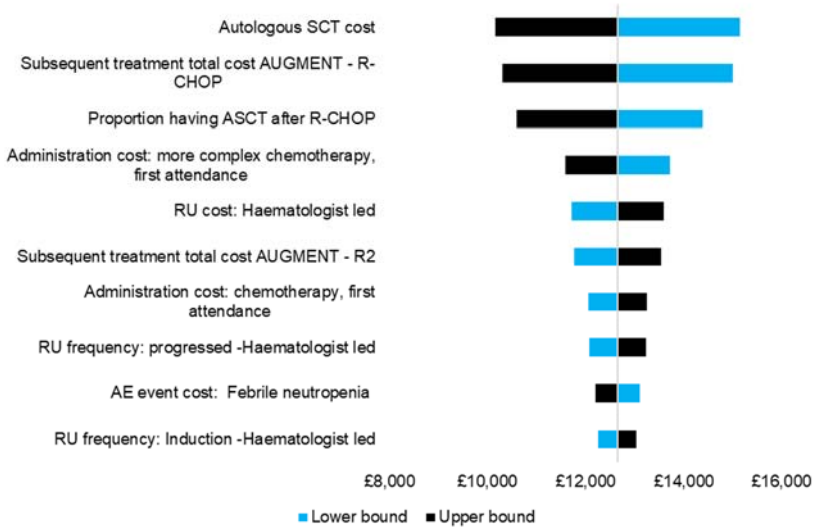
R^2 versus R-CVP



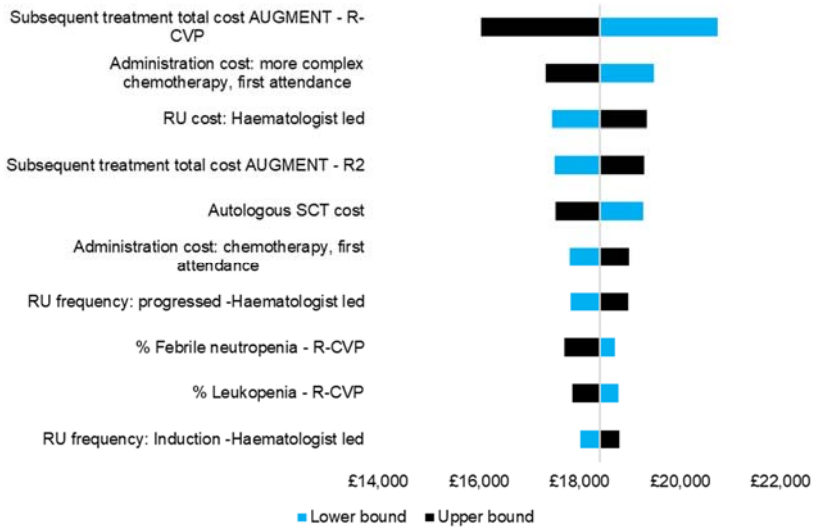
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 42: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: exponential OS R^2, exponential OS R-CHOP/CVP, generalized gamma PFS R-CHOP/CVP

R^2 versus R-CHOP



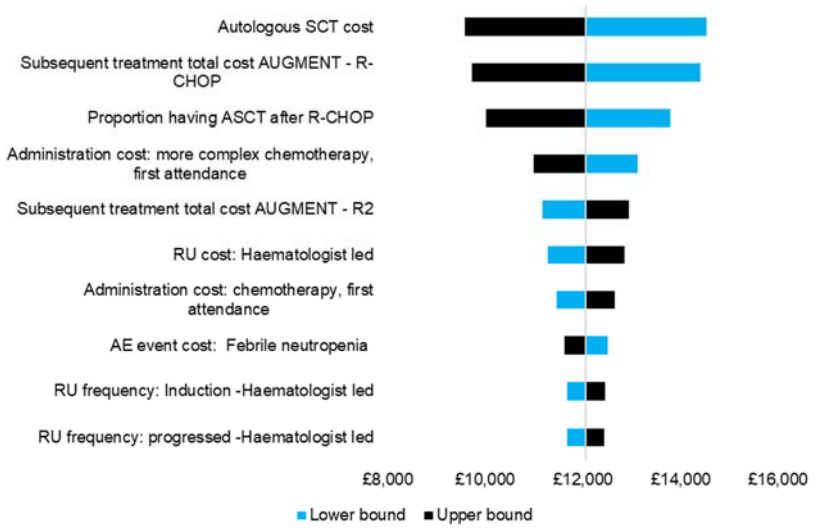
R^2 versus R-CVP



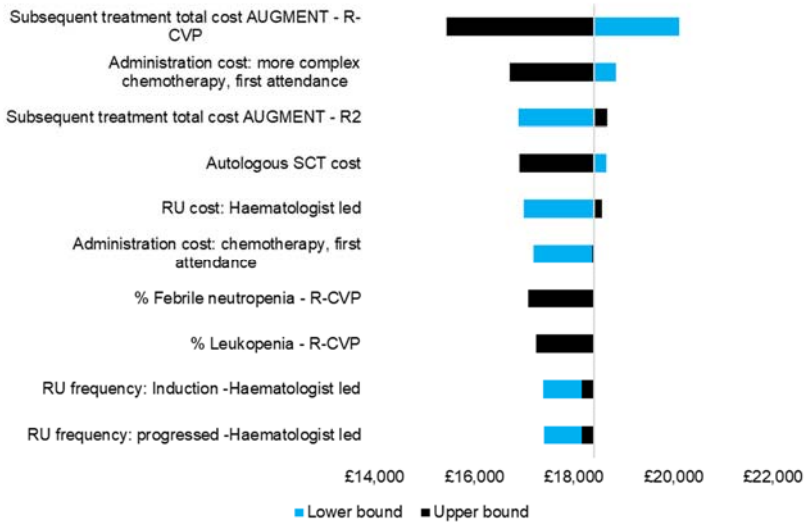
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 43: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: exponential OS R^2, exponential OS R-CHOP/CVP, Weibull PFS R-CHOP/CVP

R^2 versus R-CHOP



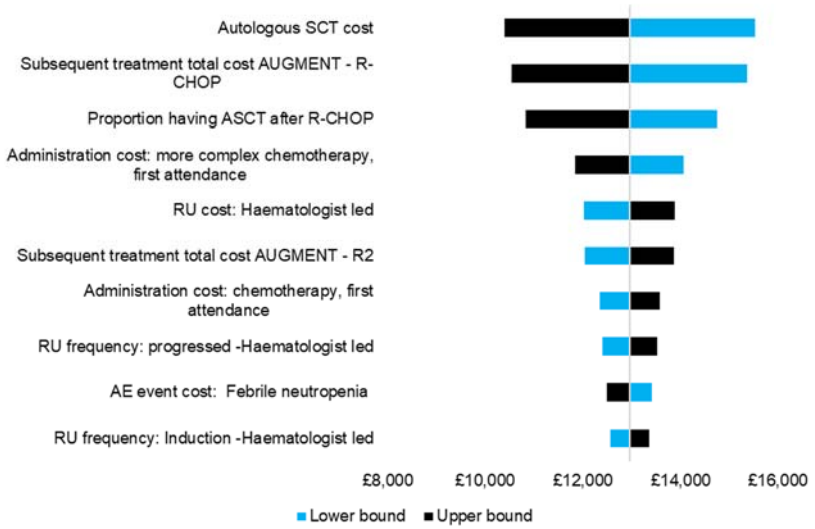
R^2 versus R-CVP



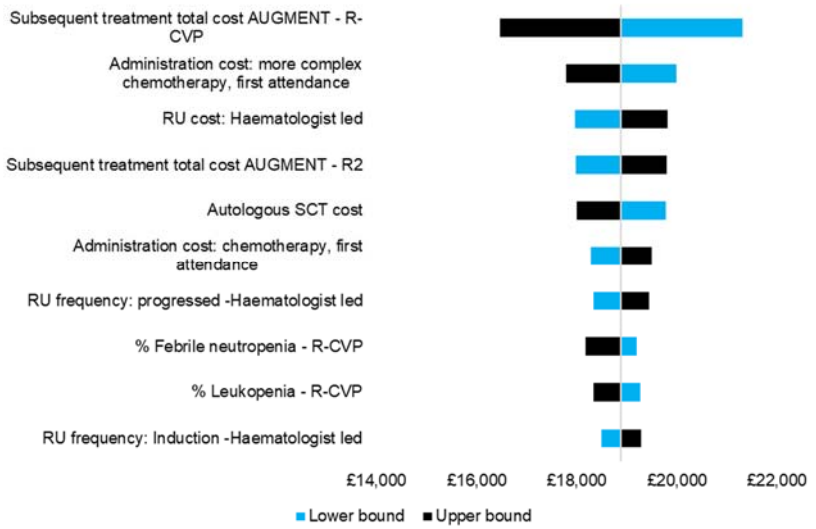
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 44: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: log-logistic OS R^2, log-logistic OS R-CHOP/CVP, generalized gamma PFS R-CHOP/CVP

R^2 versus R-CHOP



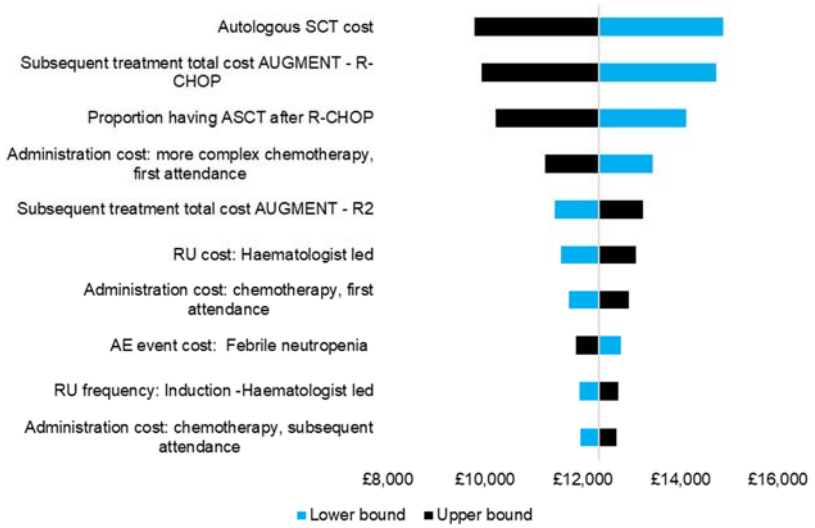
R^2 versus R-CVP



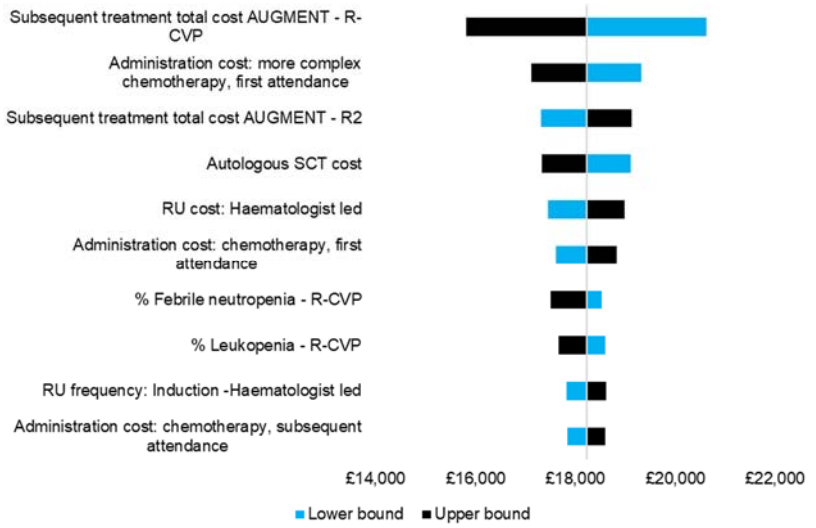
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 45: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: log-logistic OS R^2, log-logistic OS R-CHOP/CVP, Weibull PFS R-CHOP/CVP

R^2 versus R-CHOP



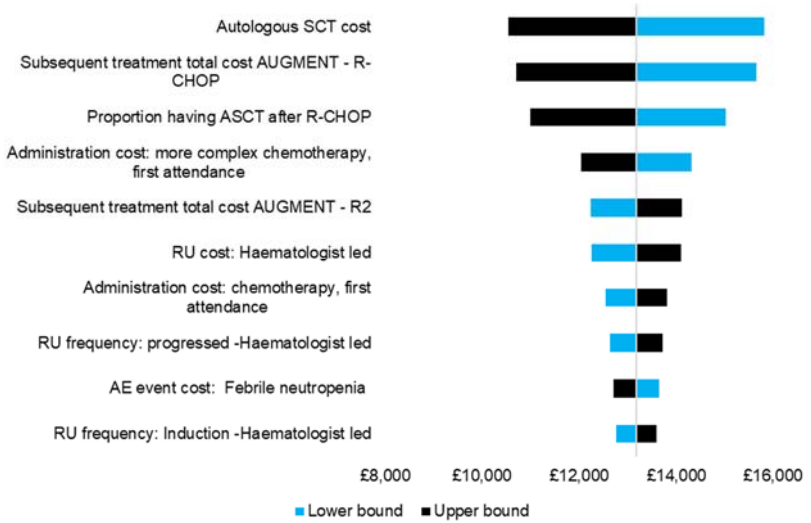
R^2 versus R-CVP



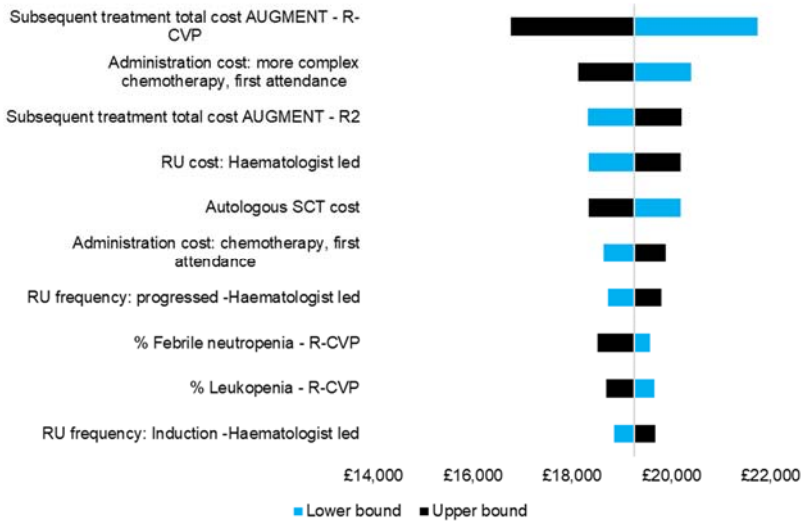
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 46: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: log-normal OS R^2, Weibull OS R-CHOP/CVP, generalized gamma PFS R-CHOP/CVP

R^2 versus R-CHOP



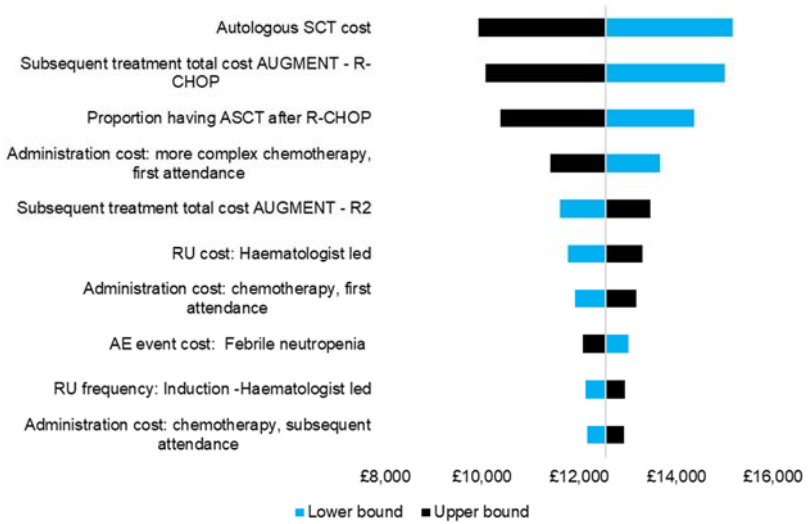
R^2 versus R-CVP



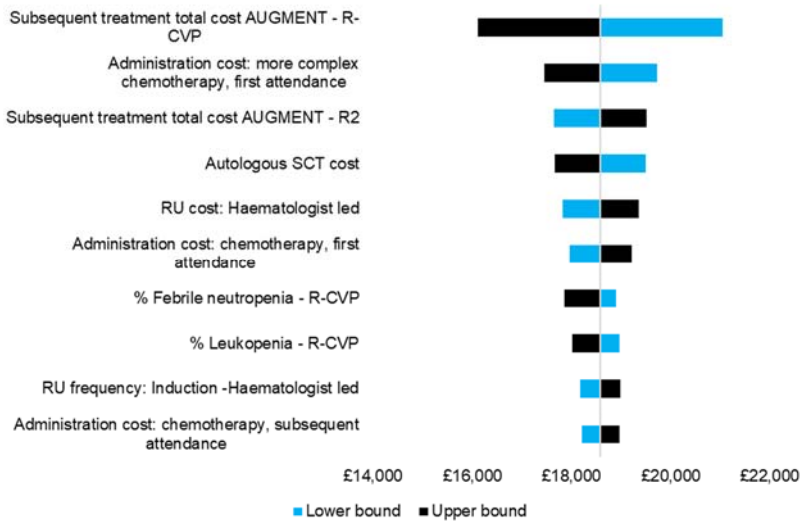
Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

Figure 47: Tornado diagrams showing OWSA results on ICER – FL only – Curve fit selection: log-normal OS R^2, Weibull OS R-CHOP/CVP, Weibull PFS R-CHOP/CVP

R^2 versus R-CHOP



R^2 versus R-CVP



Key: AE, adverse event; ASCT, autologous stem cell transplant; CHOP, cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP, cyclophosphamide, vincristine and prednisolone; FL, follicular lymphoma; ICER, incremental cost-effectiveness ratio; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival; R, rituximab; R^2, lenalidomide plus rituximab; RU, resource use; SCT, stem cell transplant.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma: Addendum to the ERG report in response to company addendum for the amended follicular lymphoma only population

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Rob Riemsma, Reviews Manager, Kleijnen Systematic Reviews Ltd, UK
Thea van Asselt, Health Economist, Maastricht UMC, Netherlands
Willem Witlox, Health Economist, Maastricht UMC
Vanesa Huertas Carrera, Systematic Reviewer, KSR Ltd
Pawel Posadzki, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Steve Ryder, Health Economist, KSR Ltd
Bram Ramaekers, Health Economist, Maastricht UMC
Sabine Grimm, Health Economist, Maastricht UMC
Xavier Pouwels, Health Economist, Maastricht UMC
Gill Worthy, Statistician, KSR Ltd
Steven Duffy, Information Specialist, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 29/10/2019

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 12/82/01.

Declared competing interests of the authors

None.

Acknowledgements

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Riemsma R, Van Asselt ADI, Witlox W, Huertas Carrera V, Posadzki P, Armstrong N, Ryder S, Ramaekers B, Grimm S, Pouwels X, Duffy S, Joore MA, Kleijnen J. Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2019.

Contributions of authors

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Thea van Asselt acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Willem Witlox, Bram Ramaekers, Sabine Grimm, Xavier Pouwels, Steve Ryder and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Vanesa Huertas Carrera and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

There is no change from the ERG report, apart from the changes detailed below.[1]

2. PURPOSE OF THE ADDENDUM TO THE ERG REPORT

In this report, the ERG provides a review of the evidence submitted in an addendum by Celgene in support of lenalidomide (Revlimid®) in combination with rituximab (MabThera®) (R^2), for the treatment of adults with treated follicular lymphoma (FL).[2] As stated in p. 2 of the addendum:

"The expected license wording for R^2 is now: [redacted]

This addendum to the ERG report ('ERG addendum' for brevity) will therefore focus on a modified decision problem, which pertains only to the FL subgroup, as opposed to the total population, constituted of FL and marginal zone lymphoma (MZL), in Section 4 for clinical effectiveness and Section 5 for cost-effectiveness evidence. Where there is no change from the original ERG report this will be noted and no further information or critique will be provided.[1]

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as inferred from the restriction to the FL subgroup)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with treated follicular lymphoma only	Adults with treated follicular lymphoma or marginal zone lymphoma	N/A	This is a change from the original ERG report, which included marginal zone lymphoma.
Intervention	Lenalidomide with rituximab (R^2)	Lenalidomide with rituximab (R^2)	N/A	No change from original ERG report
Comparator(s)	• Rituximab monotherapy (R-mono) • Rituximab in combination with chemotherapy • Established clinical management without lenalidomide (including but not limited to bendamustine)	For non-rituximab refractory patients: • Rituximab in combination with chemotherapy • Rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone (R-CHOP) • Rituximab plus cyclophosphamide vincristine prednisolone (R-CVP) For rituximab refractory patients: • Established clinical management without lenalidomide o Obinutuzumab in combination with bendamustine (O-Benda)	For non-rituximab refractory patients: • R-mono is not considered a relevant comparator as clinical expert opinion confirmed it is rarely used in the relapsed/refractory setting in the UK.[3, 4] For rituximab refractory patients: • O-Benda is included as an option for rituximab-refractory patients under the category 'Established clinical management without lenalidomide'. This is the only NICE-recommended option for this patient group (via the CDF) and clinical experts stated this is the likely treatment choice for FL	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
			patients refractory to rituximab.[3] • Bendamustine monotherapy (Benda mono) is not considered a comparator in this population given that clinical experts believe O-Benda has largely replaced use of Benda mono in rituximab refractory patients.[3]	
Outcomes	The outcome measures to be considered include: • Overall survival • Progression-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life	The outcome measures to be considered include: • Overall survival • Progression-free survival • Event-free survival • Overall response rate • Adverse effects of treatment • Health-related quality of life • Time to next anti-lymphoma treatment • Time to next chemotherapy treatment • Response rate to next anti-lymphoma treatment	Several efficacy outcomes have been presented in addition to those in the scope as several secondary and exploratory outcomes were reported in the AUGMENT and MAGNIFY studies that provide additional insight into the efficacy of R[2]	No change from original ERG report.
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.	Adhering to the reference case, the cost effectiveness of treatments is expressed in terms of incremental cost per quality adjusted life year.	Confidential PAS schemes that apply to relevant subsequent comparator therapies are not included in these analyses as Celgene is not privy to such information	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any PAS for the intervention or comparator technologies will be taken into account.	Adhering to the reference case, a lifetime horizon is used. Adhering to the reference case the economic analyses has been conducted from an NHS and Personal Social Services perspective Adhering to the reference case, the PAS has been applied in all economic analysis for all Celgene products.		
Subgroups to be considered	None listed in scope	No specific subgroups	N/A	
Source: CS, Table 1, pages 7-9. Amended from Evidence Review Group (ERG) report.[1] CDF = Cancer Drugs Fund; FL = follicular lymphoma; MZL = marginal zone lymphoma; NICE = National Institute for Health and Care Excellence.				

3.1 Population

The population defined in the addendum is: adults with previously treated follicular lymphoma only, i.e. it does not include marginal zone lymphoma and is therefore a deviation from the NICE scope. [5]

3.2 Intervention

There is no change from the ERG report.[1]

3.3 Comparators

There is no change from the ERG report.[1]

3.4 Outcomes

There is no change from the ERG report.[1]

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

There is no change from the ERG report.[1]

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Included studies

The company identified three randomised controlled trials (RCTs) of the intervention of interest (lenalidomide in combination with rituximab, R^2): the AUGMENT trial,[6] the MAGNIFY trial,[7] and the ALLIANCE trial;[8] and one non-RCT: Tuscano 2014.[9] The ERG report focused on the AUGMENT trial,[6] because this provides a head-to-head comparison of the intervention of interest (lenalidomide in combination with rituximab, R^2) versus a relevant comparator according to the NICE scope (rituximab monotherapy).[1] The company addendum also presented results for MAGNIFY and again they are ignored for the same reasoning as in the ERG report, i.e. single arm data only.[2, 7]

4.2.2 Methodology of the AUGMENT trial

There is no change from the ERG report.[1]

4.2.3 Baseline characteristics of the AUGMENT trial

Baseline characteristics for patients in the AUGMENT trial, excluding the MZL population, are presented in Table 4.1.

ERG comment: Of note, as with the total population, in the FL subgroup more patients in the R^2 arm than in the R-mono arm were female (58% vs. 46%), had Ann Arbor Stage III to IV disease ([redacted]), FLIPI score ≥3 (37% vs. 31%) and had an ECOG score of 1 or 2 (33% vs. 29%).

Table 4.1: Baseline demographic and disease characteristics, AUGMENT – FL subgroup and ITT population

	FL		Total		Overall (n=358)
	R^2 (n=147)	**R-mono (n=148)**	**R^2 (n=178)**	**R-mono (n=180)**	
Male, n (%)	61 (41.5)	80 (54.1)	75 (42.1)	97 (53.9)	172 (48.0)
Median age, years (range)	62.0 (26.0-86.0)	61.0 (35.0-88.0)	64.0 (26.0-86.0)	62.0 (35.0-88.0)	62.5 (26.0-88.0)
Age distribution, n (%)					
<65	86 (58.5)	94 (63.5)	96 (53.9)	107 (59.4)	203 (56.7)
≥65	61 (41.5)	54 (36.5)	82 (46.1)	73 (40.6)	155 (43.3)
≥70	34 (23.1)	32 (21.6)	47 (26.4)	44 (24.4)	91 (25.4)
Race, white (%)	[redacted]	[redacted]	118 (66.3)	115 (63.9)	233 (65.1)
Histology (investigator review), n (%)					
FL	[redacted]	[redacted]	147 (82.6)	148 (82.2)	295 (82.4)
Grade 1	[redacted]	[redacted]	50 (28.1)	62 (34.4)	112 (31.3)
Grade 2	[redacted]	[redacted]	75 (42.1)	61 (33.9)	136 (38.0)
Grade 3a	[redacted]	[redacted]	22 (12.4)	25 (13.9)	47 (13.1)
MZL	N/A	N/A	31 (17.4)	32 (17.8)	63 (17.6)
MALT	N/A	N/A	14 (7.9)	16 (8.9)	30 (8.4)
Nodal	N/A	N/A	8 (4.5)	10 (5.6)	18 (5.0)
Splenic	N/A	N/A	9 (5.1)	6 (3.3)	15 (4.2)
Ann Arbor stage, n (%)					
I	[redacted]	[redacted]	15 (8.4)	18 (10.0)	33 (9.2)
II	[redacted]	[redacted]	26 (14.6)	38 (21.1)	64 (17.9)
III	[redacted]	[redacted]	73 (41.0)	65 (36.1)	138 (38.5)
IV	44 (29.9)	46 (31.1)	64 (36.0)	59 (32.8)	123 (34.4)

	FL		Total		Overall (n=358)
	R^2 (n=147)	R-mono (n=148)	R^2 (n=178)	R-mono (n=180)	
Low (0,1)	[redacted]	[redacted]	52 (29.2)	67 (37.2)	119 (33.2)
Intermediate (2)	[redacted]	[redacted]	55 (30.9)	58 (32.2)	113 (31.6)
High (≥3)	[redacted]	[redacted]	69 (38.8)	54 (30.0)	123 (34.4)
0	99 (67.3)	105 (70.9)	116 (65.2)	128 (71.1)	244 (68.2)
1	47 (32.0)	42 (28.4)	60 (33.7)	50 (27.8)	110 (30.7)
2	[redacted]	[redacted]	2 (1.1)	2 (1.1)	4 (1.1)
Yes	39 (26.5)	43 (29.1)	45 (25.3)	49 (27.2)	94 (26.3)
No	107 (72.8)	105 (70.9)	132 (74.2)	131 (72.8)	263 (73.5)
Yes	77 (52.4)	68 (45.9)	97 (54.5)	86 (47.8)	183 (51.1)
No	70 (47.6)	80 (54.1)	81 (45.5)	94 (52.2)	175 (48.9)
1	[redacted]	[redacted]	102 (57.3)	97 (53.9)	199 (55.6)
>1	[redacted]	[redacted]	76 (42.7)	83 (46.1)	159 (44.4)
Yes	26 (17.7)	25 (16.9)	30 (16.9)	26 (14.4)	56 (15.6)
No	121 (82.3)	123 (83.1)	148 (83.1)	154 (85.6)	302 (84.4)
Yes	[redacted]	[redacted]	56 (31.5)	61 (33.9)	117 (32.7)
No	[redacted]	[redacted]	122 (68.5)	118 (65.6)	240 (67.0)

Source: CS, Table 5, pages 34-35. FL data also reproduced in company addendum Table 1.[2]

ECOG = Eastern Cooperative Oncology Group; FL = follicular lymphoma; FLIPI = follicular lymphoma international prognostic index; GELF = Groupe d'Etude des Lymphomes Folliculaires; LDH = lactate dehydrogenase; MZL = marginal zone lymphoma; MALT = mucosa associated lymphatic tissue; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo.

Notes: [a]) POD24 is defined as relapse within two years of initial chemoimmunotherapy.

4.2.4 Statistical analyses of the AUGMENT trial

There is no change from the ERG report.[1]

4.2.5 Results of the AUGMENT trial (FL subgroup)

As presented in the ERG report, the data presented in the CS are based on the 22 June 2018 data cut-off for the primary analysis.[1, 2] Efficacy analyses were conducted in the ITT population and based on data from IRC review, using the modified 2007 IWGRC. EMA censoring rules were applied to the analyses.[2]

Table 4.2 presents a summary of the main results for FL, which were also reported in Appendix 1 of the ERG report.[1]

Table 4.2: Summary of results from the AUGMENT trial: FL subgroup

Endpoint	**FL**	
	R^2 (n=147)	**R-mono (n=148)**
Median OS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	0.45 (0.22, 0.92)[e]	
Median PFS, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted][c]	
Best response, n (%)		
ORR (CR+PR)	118 (80.3)	82 (55.4)
95% CI[d]	72.9, 86.4	47.0, 63.6
p-value	<0.0001[f]	
CR rate	51 (34.7)	29 (19.6)
95% CI[d]	27.0, 43.0	13.5, 26.9
p-value	0.0040[f]	
PR	67 (45.6)	53 (35.8)
SD	14 (9.5)	44 (29.7)
PD/ death	7 (4.8)	19 (12.8)
No evidence of disease	3 (2.0)	2 (1.4)
Unknown/ND/Missing	5 (3.4)	1 (0.7)
Median TTNLT, months (95% CI)[a]	NE (NE, NE)	28.2 (20.8, NE)
TTNLT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	0.43 (0.29, 0.65)[c]	
p-value	<0.0001[h]	
Median EFS, months (95% CI)[a]	[redacted]	13.8 (11.0, 16.0)
Hazard ratio (95% CI)	0.42 (0.31, 0.58)[c]	
p-value	<0.0001[h]	
Median TTNCT, months (95% CI)[a]	[redacted]	[redacted]
TTNCT rate at 2 years, % (95% CI)	[redacted]	[redacted]
Hazard ratio (95% CI)	[redacted]	
p-value	[redacted]	

[redacted]		
ORR, n (% [95% CI][d])	[redacted]	[redacted]
p-value	[redacted]	
CR, n (% [95% CI][d])	[redacted]	[redacted]
p-value	[redacted]	
DCRR, n/N (%)	[redacted]	[redacted]
95% CI[d]	[redacted]	[redacted]
p-value	[redacted]	
N, Median DOR, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)[c]	[redacted]	
p-value[e]	[redacted]	
N, Median DOCR, months (95% CI)[a]	[redacted]	[redacted]
Hazard ratio (95% CI)[h]	[redacted]	
p-value	[redacted]	

Source: Response to CL, Table 5, pages 19 and 20. These results were also presented in the company addendum, Tables 3 and 4.[2]

CI = confidence interval; CR = complete response; DCRR = durable complete response rate, DOCR = duration of complete response; DOR = duration of response; EFS = event-free survival; FL = follicular lymphoma; IRC = Independent Review Committee; ITT = intent-to-treat; MZL = marginal zone lymphoma; ND = not done; NE = not estimable; ORR = overall response rate; OS = overall survival; PD = progressive disease; PR = partial response; R^2 = lenalidomide plus rituximab; R-mono = rituximab plus placebo; RTNLT = response rate to next anti-lymphoma treatment; SD = stable disease; TTNLT = time to next anti-lymphoma treatment; TTNCT = time to next anti-lymphoma chemotherapy treatment.

Notes: [a]) median estimate is from Kaplan–Meier analysis; [b]) from Cox proportional hazard model adjusting for the three stratification factors: previous rituximab treatment (yes; no), time since last anti-lymphoma therapy (≤2; >2 year), and disease histology (FL; MZL). [c]) from Cox proportional hazard model; [d]) exact confidence interval for binomial distribution; [e]) from CMH test adjusting for the three stratification factors; [f]) from Fisher-Exact test; [g]) from log-rank test adjusting for the three stratification factors; [h]) from log-rank test

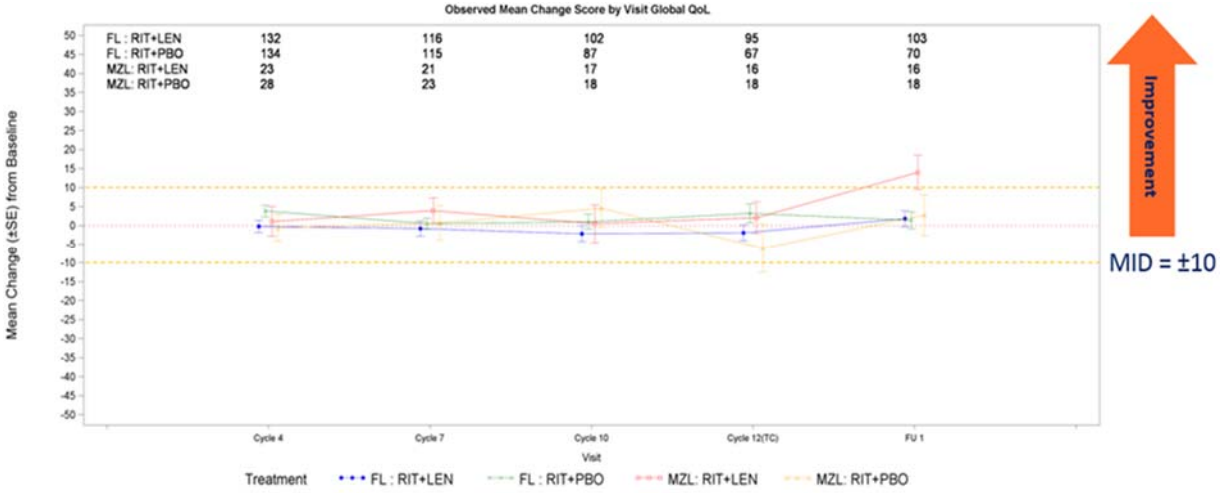
Overall, R^2 showed favourable results when compared to R-mono for PFS with a greater median PFS ([redacted]). There was also evidence of a difference in overall survival (OS) with a hazard ratio (HR) of [redacted]) for patients treated with R^2 compared to R-mono. Overall response rate (ORR) was significantly greater for R^2 compared with R-mono ([redacted]). The complete response (CR) rate was also greater for the R^2 arm compared with R-mono [redacted]

ERG comment: As was identified in the ERG report, the results for R^2 versus R-mono in FL patients are generally more favourable than in MZL patients and therefore than in the original population.[1]

Health-related quality of life (HRQoL)

The company addendum stated that details of the primary HRQoL endpoint, the global health status/quality of life (GHS/QoL) domain of the EORTC QLQ-C30, were presented in Appendix B. The company concluded: "Between-group differences in mean changes were small and not clinically meaningful across all assessment visits and did not differ between FL and MZL patients." (p.236), with results presented in Figure 4.1.[10]

Figure 4.1: Cross-sectional assessment of global health status/quality of life changes from baseline



Source: CS, Appendix P, Figure 23, page 238.
FU = follow-up; MID = minimally important difference; Len = lenalidomide; PBO = placebo; Rit = rituximab; SE = standard error; TC = treatment completion.

ERG comment

The ERG concludes, as in the ERG report, that there was no clinically meaningful change from baseline in the GHS/QoL domain of the QLQ-C30 was observed across any of the post-baseline assessment visits, regardless of treatment group.[1]

4.2.6 Adverse events

As stated in the ERG report, adverse event data from the AUGMENT trial were taken from the 22 June 2018 database cut-off; safety analyses were conducted in the safety population.[1, 2]

A summary of the treatment-emergent adverse event (TEAEs) during AUGMENT for the FL only population is presented in Table 4.3: this was originally presented in Appendix 1 of the ERG report.[1] TEAEs were reported in [redacted] patients [redacted] in the R^2 arm and [redacted] patients [redacted] in the R-mono arm. More patients in the R^2 arm ([redacted] experienced a Grade 3 or 4 TEAE compared with those in the R-mono arm ([redacted] and [redacted] patient in each treatment arm reported a Grade 5 TEAE. Additionally, a greater proportion of patients reported serious adverse events in the R^2 arm [redacted] compared with those in the R-mono arm ([redacted]

Table 4.3: Summary of treatment-emergent adverse events in AUGMENT: FL safety population

	Total population (FL only)	
	R^2 (n=146)	R-mono (n=148)
Number of patients (%)		
Any TEAE	[redacted]	[redacted]
Len related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 3–4 TEAE	[redacted]	[redacted]
Len related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Grade 5 TEAE	[redacted]	[redacted]

Any SAE	[redacted]	[redacted]
Len related	[redacted]	[redacted]
R related	[redacted]	[redacted]
Any TEAE leading to dose reduction of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to dose interruption of R	[redacted]	[redacted]
Any TEAE leading to discontinuation of Len/Pbo	[redacted]	[redacted]
Any TEAE leading to discontinuation of R	[redacted]	[redacted]

Source: Company addendum, Table 8[2]
Len = lenalidomide; Pbo = placebo; R = rituximab; R^2 = lenalidomide + rituximab; R mono = placebo, rituximab + placebo; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

In the safety population, TEAEs that occurred more frequently (≥10% difference) in the R^2 arm than the R-mono arm included the following: neutropenia ([redacted] vs. [redacted]%), diarrhoea [redacted] vs. [redacted] constipation [redacted] vs. [redacted] cough ([redacted] vs. [redacted] upper respiratory tract infection ([redacted] vs. [redacted] and leukopenia [redacted] vs. [redacted] (see Table 4.4). These were originally presented in Appendix 1 of the ERG report.[1]

The difference in the number of Grade 3 or 4 TEAEs between treatment arms (shown in Table 4.7) was largely driven by Grade 3 or 4 events of neutropenia and leukopenia. Neutropenia occurred in [redacted] patients ([redacted] in the R^2 arm compared with [redacted] patients ([redacted] in the R-mono arm, and leukopenia occurred in [redacted] patients ([redacted] in the R^2 arm compared with [redacted] patients ([redacted] in the R-mono arm.

The most common TEAEs, occurring in more than 10% of patients, are presented in Table 4.4 below.

Table 4.4: Most common treatment-emergent adverse events reported in ≥10% of patients in either treatment arm in AUGMENT: FL safety population

	FL subgroup	
	R^2 (n=146)	**R-mono (n=148)**
Number of patients (%)		
Blood and lymphatic system disorders	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Gastrointestinal disorders	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Constipation	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
URTI	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
General disorders and administration site conditions	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]

	FL subgroup	
	R^2 (n=146)	R-mono (n=148)
Pyrexia	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]
Oedema peripheral	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]
Rash	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]
Back pain	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Cough	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Investigations	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts and polyps)	[redacted]	[redacted]
Tumour flare	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]

Source: Company addendum, Table 9[2]
R^2 = lenalidomide + rituximab; R-placebo = rituximab + placebo; URTI = upper respiratory tract infection.

ERG comment: As reported in the ERG report for the total population and shown in Table 4.7, R^2 was associated with more grade 3-4 TEAEs and SAEs when compared to R-mono, especially lenalidomide/placebo related adverse events; but rituximab-related grade 3-4 TEAEs and SAEs were also more frequent in the R^2 arm than in the R-mono arm.[1] R^2 was also associated with more TEAEs leading to dose reductions, dose interruptions and discontinuations of lenalidomide/placebo or rituximab when compared to R-mono.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

There is no change from the ERG report.[1]

4.4 Critique of the indirect comparison and/or multiple treatment comparison

As for the total population, for the FL only population the company performed an indirect comparison using two types of source for the comparator data.

Firstly, the company performed an indirect comparison with data from published evidence. [1, 2] This included comparisons of R^2 with:

- R-CHOP for non-rituximab refractory patients, based on comparator data from a study by Van Oers et al. (2006)[11] comparing R-CHOP with CHOP (only the R-CHOP arm was used in the analyses).
- Established clinical management without lenalidomide, i.e. O-Benda for rituximab refractory patients, based on comparator data from a study by Sehn et al. (2016)[12] comparing O-Benda with bendamustine monotherapy (only the O-Benda arm was used in the analyses).

No further critique of the literature-based ITC is presented in the ERG addendum. This is because:

- The company did not perform any analyses for the FL subgroup versus R-CHOP because "The ERG has discounted the comparison with the Van Oers publication as a scenario due to the population not being relevant to UK clinical practice,…".[13]
- As explained in the ERG report, NICE does not consider O-Benda a relevant comparator for disease that is refractory to rituximab.[1] Therefore, as in the ERG report, this comparison will be ignored.

Secondly, the company performed an indirect comparison with data from the Haematological Malignancy Research Network (HMRN) This included a comparison of R^2 with:

- Pooled data for R-CHOP/R-CVP for non-rituximab refractory patients.

4.4.1 Indirect comparison of R^2 with R-CHOP/R-CVP based on HMRN data.

In the addendum the company stated that the methodology for the indirect comparison, i.e. a matched adjusted indirect comparison (MAIC), was the same as in the original company submission (CS). Therefore, critique of the methodology is necessary in this section only where there has been a change from the original CS, as critiqued in the ERG report.[1, 2]

As described in the ERG report and the company addendum, all subjects in the HMRN data had FL and there were 63 patients identified as receiving either R-CVP or R-CHOP as second- or later-line therapy.[1, 13] The baseline characteristics that were commonly collected by the HMRN and the AUGMENT FL subgroup are presented in Table 4.5.

Table 4.5: Covariates commonly collected across AUGMENT (FL subgroup) and HMRN datasets

Data source	HMRN	AUGMENT (FL subgroup)
Treatment	**R-CVP/R-CHOP (2L+ population)**	**R^2**
N	[redacted]	[redacted]
Age (years):		
Median	[redacted]	62
Range	[redacted]	[redacted]
n (%) Age >=60yrs	[redacted]	[redacted]
n (%) Age >=65yrs	[redacted]	61 (41.5%)

Data source	HMRN	AUGMENT (FL subgroup)
Sex, n, %		
n (%) Males	[redacted]	61 (41.5)
n (%) Females	[redacted]	86 (58.5)
Number of prior systemic anti-lymphoma regimens:		
n (%) 1	[redacted]	[redacted]
n (%) 2	[redacted]	[redacted]
n (%) ≥ 3	[redacted]	[redacted]
Prior rituximab treatment, n (%)	[redacted]	[redacted]
POD24[a], n (%)	[redacted]	[redacted]
Fully Staged, n (%)	[redacted]	NA
Bone marrow involved, n (%)	[redacted]	[redacted]
Nodal sites		
n (%) ≤4	[redacted]	[redacted]
n (%) >4	[redacted]	[redacted]
Bulky disease [b]	[redacted]	[redacted]
Stage		
n (%) I	[redacted]	[redacted]
n (%) II	[redacted]	[redacted]
n (%) III	[redacted]	[redacted]
n (%) IV	[redacted]	44 (29.9)

Source: Company addendum and ERG report (HMRN only)[1, 13]
2L+ = second or later line therapy; HMRN = Haematological Malignancy Research Network; NA = not applicable; R^2 = rituximab plus lenalidomide; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP = rituximab plus cyclophosphamide vincristine prednisolone.
[a]) POD24 is defined as relapse within two years of initial chemoimmunotherapy.
[b]) Bulky disease has different definitions in AUGMENT and the HMRN dataset. AUGMENT: At least one lesion that is ≥ 7 cm or at least 3 lesions with 3 cm or larger in the longest diameter by investigator review. HMRN: At least one lesion that is ≥ 10 cm.

The same list of potential modifiers/prognostic variables as used for the whole mixed MZL/FL subgroup was used to identify the matching variables for this comparison.[2] Therefore, as in the ERG report, Table 4.6 shows the same list of potential effect modifiers/prognostic variables (EM/PVs) that would ideally be adjusted for in a MAIC, as identified and validated by external clinical experts consulted by the company.[1]

Table 4.6: Potential EM/PVs that would ideally be adjusted for in a MAIC

Characteristic	highest priority	Included in MAIC
Previous exposure to rituximab	Yes	Yes
FLIPI components: - Age (mean, or median if mean no reported, or % >60 years if neither reported) - Ann Arbor Stage (III-IV) - Nodal sites (>4) - High LDH	Yes	Yes (as % Age ≥60yrs) Yes Yes
Refractory to last therapy	Yes	No
Prior lines of therapy 1 vs. 2 vs. >2	Yes	Yes
FLIPI risk group (low vs. intermediate vs. high)	Yes	No
FLIPI2+ components: - Serum beta-2 microglobulin high - Bone marrow involvement - Diameter of largest node >6 cm - Haemoglobin <12 dL/L		Yes
Time from last treatment		No
POD24	Yes	Yes
ECOG performance status (0–1 vs. 2+)		No
Presence of B-symptoms		No

Source: ERG report, Table 4.12[1]
EM = effect modifiers; ESS = effective sample size; FL = follicular lymphoma; FLIPI = Follicular Lymphoma International Prognostic Index; MAIC = matching-adjusted indirect comparison; MZL = marginal zone lymphoma; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone. Notes: Adjusted N is the sum of the absolute weights. The patient characteristics presented are the potential EM/PVs that were included in the matching. The following potential EM/PVs had data for all included studies but were dropped from the matching to maintain a sufficiently large effective sample size for subsequent analysis: % previous rituximab exposure. The ESS and adjusted N including these variables were 4.2 and 0.1.

As can be seen from Table 4.6, matching was performed for the following variables:

- Age ≥60 years (FLIPI component)
- Ann Arbor Stage III-IV (FLIPI component)
- Nodal sites >4 (FLIPI component)
- Prior rituximab treatment
- Prior lines of therapy (1 vs. 2 vs. >2)
- POD24 status
- Bone marrow involvement

The results of the matching for the EMs/PVs included in the matching are provided in Table 4.7 for all covariates included.

Table 4.7: Patient characteristics, observed and match-adjusted for the non-R-refractory population, comparing R^2 (AUGMENT FL subgroup) and R-CHOP/R-CVP (HMRN)

Characteristic	AUGMENT FL subgroup (R^2) (n=[redacted])	HMRN (R-CHOP/ R-CVP) (n=[redacted])	Adjusted R^2 (n=[redacted])
Patient characteristics			
% Prior rituximab	[redacted]	[redacted]	[redacted]
% Age ≥60yrs	[redacted]	[redacted]	[redacted]
% Ann Arbor stage III-IV	[redacted]	[redacted]	[redacted]
% Nodal sites ≤4	[redacted]	[redacted]	[redacted]
% 1 prior lines of therapy	[redacted]	[redacted]	[redacted]
% 2 prior lines of therapy	[redacted]	[redacted]	[redacted]
% Early relapse	[redacted]	[redacted]	[redacted]
% Bone marrow involved	[redacted]	[redacted]	[redacted]
Outcomes			
OS	Not estimable	63, [redacted] (NR)	NR
PFS (N, median (95% CI))	178, 39.4 months (NR)	63, [redacted] (NR)	NR

Source: Company addendum, Table 13 and Table 44[2]
EM = effect modifiers; ESS = effective sample size; FLIPI = Follicular Lymphoma International Prognostic Index; NR = not reported; PV = prognostic variables; R = rituximab; R^2 = lenalidomide plus rituximab; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone.

ERG comment: One important addition to the variables included previously is bone marrow involvement. In the ERG report it was stated that the company were asked why 'bone marrow involved' was not included in the matching. The company responded that 'bone marrow involved' should have been considered as a matching variable given that it was identified as being a potential prognostic factor and/or treatment effect modifier.[14] In response to Question A17b, the company performed the comparison to R-CVP/R-CHOP with additional adjustment for bone marrow involvement, and concluded that the addition of this extra variable has had little impact on the results.[14]

In conclusion, as in the ERG report, the main concerns are that the set of covariates included in the MAIC does not reflect the complete set of all possible covariates which affects the reliability of the OS and PFS results.[1] Although bone marrow involvement has been added, it remains a serious limitation which affects the reliability of the MAIC results.

Results of the MAIC are presented in the Kaplan-Meier curves for OS, PFS, and TTNLT (Company addendum, Figures 5-7).[2] Hazard ratios (HRs) from the Cox Proportional-Hazard models comparing R^2 and R-CHOP/R-CVP are reproduced in Table 4.8. R^2 had a significant improvement in OS and TTNLT compared to R-CHOP/R-CVP, but no with significant difference in PFS.

Table 4.8: Results from Cox Proportional Hazard models comparing R^2 and R-CVP/R-CHOP

Outcome	R^2, adjusted N	R-CHOP/R-CVP, N	HR (95% CI)[a]
OS	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]
TTNLT	[redacted]	[redacted]	[redacted]

CI = confidence interval; HR = hazard ratio; N = number of patients; OS = overall survival; PFS = progression-free survival; R^2 = rituximab plus lenalidomide; R-CHOP; rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP = rituximab plus cyclophosphamide, vincristine, prednisolone; TTNLT = time to next anti-lymphoma treatment.

[a]) bootstrapped CI.

ERG comment: As in the ERG report, in conclusion the results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were excluded from the matching models, small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources.[1] The analysis also used an unanchored MAIC involving two single treatment arms from different studies, as there was no relevant comparative trial data. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

4.5 Additional work on clinical effectiveness undertaken by the ERG

As with the ERG report, no further additional work on clinical effectiveness has been undertaken by the ERG.[1]

4.6 Conclusions of the clinical effectiveness section

The addendum presented by the company focused on a modified decision problem, which pertains only to the FL subgroup.[2] No further literature review was conducted. The ERG critique in this addendum, as in the ERG report, focused on the AUGMENT trial,[6] because this provides a head-to-head comparison of the intervention of interest (lenalidomide in combination with rituximab, R^2) versus a relevant comparator according to the NICE scope (rituximab monotherapy).[1] The company addendum also presented results for MAGNIFY and again they are ignored for the same reasoning as in the ERG report, i.e. single arm data only.[2, 7]

The same issues as identified in the ERG report regarding the total population in AUGMENT trial[6] still apply to the FL subgroup.[1] The trial did not include any patients from the UK and there is some disparity in the baseline characteristics between the R^2 and the R-mono arms.

As with the total population, the FL subgroup results from the AUGMENT trial show favourable results for R^2 when compared to R-mono in terms PFS with a greater median PFS ([redacted]).[2] However, unlike for the total population, there was also evidence of a difference in OS with a HR of [redacted]) for patients treated with R^2 compared to R-mono.[2] In terms of health-related quality of life, as with the total population, no clinically meaningful change from baseline in the GHS/QoL domain of the QLQ-C30 was observed across any of the post-baseline assessment visits, regardless of treatment group. As reported in the ERG report for the total population, R^2 was associated with more grade 3-4 TEAEs and SAEs and with more TEAEs leading to dose reductions, dose interruptions and discontinuations of lenalidomide/placebo or rituximab when compared to R-mono.[2]

In their addendum, the company updated the unanchored indirect comparison using data from HMRN:[2]

- R^2 from the FL subgroup versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN.

This also involved adding one more variable for matching, i.e. bone marrow involvement. Results from the MAIC (R^2 from FL subgroup versus pooled data for R-CHOP/R-CVP for non-rituximab refractory patients using data from HMRN) show a significant improvement in OS (HR = [redacted]) and TTNLT (HR = [redacted] compared to R-CHOP/R-CVP, but no statistically significant difference in PFS (HR = [redacted]).

The conclusion of the ERG report still applies that the results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were not included in the matching models, small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources. The analysis also used an unanchored MAIC involving two single treatment arms from different studies, as there was no relevant comparative trial data. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

There is no change from the ERG report.[1]

5.1.1 Searches performed for cost effectiveness section

There is no change from the ERG report.[1]

5.1.2 Inclusion/exclusion criteria used in the study selection

There is no change from the ERG report.[1]

5.1.3 Included/excluded studies in the cost effectiveness review

There is no change from the ERG report.[1]

5.1.4 Conclusions of the cost effectiveness review

There is no change from the ERG report.[1]

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.1: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS or company addendum)
Model	Partitioned survival model	Makes use of the PFS and OS data directly, ensuring that estimated survival outcomes versus observed outcomes are matched.	B.3.2, CS
States and events	Progression-free, post-progression, death		B.3.2, CS
Comparators	Non-rituximab-refractory patients: R-CHOP and R-CVP Rituximab-refractory patients: O-Benda	Expert opinion	3.1, Company addendum
Population	The patient population considered in the model is, adult patients with previously treated FL). The model is split into two subpopulations: non-rituximab-refractory and rituximab-refractory patients.	In line with the licence.	Company addendum
Treatment effectiveness	Non-rituximab-refractory: Unanchored MAIC using AUGMENT and HMRN Rituximab-refractory:		3.1, Company addendum

	Approach	Source/Justification	Signpost (location in CS or company addendum)
	Unanchored MAIC using MAGNIFY and GADOLIN		
Adverse events	Grade 3 and 4 based on trial data		3.1, Company addendum
Health related QoL	EQ-5D-3L data from AUGMENT	NICE reference case	3.1, Company addendum
Resource utilisation and costs	NHS and Personal Social Services	NICE reference case	3.1, Company addendum
Discount rates	3.5% discount rate was used for utilities and costs	NICE reference case	B.3.2, CS
Subgroups	non-rituximab-refractory and rituximab-refractory patients		NA
Sensitivity analysis	Probabilistic and deterministic sensitivity analyses and scenario analyses	NICE reference case	3.2, Company addendum

FL = follicular lymphoma; MZL = marginal zone lymphoma; R-CHOP = rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP = rituximab plus cyclophosphamide vincristine prednisolone; O-Benda = obinutuzumab plus bendamustine; MAIC = matching-adjusted indirect comparison; HMRN = Haematological Malignancy Research Network.

5.2.1 NICE reference case checklist (TABLE ONLY)

There is no change from the ERG report.[1]

5.2.2 Model structure

There is no change from the ERG report.[1]

5.2.3 Population

The patient population considered in the model is in line with the license: adult patients with previously treated FL.[2] The patient starting age and gender were matched to the data source used for the comparator arm, i.e. for non-R refractory patients this was the HMRN: mean age [redacted] years, percentage female [redacted]). Body surface area (BSA) data were taken from individual patients in the AUGMENT study (mean BSA 1.85 m^2).

ERG comment: The main concern of the ERG in the ERG report related to pooling the FL and MZL populations throughout the economic analysis.[1] This does not apply to the company addendum.

5.2.4 Interventions and comparators

As presented in the original CS, the R[2] dosing regimen within the model is lenalidomide 20 mg orally once daily on days 1–21 of repeated 28-day cycles for up to 12 cycles of treatment.[15] Rituximab is given as 375 mg/m2 every week in Cycle 1 (days 1, 8, 15 and 22) and Day 1 of every 28-day cycle for Cycles 2–5. This is in line with the recommended dose in the SmPC.[16] Patients with moderate renal impairment start on a dose of 10 mg of lenalidomide if CrCl is ≥30 ml/min but <60 ml/min. These criteria were met by [redacted]% of patients in AUGMENT and [redacted]% in MAGNIFY (R-refractory population), and these proportions are used to inform the starting dose in the model for the non-R-refractory and R-refractory

populations, respectively. In the company addendum, it was stated that "As per the request in clarification question B3, the FL only analysis has also been conducted for the R^2 versus R-mono comparison based on the head-to-head AUGMENT trial."[2 3, 4]

ERG comment: The main concerns of the ERG in the ERG report related to: a) the inclusion of O-Benda as a comparator while NICE have explicitly stated it is not considered a relevant comparator for disease that is refractory to rituximab, b) omitting R-mono as a comparator (based on expert opinion) although listed in the scope and given the direct evidence available. As R-mono was included as a comparator, this is no longer a concern. The issue concerning O-Benda still applies.[2]

5.2.5 Perspective, time horizon and discounting

There is no change from the ERG report.[1]

5.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for intervention and comparators was the AUGMENT study for R^2 and HMRN data for R-CHOP and R-CVP.[17, 18] The AUGMENT study is a Phase III, multicentre, double-blind, randomised study comparing R^2 versus R-mono in patients with non-R-refractory/relapsed FL or MZL. Only data from FL-patients in the R^2 arm and from the 22 June 2018 data cut-off were used in the model. The HMRN is a population-based cohort covering the Yorkshire and Humber & Yorkshire Cancer Networks for all patients newly diagnosed with a haematological malignancy between 2004 and 2016 and contained only data on FL patients.

The Phase III study by Van Oers et al. (2006) on R-CHOP,[11] which was used as a scenario in the original submission was discarded for the work in the addendum as the population in the Van Oers study was 100% rituximab naïve and therefore not considered relevant, also because prior rituximab exposure was a key treatment effect modifier in the MAIC.[15]

The matching methodology used was equivalent to that previously described in section B.2.9 of the company submission and in the ERG report.[1, 15] HMRN data for R-CHOP and R-CVP were pooled, and data from FL patients in AUGMENT (n=76) were then matched to the pooled data from HMRN for R-CHOP and R-CVP ([redacted]). In addition to what was presented in the original company submission, the FL only analysis was now also conducted for the R^2 versus R-mono comparison based on the head-to-head AUGMENT data.

Parametric survival curves were fitted to the patient level data from AUGMENT and HRMN in the same way as described in the ERG report.[1] The treatment waning effect was maintained at 5 years, as in the original submission.

Overall survival

Parallel to the original submission,[15] the parametric survival curve selected for the base-case was the Weibull. The company argued that they had sought clinical expert opinion to judge the appropriateness of model extrapolations for OS in R-CHOP/R-CVP. According to this expert opinion (for which no reference was provided by the company), the bottom three curves (exponential, Weibull, and log-logistic) were considered more plausible than the top three (generalized gamma, Gompertz, and lognormal). The clinical experts were said to have expressed a preference for the middle of the bottom three curves (Weibull) as the 20-year OS estimate seemed reasonable. Other arguments the company mentioned for choosing Weibull were that its AIC and BIC suggested a reasonable fit, and the Weibull distribution had been considered appropriate for the relapsed/refractory setting in TA137.[19]

The company considered these arguments sufficient to select the Weibull distribution for base-case in R-CHOP/R-CVP, and then also applied it to the R^2 arm. The exponential and log-logistic curves were used in a scenario.

For R^2 versus R-mono, parametric curves were fit directly to the AUGMENT FL only trial data. Based on AIC and BIC, generalized gamma and exponential seemed to fit best the R^2 data, and log-normal and exponential seemed to fit best the R-mono data. Clinicians at the advisory board [3] were shown the R-mono survival curve extrapolations for the FL+MZL population. As they judged the lower curves to be most plausible (having 40-50% survival at 10 years), the Weibull distribution was chosen for both R^2 and R-mono in the FL+MZL population. The company then also chose the Weibull distribution for both arms in the FL only analysis, in order to be consistent, they argue. The log-logistic and log-normal distributions were presented in a scenario analysis.

Progression free survival

For PFS, similar reasoning as in the original submission was followed to arrive at the selection of curves. In summary, for R-CHOP/R-CVP, in spite of the Weibull distribution making the best fit, the generalized gamma was selected for the base-case as the Weibull would cross the TTNLT curve in R-CHOP/R-CVP at approximately eight years, which was considered to be clinically implausible. The Weibull was explored in a scenario analysis. For R^2, the company decided to model PFS using the KM data until the maximum follow-up of 46.7 months, and then applied the comparator hazard to extrapolate further. The company considered this approach to ensure the relative treatment effect of R^2 versus R-CHOP/R-CVP based on the MAIC would be accurately reflected.

For the R-mono comparison, the log-logistic distribution was chosen for both arms to be consistent, as the company stated, with the FL+MZL population. The log-normal distribution would make a better fit based on AIC and BIC, but the company argued that the log-normal and log-logistic distributions were very similar and so effect on the ICER would be minimal.

Time to next anti-lymphoma treatment

From the cumulative hazard plot, proportional hazards seemed reasonable but to be consistent with OS and PFS, stratified models were used. Where unstratified models were explored in a scenario in the FL+MZL population, this was not done in the FL only population now. Based on the AIC/BIC, the log-normal distribution was selected for the base-case as it best fitted the R-CHOP/R-CVP data and also had the lowest AIC in R^2.

For the R-mono comparison, differently from the R-mono analyses the company provided in response to clarification[14] the log-normal distribution was used for both arms, based on AIC and BIC.

Time on treatment

As in the FL+MZL population, because the company considered the parametric survival curves in general to produce a poor fit to the data, KM data were used to directly in the model to inform the proportion of patients on treatment.

For the R-mono comparison the same approach was used, that is, KM data were used.

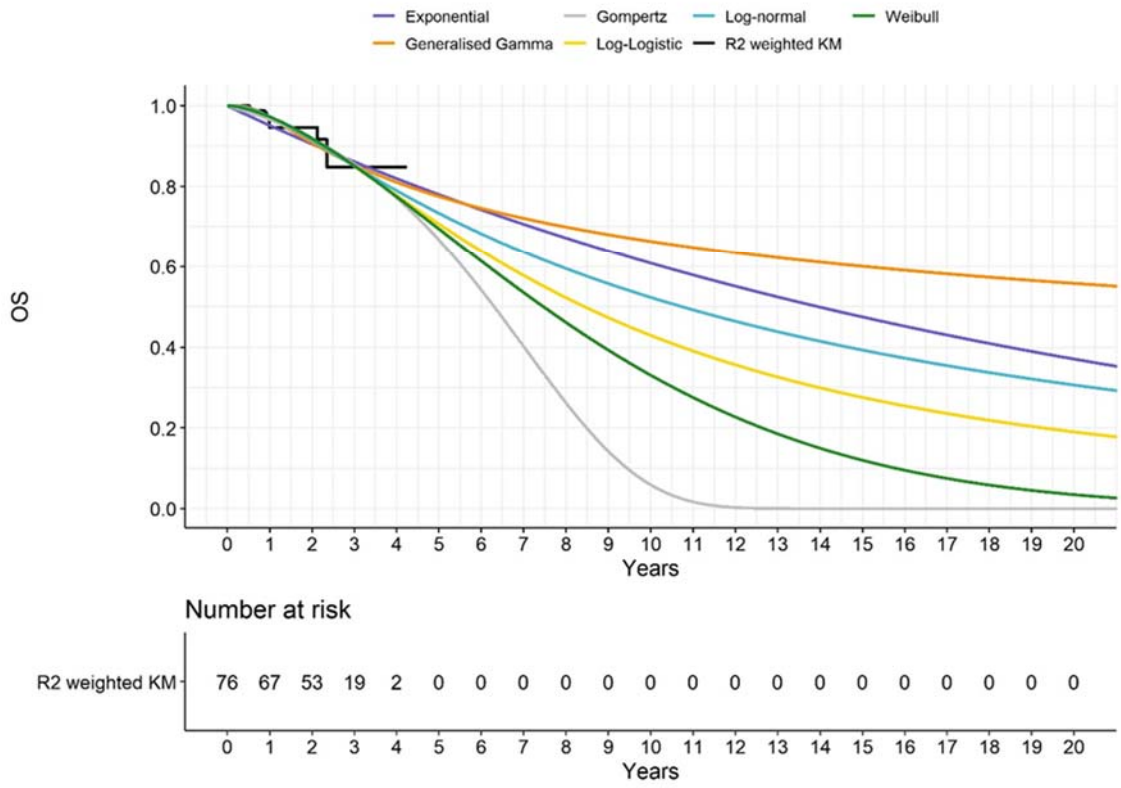
ERG comment: The main concerns of the ERG relate to: a) the Weibull OS curve producing vastly lower survival rates in R^2 than in the FL+MZL population b) the fact that the choice for Weibull in R^2 was not backed up by clinical opinion c) the divergence of the OS curves for R^2, making survival estimates highly uncertain d) the exclusion of the Van Oers scenario while the Van Oers data, being FL

only, may apply better to the FL only analyses than it did to the FL+MZL analysis, even though the ERG agrees it is still not representative for UK patients e) the exclusion of the scenario with unstratified TTNLT curves, while the log-cumulative hazard plot suggests that the proportional hazards assumption is very reasonable

In addition, the points of critique concerning treatment effectiveness and extrapolation raised in the ERG report still considered relevant in the FL only comparison are f) the uncertainty introduced by the indirect comparison of R^2 with R-CHOP and R-CVP based on only [redacted] patients - which seems to be underlined by h) the counterintuitive results for the R-mono comparison i) the lack of justification for the choice of time-point at which treatment effect ends, and j) the extrapolation of PFS.[1] Elaborate discussion of these issues can be found in the ERG report.[1]

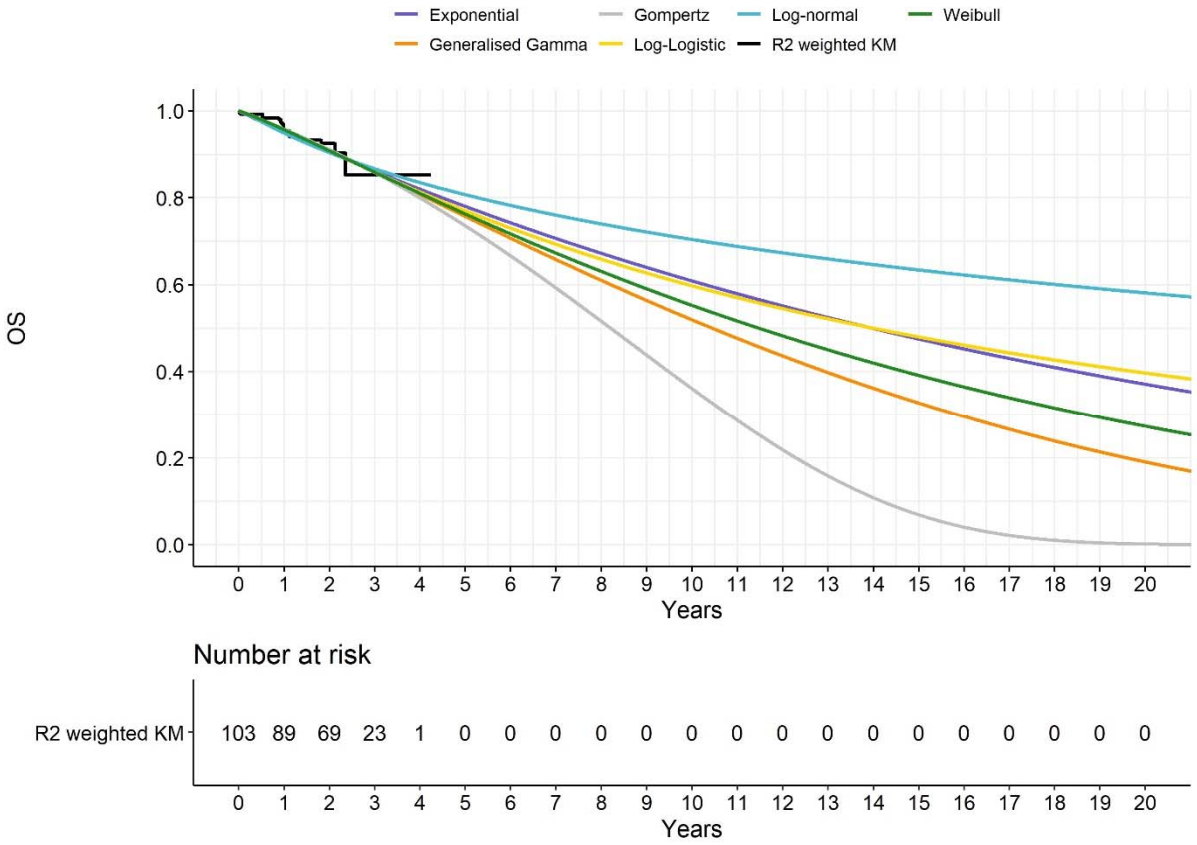
a) As can be seen from Figures 5.1 and 5.2 below, showing OS parametric curves for the FL only and FL+MZL populations, respectively, the survival estimates between these populations appear to be very different. For the Weibull distribution for instance (the distribution used in the company base-case), in the FL+MZL population, OS at 10 and 20 years was estimated at around 55% and 28% respectively. In the FL only population these percentages were 35% and 5%. These differences are even more striking as the FL only population is supposed to have a better prognosis as they are younger and fitter. The ERG considers this difference in OS estimates between the populations not plausible and as the company did not offer an explanation for this phenomenon, it contributes to the uncertainty around the model results.

Figure 5.1: OS parametric curves for R2 (FL only, weighted) in the non-rituximab refractory population



Source: addendum Figure 10[2]
FL = follicular lymphoma; KM = Kaplan–Meier; OS = overall survival; R = rituximab; R^2 = lenalidomide plus rituximab.

Figure 5.2: OS parametric curves for R^2 (FL+MZL, weighted) in the non-rituximab refractory population



Source: company submission Figure 24[15]
FL = follicular lymphoma; MZL = marginal zone lymphoma; KM = Kaplan–Meier; OS = overall survival; R = rituximab; R^2 = lenalidomide plus rituximab.

b) The company stated that clinical experts had been presented with the R-CHOP/R-CVP OS curves and that they considered the bottom three curves most plausible. The bottom three curves for R-CHOP/R-CVOP however were not comparable to what was seen in R^2. The bottom curve in R-CHOP had around 17% survival at 20 years, for instance, while the bottom two curves in R^2 had survival between 0% and 5%. Clinical validation of only the R-CHOP/R-CVP curves is not sufficient to also base the choice of R^2 curve on, in particular when the differences between the two sets of curves are as large as observed here. The ERG feels the company should have sought and presented clinical expert opinion on the quite deviant survival estimates for R^2 separately. The absence of clinical validation of R^2 OS curves increases the uncertainty around model results, although it is not possible to say in which direction.
c) Apart from the differences between extrapolations for MZL+FL and FL only, the OS estimates for the various survival extrapolations for R^2 within FL only were quite divergent. At 10 years, lowest and highest survival were estimated at around 3% and almost 70% and at 20 years this was 0% and around 67%. This implies that the selection of the right parametric curve is crucial,

as selecting any other curve would have a large impact on the ICER. The direction of the uncertainty introduced by this issue is unclear. Given the combined effect of issues a) b) and c) the ERG has serious doubts as to which OS curve for the R^2 arm is most suitable. As any choice could be considered arbitrary to some extent and may have substantial consequences for the ICER, the ERG decided to present six different base-cases, each based on one of the parametric curves for OS.

d) The scenario using data from the Van Oers publication (2006) was no longer included in the FL only work,[11] as, stated the company, the ERG discounted it as a scenario due to the population not being relevant to UK clinical practice. The ERG however feels that this scenario with the Van Oers data, being from an FL only population, may apply better to the FL only analyses than it did to the FL+MZL analysis. The ERG therefore performed the Van Oers scenario conditional on the ERG base-case.

e) In the MZL+FL analyses, a scenario with unstratified TTNLT curves was included as the log-cumulative hazard plot suggested that the proportional hazards assumption would not be unreasonable. For the FL population, the log-cumulative hazard plot was even more suggestive of proportional hazard but the scenario with unstratified curves was not included anymore by the company. The ERG feels this scenario is relevant and has tested it conditional on the ERG base-case, but the impact was very minimal and therefore it was not included as a scenario in the current analyses.

5.2.7 Adverse events

The methods for including adverse events in the model were similar to those described in the ERG report with the only difference being that the adverse events from AUGMENT were now taken from the FL only population for both R^2 and R.[1] Table 5.2 shows the incidence for grade 3/4 adverse events as used in the model.

Table 5.2: Grade 3/4 AE incidence: non-rituximab (FL only) refractory population

AE	R^2 (n=146)		R-CHOP (%)	R-CVP (%)	R-mono (n=148)	
	n	%			n	%
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lymphocyte count decreased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Febrile neutropenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
White blood cell count decreased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypokalaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary embolism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AE	R^2 (n=146)		R-CHOP (%)	R-CVP (%)	R-mono (n=148)	
	n	%			n	%
Nausea and emesis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Allergic reaction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alopecia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abdominal pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute kidney injury	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total cost	**£1,796**		**£3,471**	**£2,714**	**£393**	
Source	AUGMENT		RELEVANCE (adjusted)[a]		AUGMENT	

Source: Table 23 of the company addendum[2]
AE = adverse event; CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP = cyclophosphamide, vincristine and prednisolone; FL = follicular lymphoma; mono = monotherapy; NR = not reported; R = rituximab; R^2 = lenalidomide plus rituximab.
Notes: [a] Comparator AE incidence = (AECOMPARATOR incidence in RELEVANCE/AER[2] incidence in RELEVANCE) x AER[2] incidence in AUGMENT.

ERG comment: The original concern of the ERG relating to the omission of AEs related to ASCT and subsequent R-mono therapy in the R^2 arm is no longer applicable as the company has pointed out in the FAC[20] that in fact AEs related to subsequent therapies were not included in the model altogether. The ERG therefore decided to take this fixing violation out of the ERG base-case, which only had a very small impact on the ICER.

The other original concern of the ERG, which related to the RELEVANCE population being exclusively patients that were previously untreated, still applies. Elaborate discussion of this issue can be found in the ERG report.[1]

5.2.8 Health-related quality of life

The EQ-5D meta-regression was re-ran from AUGMENT using only the FL population according to the same methodology previously described in Section B.3.4 of the original company submission.[15] The resulting final utility values per health state as used in the model are displayed in Table 5.3.

ERG comment: The main concerns of the ERG relate to the fact that the utility values that the ERG already considered very high in the original company submission were now even further increased. Even though it may be thought plausible that quality of life would be slightly higher in an FL only population, given their better condition, the ERG still considers it unlikely that utilities would be higher than in the general population. Therefore, the ERG maintained their base-case which would cap utility values at the level of the general population as in the ERG report.[1]

Table 5.3: Health state utility values

State	Utility value R^2 versus R-CHOP/CVP	Utility value R^2 versus R-mono	Reference	Justification
PF	0.867	0.846	Section 3.1, page 62 of the addendum	EQ-5D values derived from a relevant patient population and
PP (off treatment)	0.841	0.820		
PP (on treatment)	0.806	0.785		

State	Utility value R^2 versus R-CHOP/CVP	Utility value R^2 versus R-mono	Reference	Justification
				model specific health states.
Source: based on Table 25 of the company addendum[2]				

Adverse event related disutility values

There is no change from the ERG report.[1]

5.2.9 Resources and costs

Resource use and costs data identified in the review

There is no change from the ERG report.[1]

Drug acquisition costs (with PAS)

Drug acquisition costs were calculated as in CS section B.3.5,[15] but for the FL only analysis the dose reductions for lenalidomide have been taken directly from the AUGMENT trial using only the FL patients for the analysis. These are then used to calculate the cost per cycle applied to the proportion of patients on treatment.

Administration costs

There is no change from the ERG report.[1]

Treatment-specific monitoring

There is no change from the ERG report.[1]

Health state costs

There is no change from the ERG report.[1]

Disease monitoring

There is no change from the ERG report.[1]

Stem cell transplant (pre-progression)

There is no change from the ERG report.[1]

Adverse event related costs and costs of terminal care

There is no change from the ERG report.[1]

Costs of subsequent treatments

Subsequent treatments usage from the FL only population in AUGMENT were used to inform the costs of subsequent treatment for the R^2 arm. The durations of the subsequent treatments and cost per treatment remain the same as per section B.3.5 in the original company submission.[15] See Table 5.4 for included subsequent treatments and costs.

Table 5.4: Subsequent treatments and costs

Subsequent treatment	R^2 (FL only AUGMENT) n (%)[a] n=147	R-mono (FL only AUGMENT) n (%)[a] n=148
R-mono	[redacted]	[redacted]
R-Benda	[redacted]	[redacted]
R-CHOP	[redacted]	[redacted]
R-CVP	[redacted]	[redacted]
Other R-chemo	[redacted]	[redacted]
O-Benda	[redacted]	[redacted]
Bendamustine	[redacted]	[redacted]
Other chemotherapy	[redacted]	[redacted]
Targeted therapies	[redacted]	[redacted]
Radiotherapy	[redacted]	[redacted]
Other	[redacted]	[redacted]
ASCT	[redacted]	[redacted]
Total weighted treatment cost (R^2)	**£3,128**	**£6,126**
Total weighted administration cost (R^2)	**£349**	**£524**

Source: Table 26 of the company addendum[2]
ASCT = autologous stem cell transplant; Benda = bendamustine; chemo = chemotherapy; CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP = cyclophosphamide, vincristine and prednisolone; mono = monotherapy; FL = follicular lymphoma; R = rituximab; R^2 = lenalidomide plus rituximab.
Notes: [a]) Percentages include multiple lines therefore total may be over 100%.

5.2.10 Cost effectiveness results

R^2 versus R-CHOP and R-CVP

In the deterministic base-case analysis, total LYs and QALYs gained were larger for R^2 than for R-CHOP and R-CVP. Incremental QALYs ([redacted]) were mainly driven by QALY gains in the PP (off treatment) health state. Total costs were also higher for R^2 than for R-CHOP and R-CVP. Incremental costs ([redacted] and [redacted] respectively) mainly resulted from higher drug acquisition (induction) costs. The deterministic incremental cost effectiveness ratio (ICER) amounted to £15,909 per QALY gained for R^2 versus R-CHOP and £23,746 per QALY gained for R^2 versus R-CVP (see Table 5.5).

R^2 versus R-mono

In the deterministic base-case analysis, using data from the head-to-head comparison in AUGMENT, R^2 was more expensive and generated less QALYs than in the indirect comparison used for the analysis with R-CHOP and R-CVP. Still, total LYs and QALYs gained were larger for R^2 than for R-mono. Incremental QALYs ([redacted]) were mainly driven by QALY gains in the PF health state. Total costs were also higher for R^2 than for R-mono. Incremental costs ([redacted]) mainly resulted from higher drug acquisition (induction) costs. The deterministic cost effectiveness ratio (ICER) amounted to £20,274 per QALY gained (see Table 5.5).

Table 5.5: Company's deterministic base-case results for FL only

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
R^2 versus R-CHOP							
R^2	[redacted]	[redacted]	[redacted]				
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R^2 versus R-CVP							
R^2	[redacted]	[redacted]	[redacted]				
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R^2 versus R-mono							
R^2	[redacted]	[redacted]	[redacted]				
R-mono	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£20,274

Source: Based on Table 27 of the addendum.[2]
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALY = quality adjusted life year.

5.2.11 Sensitivity analyses

As in the original submission, the company performed a probabilistic sensitivity analysis (PSA) and deterministic sensitivity analysis (DSA) to show the uncertainty surrounding the base-case results.

R^2 versus R-CHOP and R-CVP

Compared with the deterministic results, the PSA with 1,000 iterations showed substantially lower incremental QALYs and costs for both R-CHOP and R-CVP, which resulted in increased ICERs of £27,768 (versus R-CHOP) and £41,602 (versus R-CVP) (see Table 5.6). The cost effectiveness acceptability curve in the economic model showed that R^2 had an 63% (versus R-CHOP) and 54% (versus R-CVP) probability of being cost effective at a willingness-to-pay (WTP) threshold of £30,000 (see Figures 5.3 and 5.4).

R^2 versus R-mono

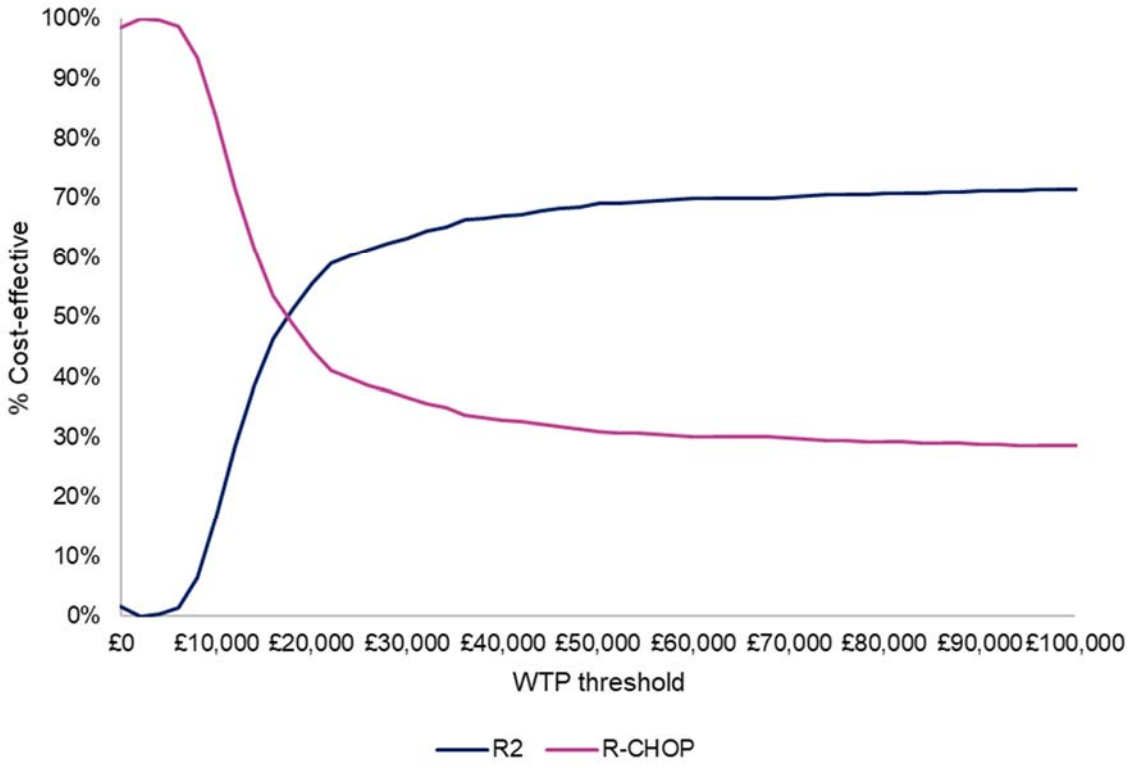
Compared with the deterministic results, the PSA with 1,000 iterations showed lower incremental QALYs and costs, which resulted in an increased ICER of £23,412) (see Table 5.6). The cost effectiveness acceptability curve in the economic model showed that R^2 had a 77% probability of being cost effective at a willingness to pay (WTP) threshold of £30,000 (see Figure 5.5).

Table 5.6: Company's base-case results (probabilistic, 1,000 iterations)

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
R^2 versus R-CHOP					
R^2	[redacted]	[redacted]	--	--	--
R-CHOP	[redacted]	[redacted]	[redacted]	[redacted]	£27,768
R^2 versus R-CVP					
R^2	[redacted]	[redacted]			
R-CVP	[redacted]	[redacted]	[redacted]	[redacted]	£41,602
R^2 versus R-mono					
R^2	[redacted]	[redacted]	--	--	--

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
R-mono	[redacted]	[redacted]	[redacted]	[redacted]	23,412
Source: Based on Table 29 of the company addendum. ICER = incremental cost effectiveness ratio; QALY = quality adjusted life year.					

Figure 5.3: Cost-effectiveness acceptability curve – R^2 versus R-CHOP - FL only

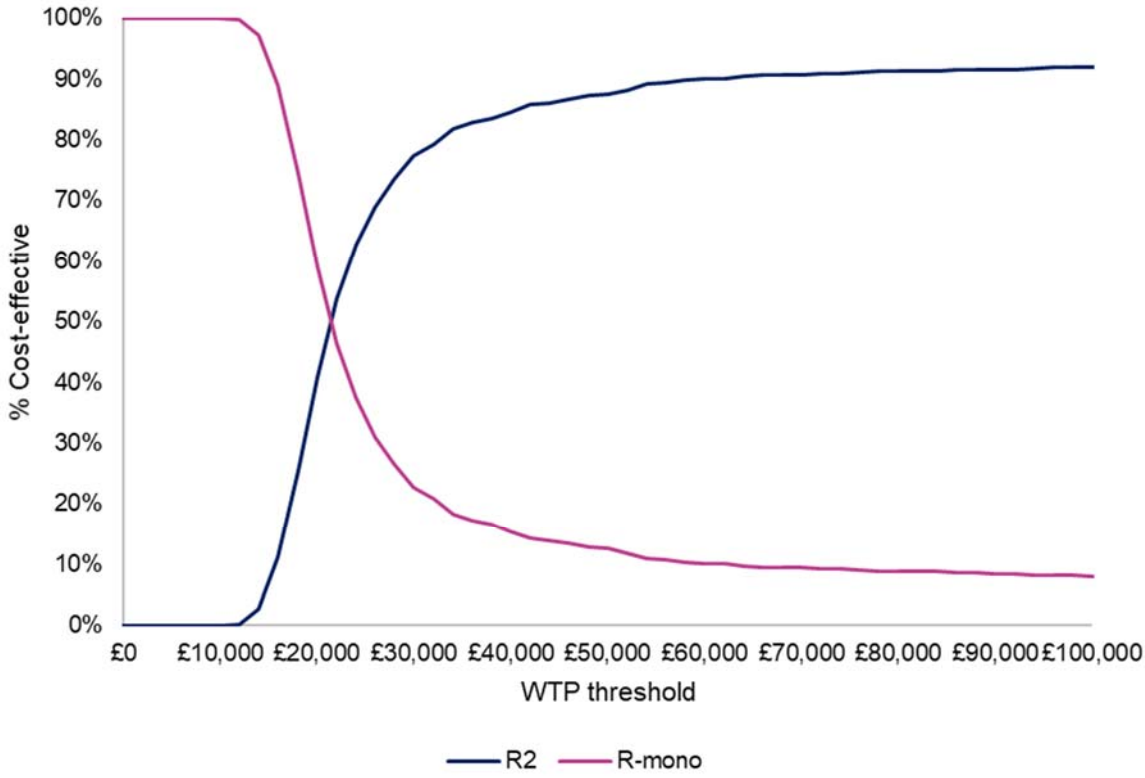


Key: CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; FL = follicular lymphoma; R = rituximab; R^2 = lenalidomide plus rituximab; WTP = willingness to pay.

Figure 5.4: Cost-effectiveness acceptability curve – R^2 versus R-CVP - FL only

Key: CVP = cyclophosphamide, vincristine and prednisolone; FL = follicular lymphoma; R = rituximab; R^2 = lenalidomide plus rituximab; WTP = willingness to pay.

Figure 5.5: Cost-effectiveness acceptability curve – R^2 versus R-mono - FL only



Key: FL = follicular lymphoma; mono = monotherapy; R = rituximab; R^2 = lenalidomide plus rituximab; WTP = willingness to pay.

Curve fit scenario analyses

The company also performed a number of scenarios using alternative distributions for PFS and OS, given the following arguments:

- Weibull made the best fit to R-CHOP/R-CVP PFS
- Exponential and log-logistic were considered plausible for R-CHOP/R-CVP OS
- Log-normal for R^2 OS made the best fit according to the AIC

Tables 5.7 and 5.8 provide the results for these curve fit scenarios in the R-CHOP and R-CVP comparisons, respectively.

Table 5.7: Results of key curve fit scenario analyses (with PAS) - R^2 versus R-CHOP – FL only

Curve fit selection			ICER (£) versus baseline (QALYs)	
OS		PFS		
R^2	R-CHOP/CVP	R-CHOP/CVP	PSA	Deterministic
Base case				
Weibull	Weibull	Generalized gamma	£27,768	£15,909
Plausible scenarios				
Weibull	Weibull	Weibull[a]	£26,827	£15,105
Exponential[b]	Exponential[b]	Generalized gamma	£12,953	£12,651
Exponential[b]	Exponential[b]	Weibull[a]	£12,002	£12,061
Log-logistic[b]	Log-logistic[b]	Generalized gamma	£15,925	£12,955
Log-logistic[b]	Log-logistic[b]	Weibull[a]	£14,800	£12,322
Log-normal[c]	Weibull	Generalized gamma	£15,894	£13,176
Log-normal[c]	Weibull	Weibull[a]	£14,835	£12,541

Source: Based on Table 30 of the company addendum.[2]

CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP = cyclophosphamide, vincristine and prednisolone; ICER = incremental cost-effectiveness ratio; KM = Kaplan-Meier; OS = overall survival; PAS = patient access scheme; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; QALYs = quality-adjusted life years; R = rituximab; R^2 = lenalidomide plus rituximab.

Notes: [a]) Weibull based on best PFS AIC/BIC for R-CHOP/CVP and fit the data well; [b], based on other clinically plausible OS extrapolations from clinical expert opinion; [c], Based on best OS AIC/BIC for R^2.

Table 5.8: Results of key curve fit scenario analyses (with PAS) - R^2 versus R-CVP – FL only

Curve fit selection			ICER (£) versus baseline (QALYs)	
OS		PFS		
R^2	R-CHOP/CVP	R-CHOP/CVP	PSA	Deterministic
Base case				
Weibull	Weibull	Generalized gamma	£41,602	£23,746
Scenarios				
Weibull	Weibull	Weibull[a]	£40,630	£22,886
Exponential[b]	Exponential[b]	Generalized gamma	£18,555	£18,394
Exponential[b]	Exponential[b]	Weibull[a]	£18,514	£17,774

Curve fit selection			ICER (£) versus baseline (QALYs)	
OS		PFS		
R^2	R-CHOP/CVP	R-CHOP/CVP	PSA	Deterministic
Log-logistic[b]	Log-logistic[b]	Generalized gamma	£26,740	£18,886
Log-logistic[b]	Log-logistic[b]	Weibull[a]	£25,172	£18,218
Log-normal[c]	Weibull	Generalized gamma	£23,321	£19,253
Log-normal[c]	Weibull	Weibull	£22,089	£18,584

Source: Based on Table 31 of the company addendum.[2]
CHOP = cyclophosphamide, doxorubicin hydrochloride, vincristine and prednisolone; CVP = cyclophosphamide, vincristine and prednisolone; ICER = incremental cost-effectiveness ratio; OS = overall survival; PAS = patient access scheme; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; QALYs = quality-adjusted life years; R = rituximab; R^2 = lenalidomide plus rituximab.
Note: [a]) Weibull based on best PFS AIC/BIC for R-CHOP/CVP and fit the data well; [b]) Based on other clinically plausible OS extrapolations from clinical expert opinion; [c]) Based on best OS AIC/BIC for R^2.

Scenario analyses

The company conducted several scenario analyses. The results for R^2 versus R-CHOP showed ICERs ranging between £4,598 and £21,782 per QALY gained, excluding the scenarios assessing different time horizons. The three most influential scenarios that increased the ICER were using the Weibull (£21,782) and Gompertz (£21,312) distributions for R^2 PFS and applying equal subsequent treatment costs for R^2 and R-CHOP (£20,729). The three most influential scenarios that decreased the ICER were using the exponential distribution for R^2 ToT (£4,598) and using the exponential (£11,108) and generalised gamma (£11,340) distributions for R^2 OS.

The results for R^2 versus R-CVP showed ICERs ranging between £12,413 and £30,043 per QALY gained, excluding the scenarios assessing different time horizons. The three most influential scenarios that increased the ICER were using the Weibull distribution for R^2 PFS (£30,043), using the exponential distribution for R-CVP OS (£29,994) and using the Gompertz distribution for R^2 PFS (£29,541). The three most influential scenarios that decreased the ICER were using the exponential distribution for R^2 ToT (£12,413), and using the exponential (£15,854) and generalised gamma (£16,236) distributions for R^2 OS.

The results for R^2 versus R-mono showed ICERs ranging between £12,085 and £33,912 per QALY gained, excluding the scenarios assessing different time horizons. The three most influential scenarios that increased the ICER were applying R-mono hazard to R^2 after 3 years (£33,912), using the Gompertz distribution for R-mono OS and using a 6.0% discount rate for QALYs (£23,953). The three most influential scenarios that decreased the ICER were using the exponential distribution for R^2 ToT (£12,085), applying R-mono hazard to R^2 after 10 years (£12,334) and using a 0.0% discount rate for QALYS (£15,638).

ERG comment: The main comments of the ERG relate to: a) the substantial difference between deterministic and probabilistic results in the company base-case which is likely caused by the instability of the Weibull OS curve estimates used in the company base-case. In the curve fit scenarios presented by the company, the difference between probabilistic and deterministic ICERS is substantially reduced

when other parametric curves for OS were used. The ERG presents 6 base-cases, one for each of the parametric curves estimated for OS.

5.2.12 Model validation and face validity check

There is no change from the ERG report.[1]

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.9 summarises the main issues highlighted by the ERG in Section 5.2 of this report, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case. Most of the items in this table were already in the original ERG report, but are repeated here for completeness.

Table 5.9: Main ERG critique of company's submitted economic evaluation

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
Partitioned survival analysis, no alternative results from state transition model provided for comparison	+/-	No	Requested but not provided
Treatment effectiveness and extrapolation (section 5.2.6)			
Indirect comparison seems to inflate R^2 efficacy and lower costs relative to R^2 in direct comparison based on AUGMENT	+	No	No
Substantial uncertainty concerning extrapolation of OS curves. Company base-case not based on best fit, nor solid other justification	+/-	Base-case (MJ), scenarios	Scenarios
Curves for PFS extrapolation do not provide best fit, choice for KM + comparator hazard approach is not sufficiently justified	+/-	Base-case (MJ), scenarios	Scenarios
Curves for TTNLT extrapolation do not provide best fit, choice is not sufficiently justified	+/-	Base-case (MJ), scenarios	Scenarios
Cut-off for treatment effectiveness at 5 years not supported by evidence	+/-	Scenarios	Scenarios
Adverse events (section 5.2.7)			
Incidence for adverse events in R-CHOP and R-CVP taken from published source on a previously untreated population	+/-	Scenario	Scenario
Health-related quality of life (section 5.2.8)			
Utility scores for all health states are likely high	+	Base-case (FV)	Yes, scenarios allow for alternative values
Utility decrement post progression low	+	Scenarios	
Resources and costs (section 5.2.9)			
One-off costs for subsequent treatment likely underestimates R^2 costs	+	Scenario	Scenario using same subsequent treatment costs
Incidences of subsequent treatments for R-CHOP and R-CVP were taken from the mixed R-chemo group of HMRN, which likely is an overestimate	+	Base-case (FV)	No

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Consolidation ASCT in R^2 arm assumed zero, data on observed number of ASCTs was not provided in CS	+	No	No
Cost effectiveness analyses (sections 5.2.10 and 5.2.11)			
Discrepancy between probabilistic and deterministic results	+/-	No	No
PSA does not allow for full incremental analysis	+/-	No	No

Footnotes: [a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; ICER = incremental cost effectiveness ratio; MJ = matters of judgement; NA = not applicable.

Based on all considerations in Section 5.2 of this report (summarised in Table 5.9), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016[21])

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred).

Most importantly, the ERG now presents six base-cases, one for each parametric curve for OS: Weibull (company base-case), exponential, log-normal, log-logistic, generalized gamma, and Gompertz.

The adjustments apply to the R-CHOP and R-CVP comparisons. For the R-mono comparison, these adjustments may be different or do not apply. In the list below, when nothing is mentioned on R-mono, this implies that this particular adjustment was similarly applied to the R-mono comparison.

Fixing errors

1. Error cells when using 'Van Oers' as input for R-CHOP efficacy (section 5.2.6).
   The ERG replaced dots by commas in the Van Oers parameters for curves.

Fixing violations

2. Subsequent treatment rates for R-CHOP/R-CVP taken from mixed R-chemo population (section 5.2.9). The ERG used pooled R-CVP/R-CHOP subsequent treatment rates instead of R-chemo. (Not applicable in the R-mono comparison)
3. Utilities in all health states were higher than or comparable to general population levels (section 5.2.8). The ERG capped utilities at the general population level

Matters of judgment

4. KM+comparator hazard approach likely overestimates PFS in R^2 (section 5.2.6).
   The ERG used log-logistic for PFS in R^2 and Weibull for PFS in the comparator (not applied to R-mono comparison)
5. Lognormal curves for extrapolating TTNLT appear to be suboptimal (section 5.2.6).
   The ERG used log-logistic for TTNLT both arms (not applied to R-mono comparison)

Tables 6.1 to 6.18 show the combined effect of all abovementioned adjustments simultaneously, resulting in the (deterministic) ERG base-case. The 'fixing error' adjustments were combined and the other ERG analyses were performed also incorporating these 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

5.3.1 ERG base-case results

The results of the deterministic ERG base-case are shown in Table 5.10. The ERG wishes to emphasise that, as in the original ERG report, all ERG analyses (except those for the R-mono comparison) are conditional upon the MAIC results for which uncertainty could not be quantified or incorporated in the economic model. For R-CHOP, deterministic incremental costs varied from [redacted] to [redacted] for the various OS curves and incremental QALYs from [redacted] to [redacted] which resulted in an ICER range of £12,941 to £21,781. For R-CVP, incremental costs ranged from [redacted] to [redacted] with the same range of incremental QALYs as in the R-CHOP comparison (by definition), which resulted in an ICER that ranged from £17,312 to £30,404. Finally, for the R-mono deterministic comparison, incremental

costs varied from [redacted] to [redacted] and incremental QALYs from [redacted] to [redacted] with resulting ICERs ranging from £14,457 to £25,625.

The probabilistic ERG base-case (based on 1,000 iterations) for R^2 versus R-CHOP ranged from £16,874 to £44,888. For R^2 versus R-CVP, the ICER ranged from £23,135 to £59,810 and for R^2 versus R-mono, it ranged from £18,779 to £27,156. Compared with the deterministic base-case results, the ERG PSA resulted in higher ICERs, similar to what was seen in the company analyses. Particularly for the Weibull and Gompertz OS curves in the R-CHOP and R-CVP comparisons, the probabilistic ICER would sometimes be around twice the value of the deterministic ICER. For all the other OS curves, the differences between deterministic and probabilistic ICERs were more modest, although still considerable at times (see Table 5.11). The CEACs of all analyses are presented in Figures 5.6 to 5.23.

Table 5.10: ERG pairwise deterministic base-case results

Technologies	OS curve	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case for R^2 versus R-CHOP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£21,781
R-CHOP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£16,581
R-CHOP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£14,531
R-CHOP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£17,146
R-CHOP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£12,941
R-CHOP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£20,019
R-CHOP	Gompertz	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-CVP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£30,404
R-CVP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£22,742
R-CVP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£19,658
R-CVP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£23,529
R-CVP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£17,312
R-CVP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£27,767
R-CVP	Gompertz	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-mono						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£17,856

R-mono	Exponential	[redacted]				
R^2	Log-normal	[redacted]				£16,884
R-mono	Log-normal	[redacted]				
R^2	Log-logistic	[redacted]				£17,366
R-mono	Log-logistic	[redacted]				
R^2	Gen gamma	[redacted]				£14,457
R-mono	Gen gamma	[redacted]				
R^2	Gompertz	[redacted]				£25,625
R-mono	Gompertz	[redacted]				

ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year

Table 5.11: ERG probabilistic base-case results

Technologies	OS curve	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case for R^2 versus R-CHOP						
R^2	Weibull	[redacted]				£44,888
R-CHOP	Weibull	[redacted]				
R^2	Exponential	[redacted]				£17,138
R-CHOP	Exponential	[redacted]				
R^2	Log-normal	[redacted]				£17,177
R-CHOP	Log-normal	[redacted]				
R^2	Log-logistic	[redacted]				£20,800
R-CHOP	Log-logistic	[redacted]				
R^2	Gen gamma	[redacted]				£16,874
R-CHOP	Gen gamma	[redacted]				
R^2	Gompertz	[redacted]				£30,229
R-CHOP	Gompertz	[redacted]				
Probabilistic ERG base-case for R^2 versus R-CVP						
R^2	Weibull	[redacted]				£59,810
R-CVP	Weibull	[redacted]				
R^2	Exponential	[redacted]				£23,583
R-CVP	Exponential	[redacted]				
R^2	Log-normal	[redacted]				£23,135
R-CVP	Log-normal	[redacted]				
R^2	Log-logistic	[redacted]				£32,899
R-CVP	Log-logistic	[redacted]				
R^2	Gen gamma	[redacted]				£24,778
R-CVP	Gen gamma	[redacted]				
R^2	Gompertz	[redacted]				£43,915
R-CVP	Gompertz	[redacted]				
Probabilistic ERG base-case for R^2 versus R-mono						
R^2	Weibull	[redacted]				£24,659
R-mono	Weibull	[redacted]				

R^2	Exponential	[redacted]				£18,779
R-mono	Exponential					
R^2	Log-normal					£19,326
R-mono	Log-normal					
R^2	Log-logistic					£20,027
R-mono	Log-logistic					
R^2	Gen gamma					£26,462
R-mono	Gen gamma					
R^2	Gompertz					£27,156
R-mono	Gompertz					

ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Figure 5.6: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: Weibull OS

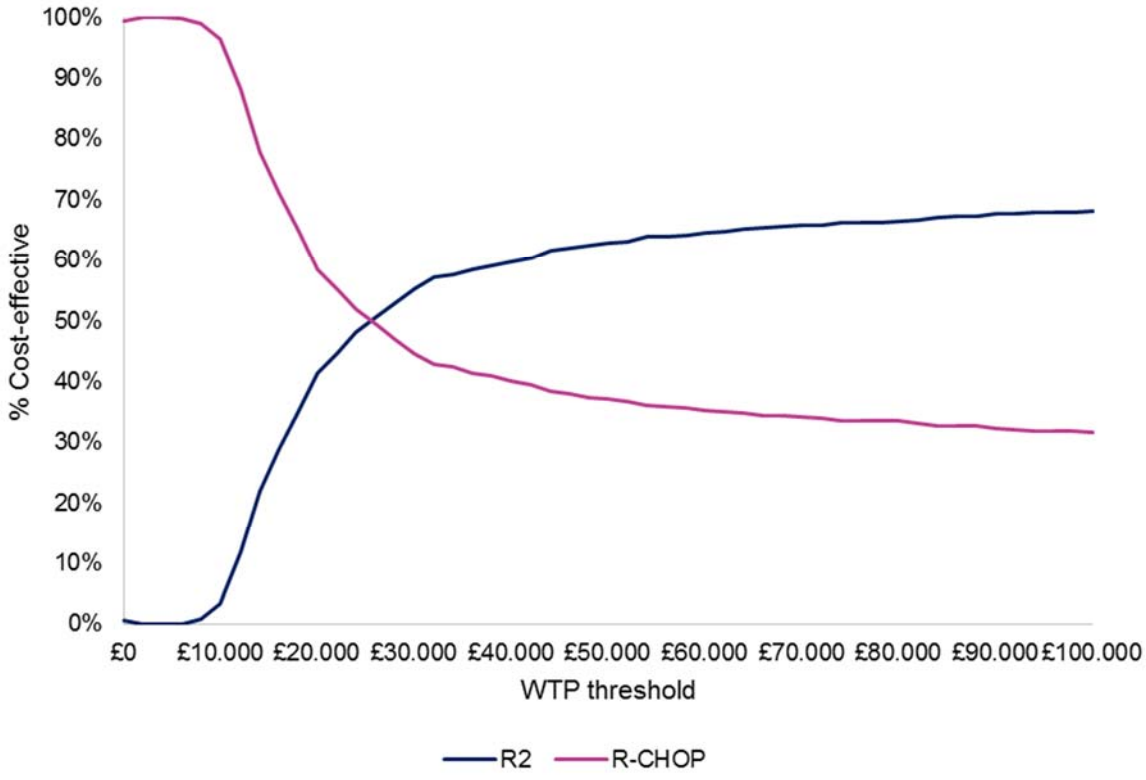


Figure 5.7: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: exponential OS

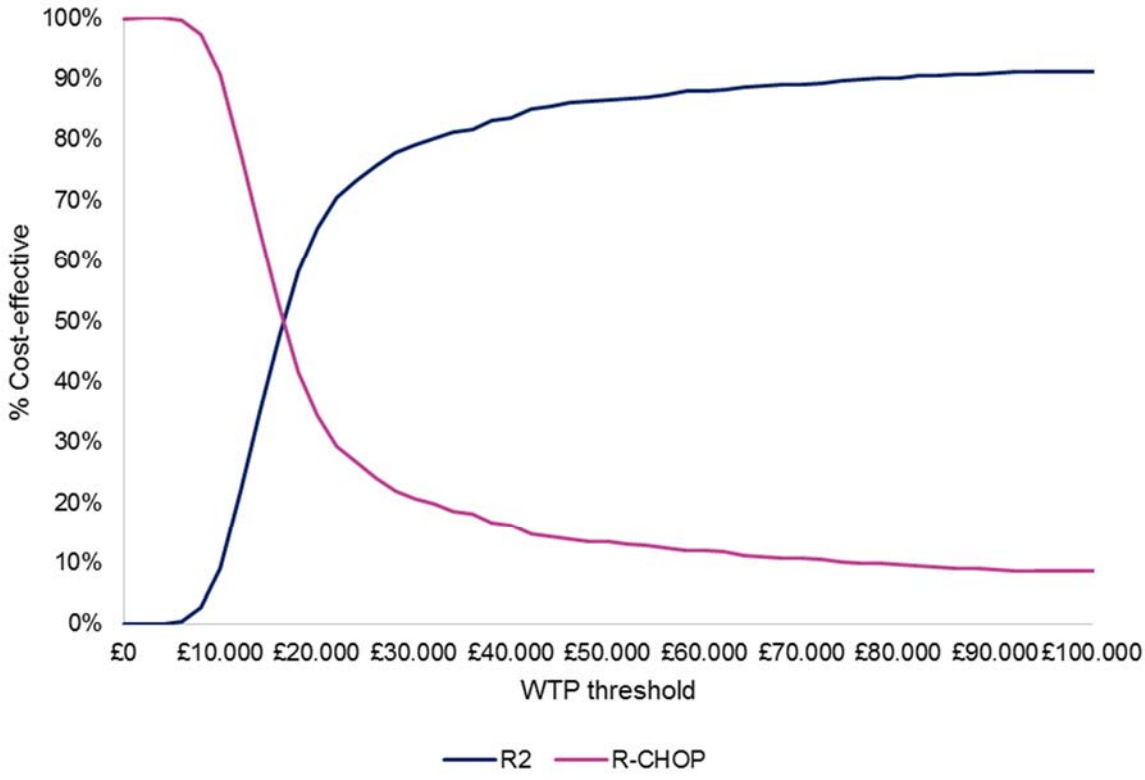


Figure 5.8: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: log-normal OS

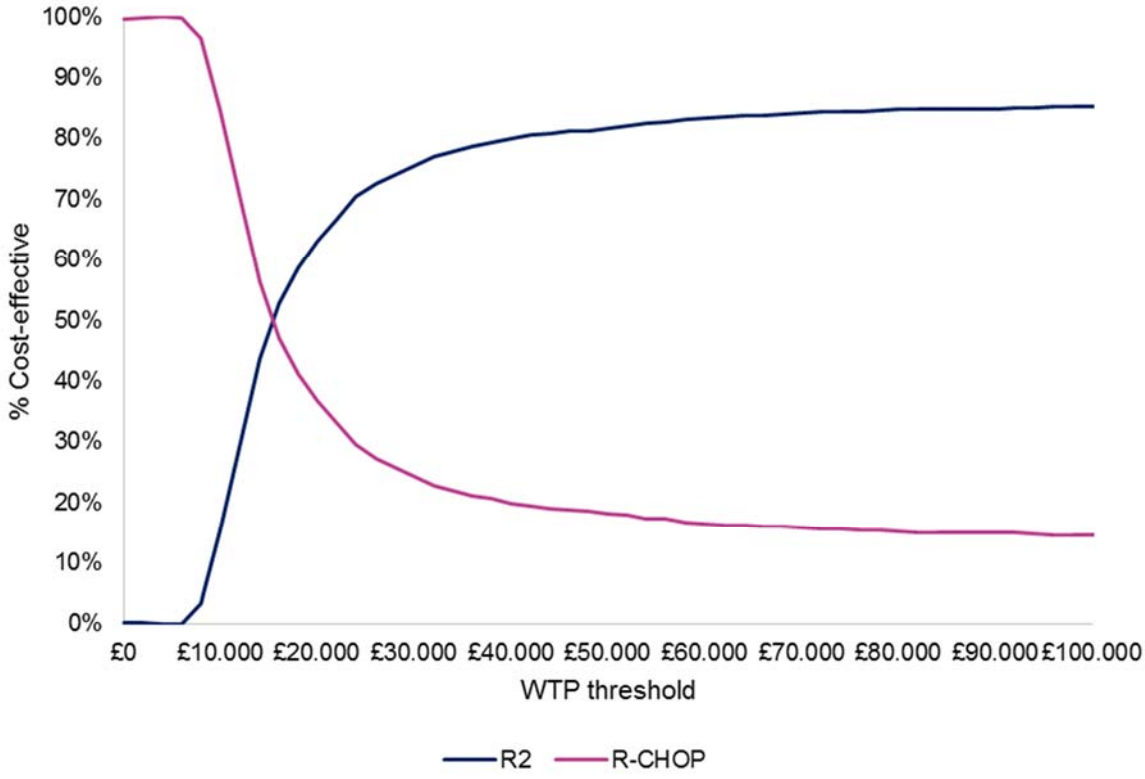


Figure 5.9: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: log-logistic OS

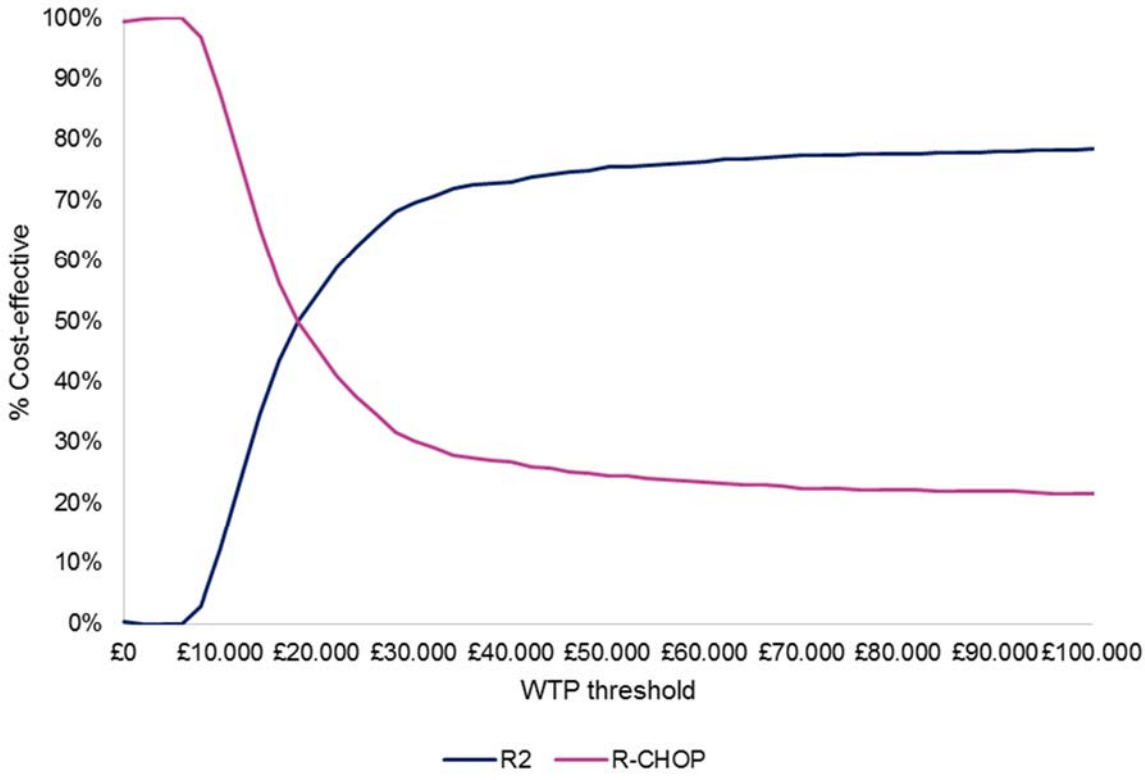


Figure 5.10: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: generalized gamma OS

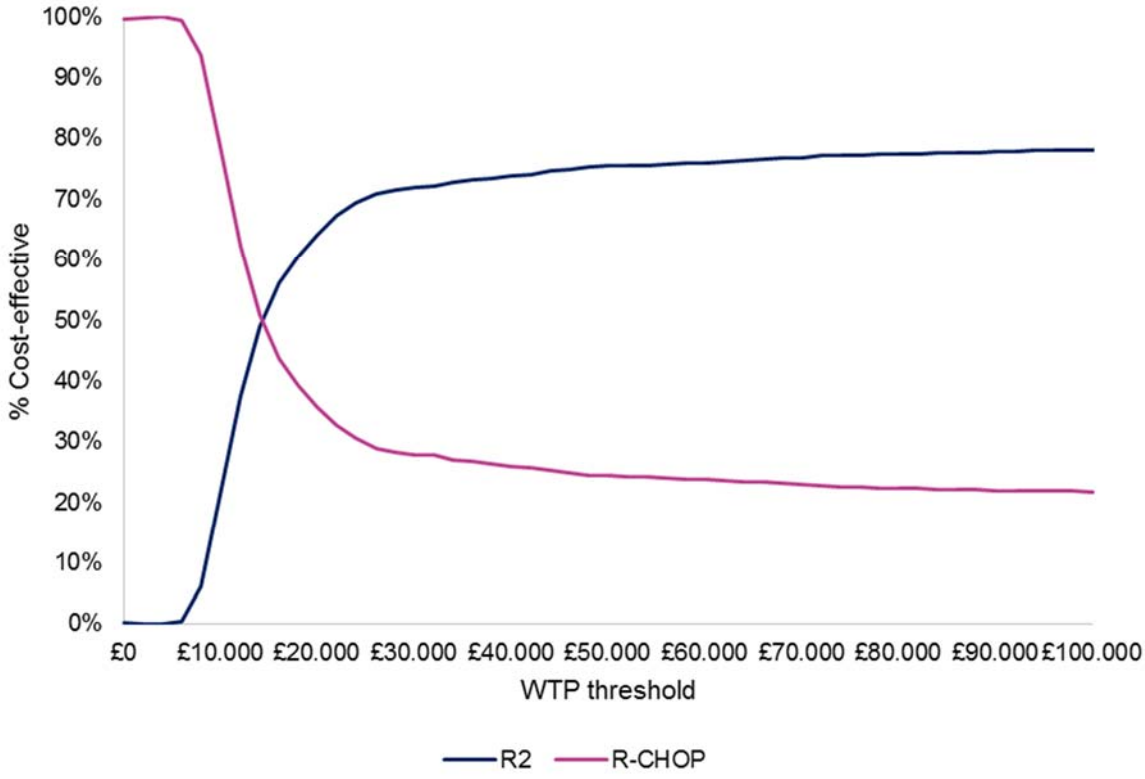


Figure 5.11: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: gompertz OS

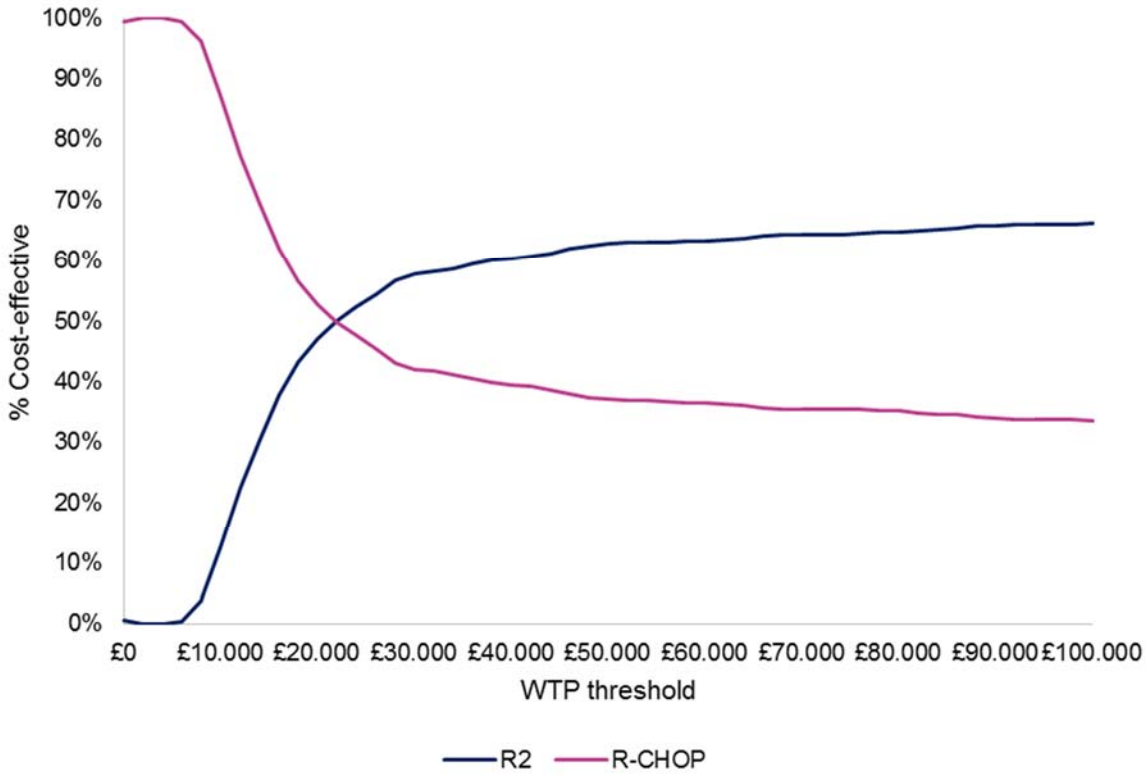


Figure 5.12: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: Weibull OS

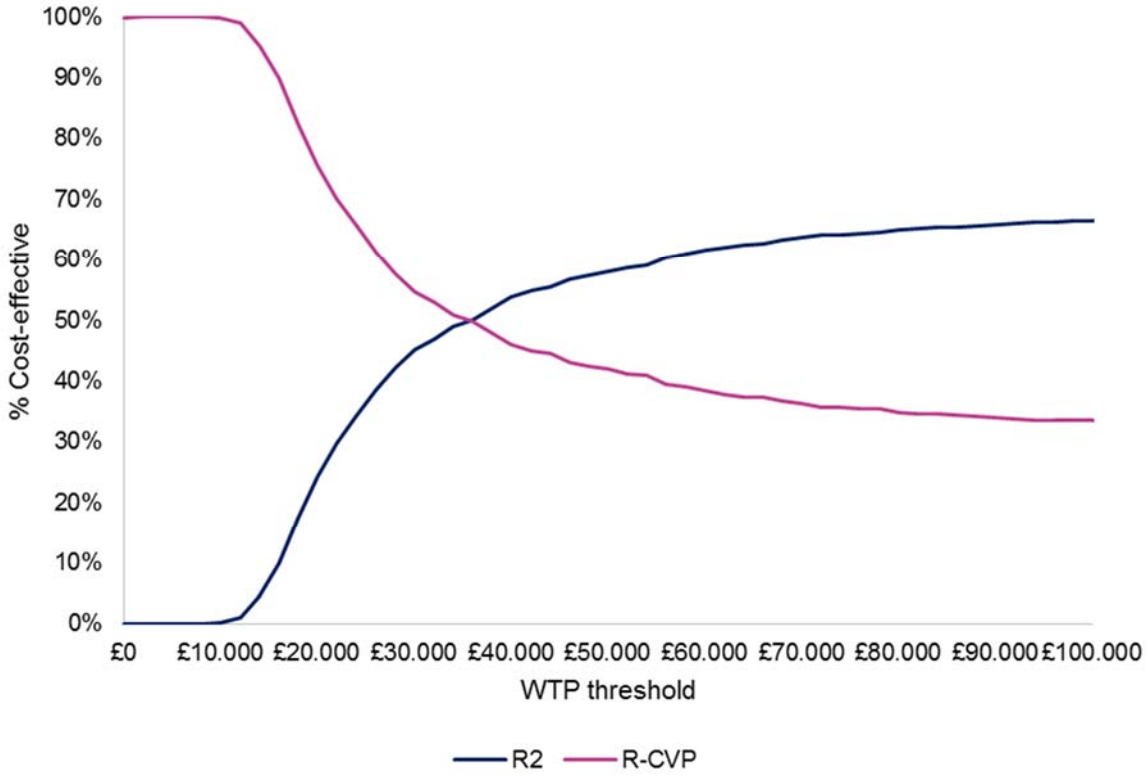


Figure 5.13: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: exponential OS

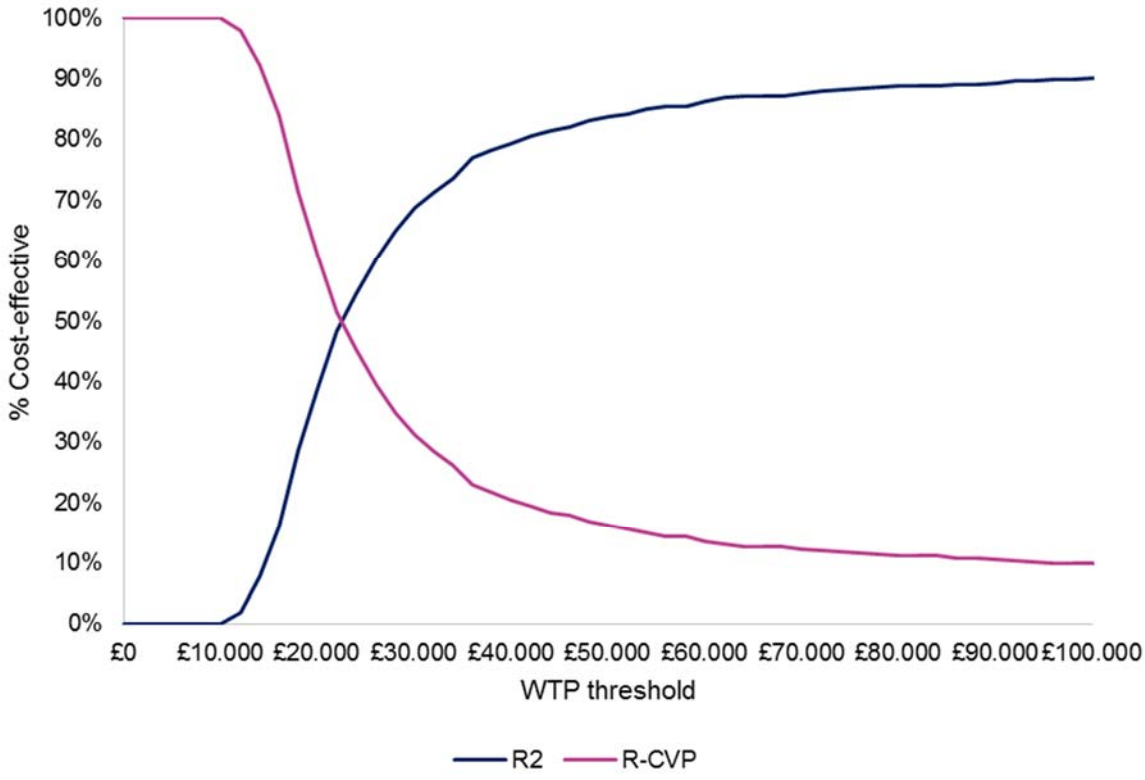


Figure 5.14: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: log-normal OS

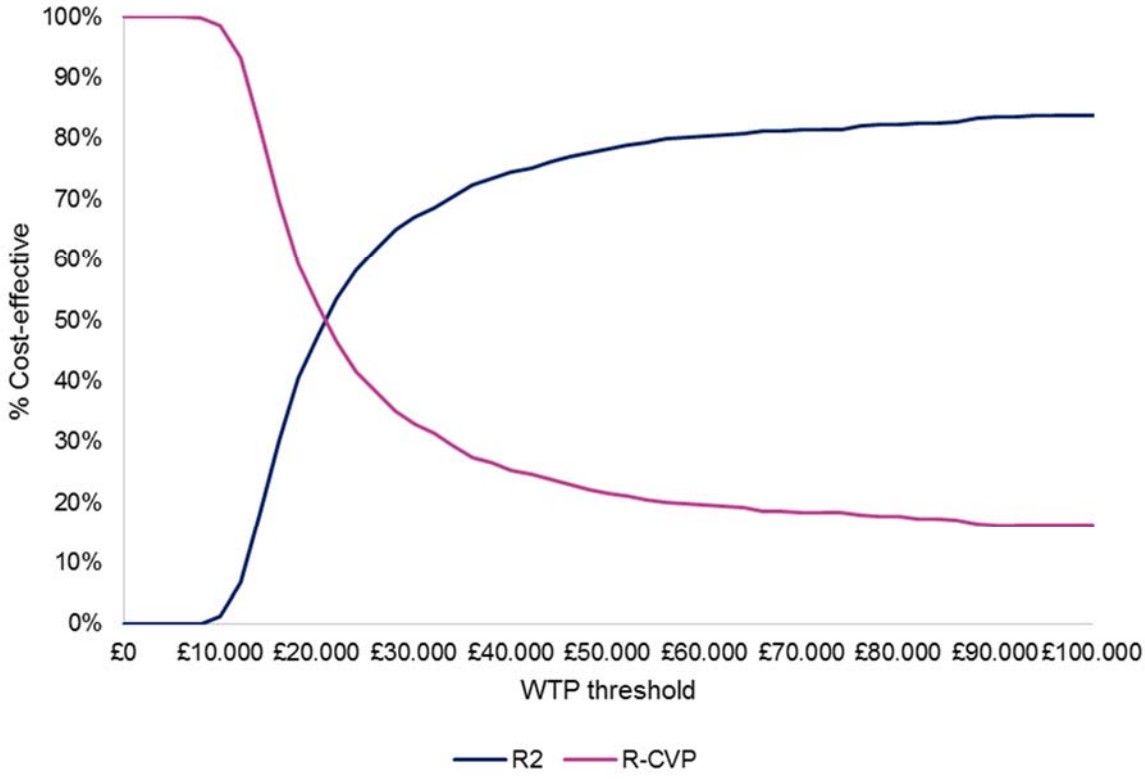


Figure 5.15: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: log-logistic OS

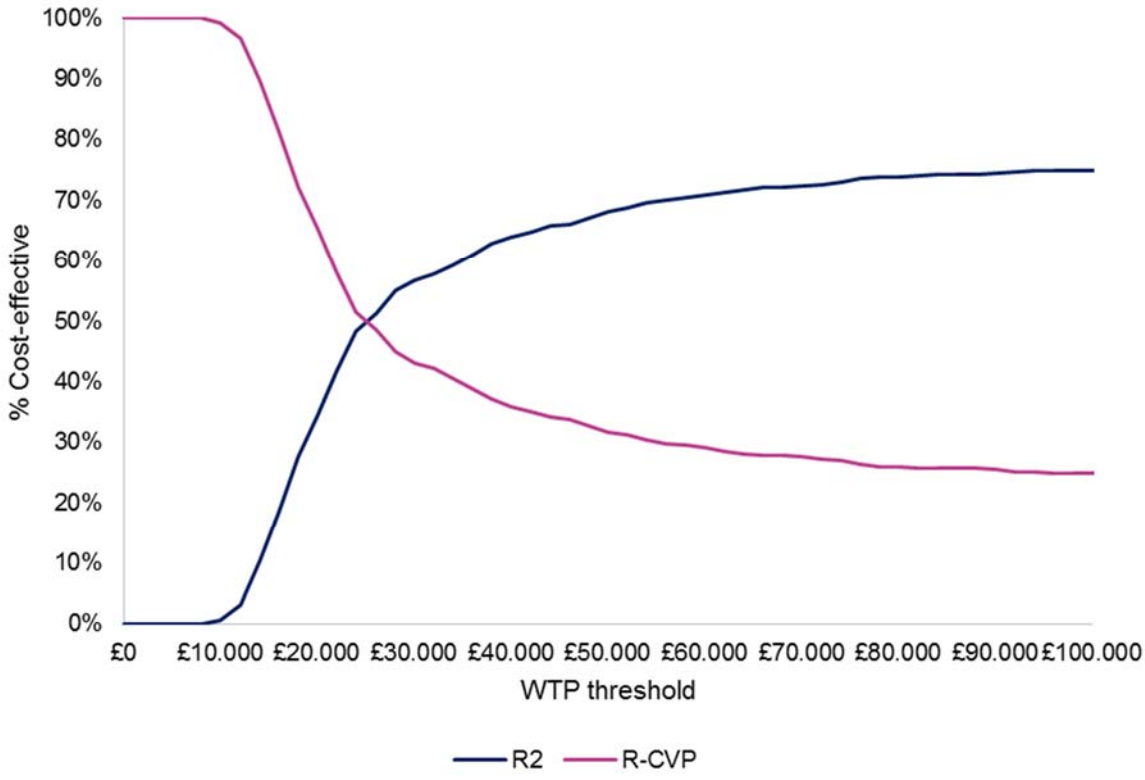


Figure 5.16: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: generalized gamma OS

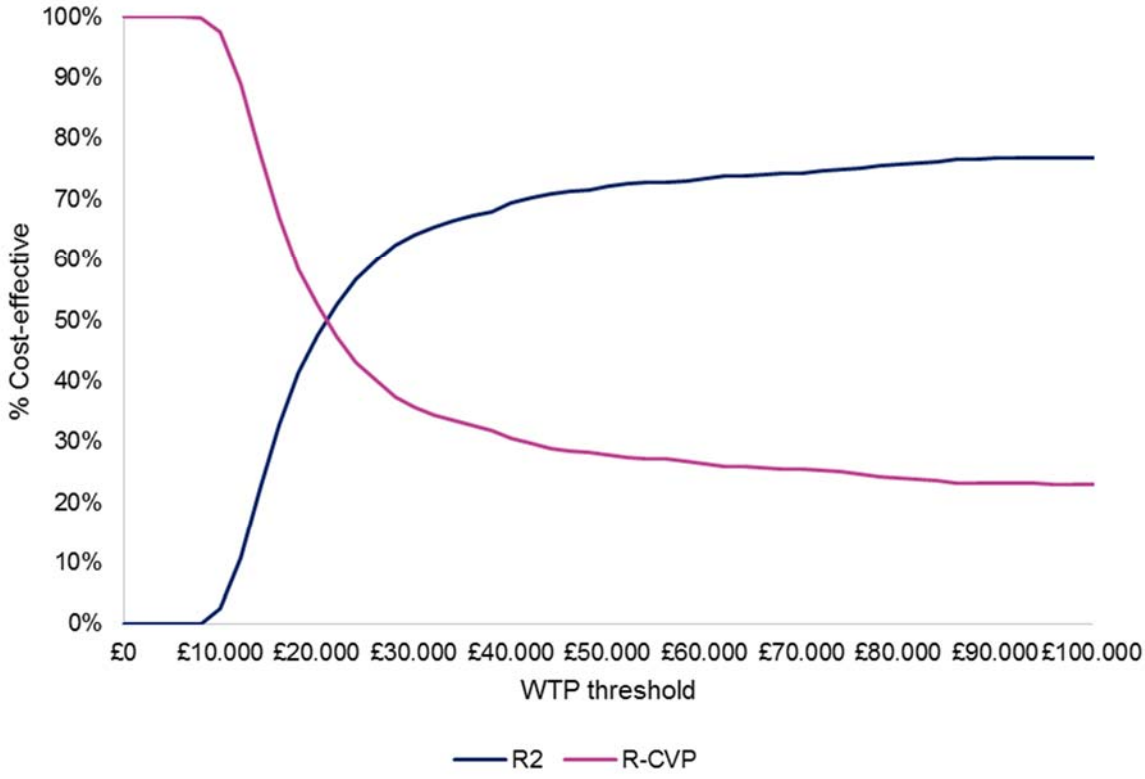


Figure 5.17: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: Gompertz OS

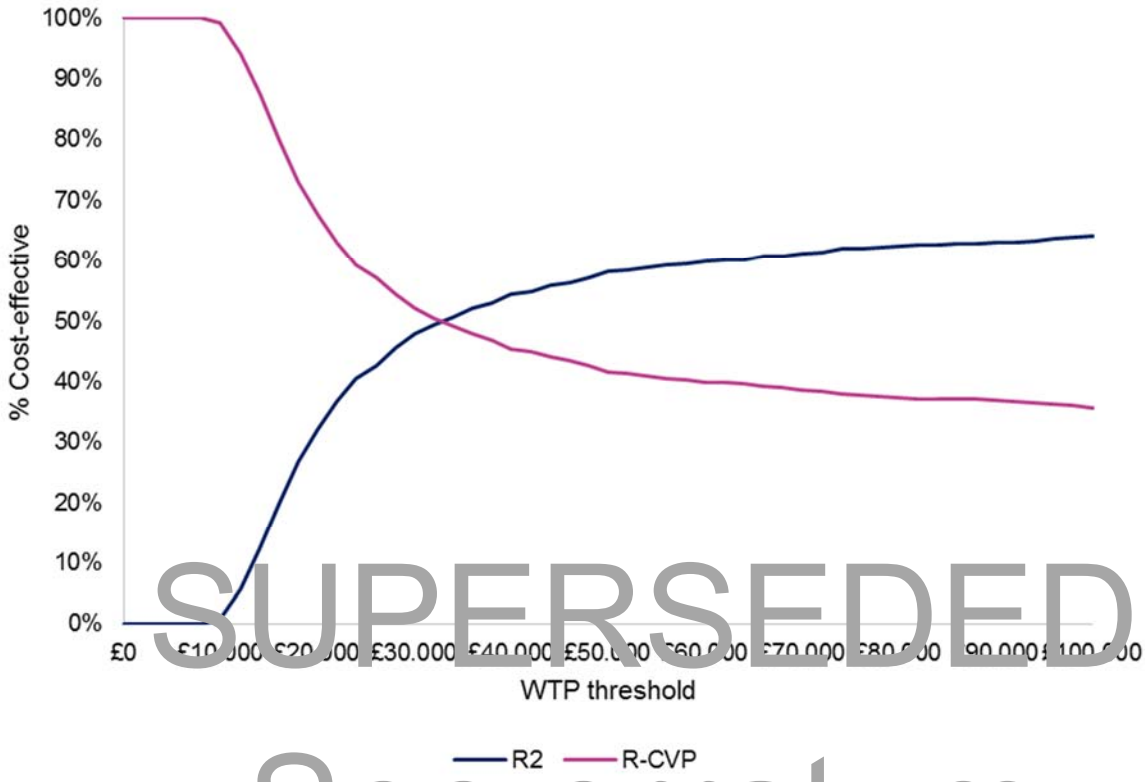


Figure 5.18: ERG base-case cost effectiveness acceptability curve for R2 versus R-mono: Weibull OS

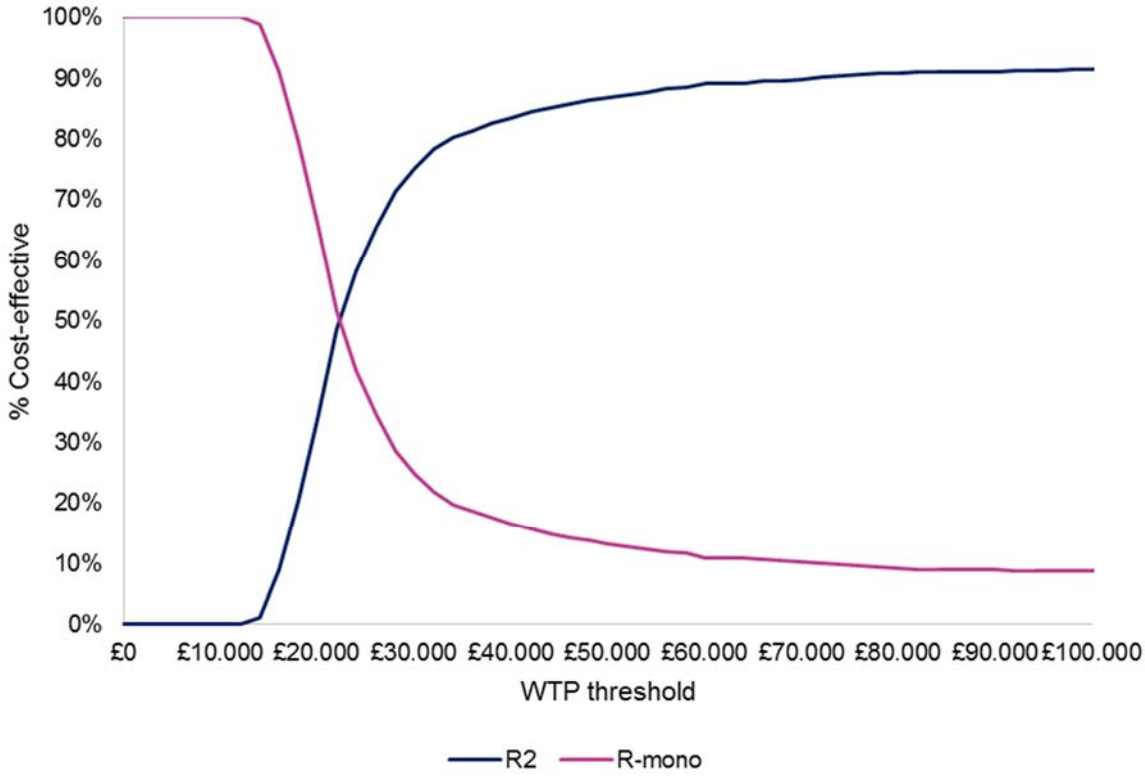


Figure 5.19: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: exponential OS

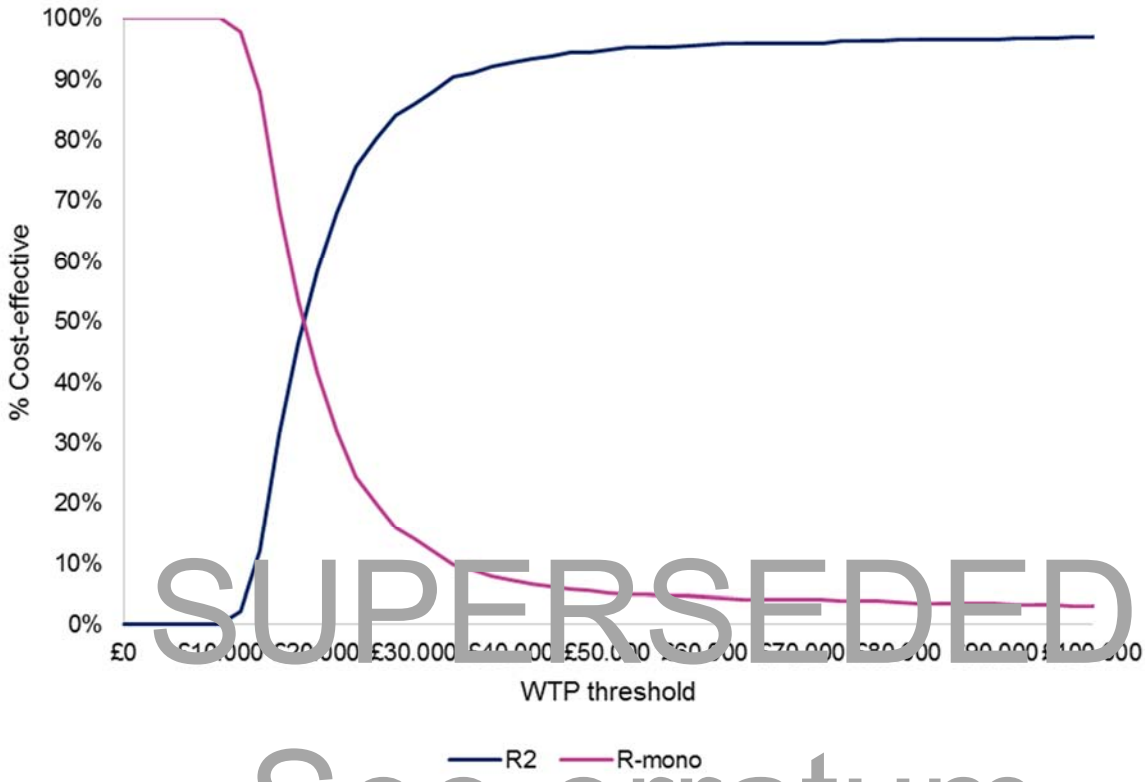


Figure 5.20: ERG base-case cost-effectiveness acceptability curve for R² versus R-mono: log-normal OS

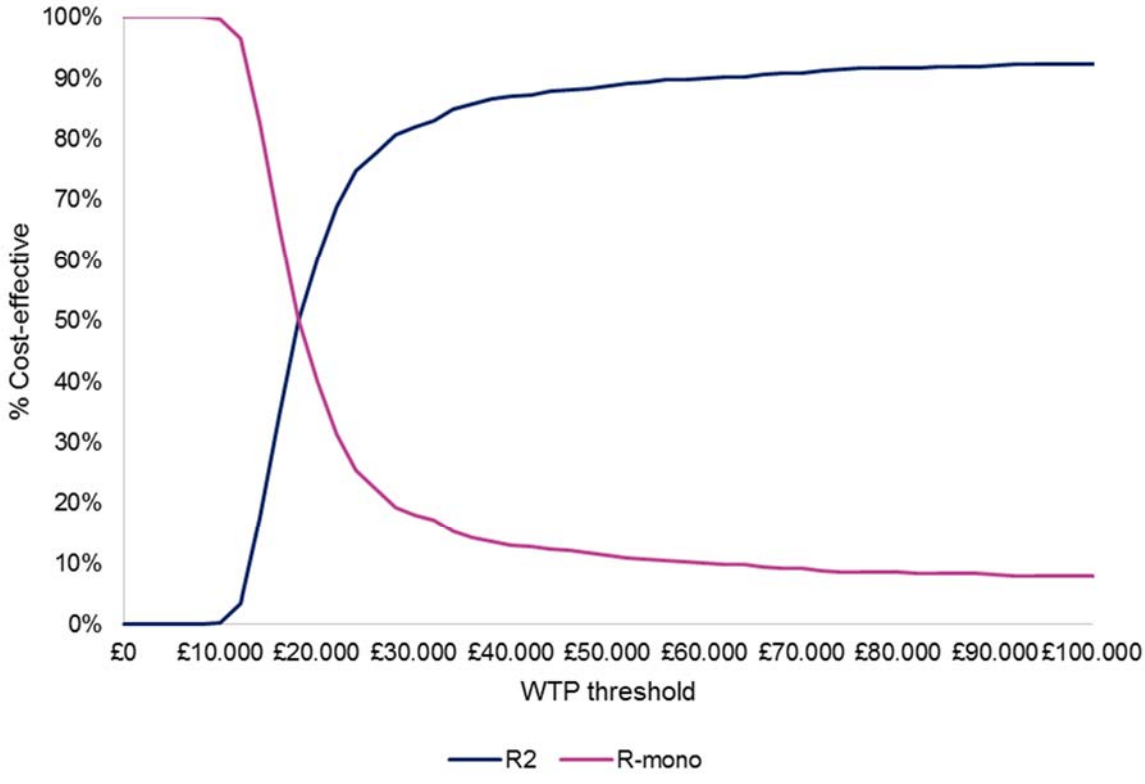


Figure 5.21: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: log-logistic OS

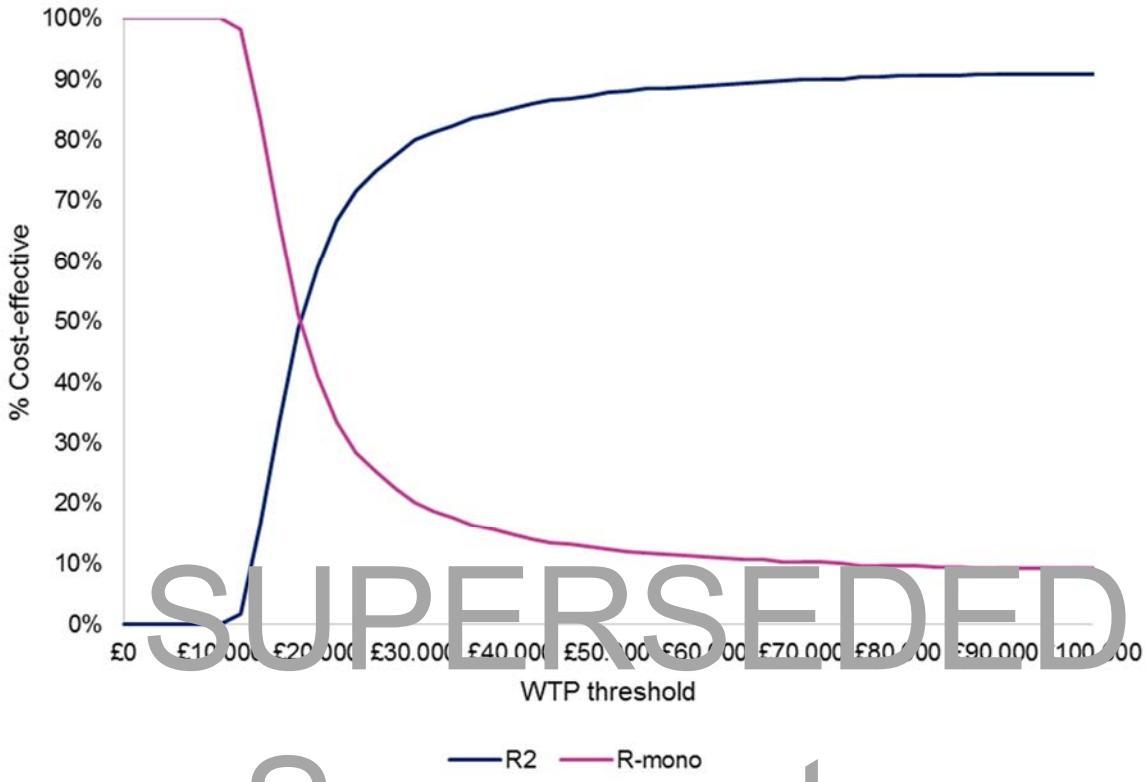


Figure 5.22: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: generalized gamma OS

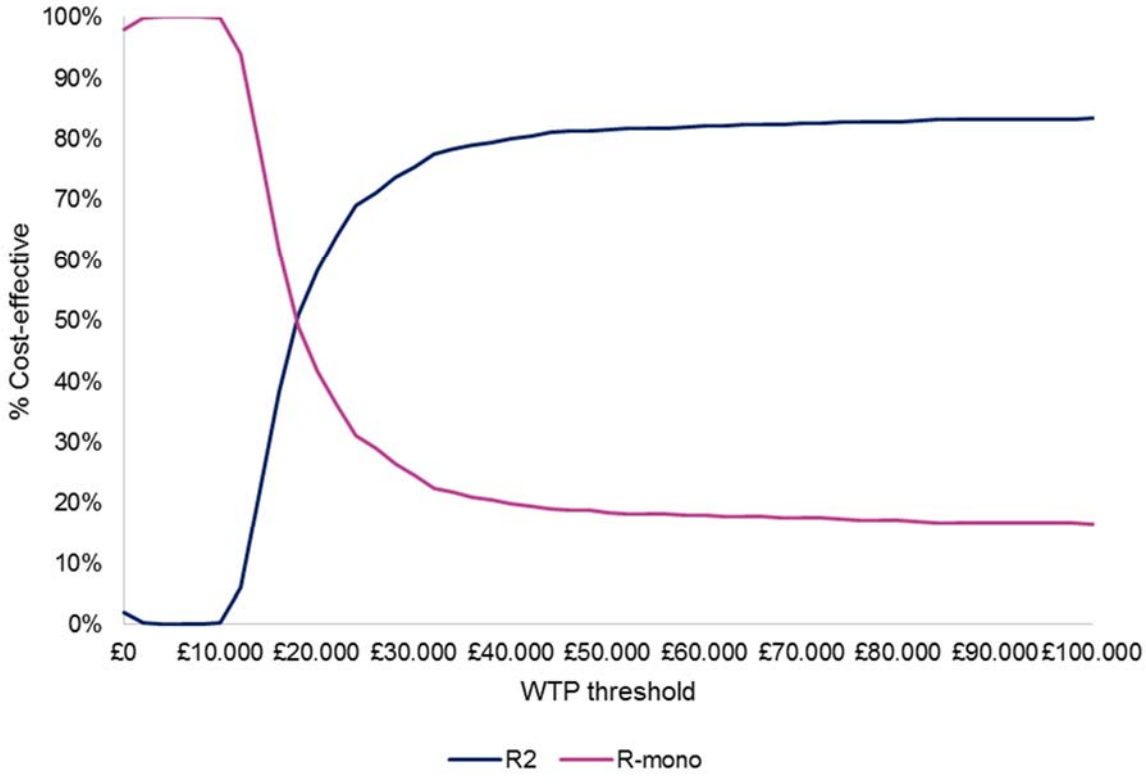
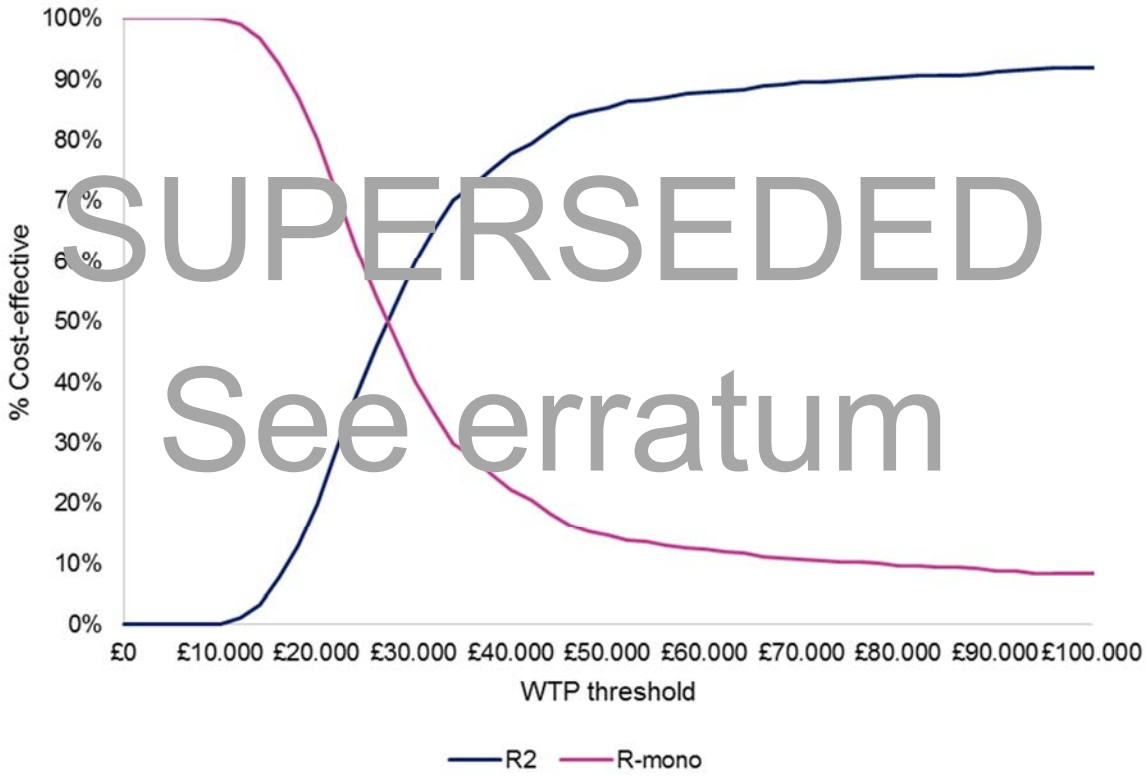


Figure 5.23: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: Gompertz OS



5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These were all performed using the ERG base-case. Results are presented in Tables 6.19 to 6.36 in Section 6 of this report.

Exploratory analyses using the ERG base-case:

1. Alternative PFS distributions: use Weibull for PFS both arms (for the R-mono comparison, generalised gamma was used as the alternative PFS distribution) (section 5.2.6)
2. Alternative PFS distributions: use exponential For PFS R^2 and Weibull for PFS comparator (not applied to R-mono comparison) (section 5.2.6)
3. Treatment waning effect after three-year cut-off (section 5.2.6)
4. Treatment waning effect after seven-year cut-off (section 5.2.6)
5. Adverse events for comparator taken from Van Oers et al. (2006)[11] (Not applicable in R-mono comparison) (section 5.2.7)
6. Apply same subsequent treatment costs for R^2 as for R-CHOP/R-CVP (Not applicable in R-mono comparison) (section 5.2.9)
7. Alternative utilities taken from Wild et al. (2006)[22] 0.805 for PF, 0.736 for PP off treatment, and 0.62 for PP on treatment (section 5.2.8)
8. Source for R-CHOP efficacy taken from Van Oers et al. (Not applicable in R-mono comparison) (section 5.2.6)
9. Alternative utilities taken for PP states taken from Pereira et al. (2010)[23] 0.45 for both PP states. (section 5.2.8)

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

The main concern of the ERG in the original ERG report[1] was the questionable trustworthiness of R^2 efficacy resulting from the indirect comparison, which seemed to be inflated relative to the direct comparison data from AUGMENT. Although the ERG did not have the necessary data to quantify this uncertainty, it may have lowered the ICER substantially. This issue still applies to all the analyses presented here. The likely overestimation of utility values also still applies.

The ERG had concerns about the way survival curves were selected and validated. For the FL only analyses presented in the company addendum, overall survival as predicted by the parametric survival curves was very different from overall survival in the original submission. No clinical validation of these new OS curves was performed.

The ERG made various adjustments to the company base-case in the addendum.[2] The probabilistic ERG base-case for R^2 versus R-CHOP ranged from £16,874 to £44,881 per QALY gained (based on 1,000 iterations). For R^2 versus R-CVP, the ICER ranged from £22,155 to £59,810 and for R^2 versus R-mono, it ranged from £18,779 to £27,156.

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost-effectiveness estimates. For the R-CHOP/R-CVP comparisons, using R-CHOP and R-CVP efficacy from van Oers et al. would change the ICER substantially, but not always in the same direction. Alternative assumptions regarding lowered utilities in the PP health states and the time point at which treatment waning start could also change the ICER substantially, dependent on the OS curves chosen. In general, for the R-CHOP/R-CVP comparison it can be said that the model seems instable and results are highly dependent on the assumptions applied, with ICERs ranging between dominant and dominated. For the R-mono comparison, the ICERs are much less volatile, but still ranging between £11,539 and £42,311.

Of note, a full incremental analysis would result in R-CHOP being strictly dominated by definition (being equally effective and more costly than R-CVP), and the relevant ICER would therefore always be R^2 versus R-CVP. For R-mono, a full incremental analysis is not applicable, because costs and QALYs for R^2 are different in this comparison.

The main conclusion of the original ERG report[1] still applies, that is, even though the ERG base-case ICER for R-CHOP was below £20,000, the uncertainty around the cost effectiveness of R^2 is substantial, mainly caused by the possible bias introduced by the indirect treatment comparison, which could not be accounted for in the ERG analyses. In addition, specific to the FL only population analyses presented in the company addendum,[2] the uncertainty around the OS estimates and the lack of clinical validation of these estimates would warrant even more caution in the interpretation of results. The ICER for R-CVP is higher and suffers from the same uncertainty.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

6.1 Analyses undertaken by the ERG

In Section 5.3 the ERG base-case was presented, which was based on various changes compared to the company base-case. Tables 6.1 to 6.18 show how individual changes impact the results plus the combined effect of all changes simultaneously, for the R-CHOP, R-CVP, and R-mono comparators, and all possible OS curves, respectively. The exploratory scenario analyses are presented in Tables 6.19 to 6.36 respectively. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.1 to 6.18 correspond to the analyses numbers reported in Section 5.3 of this report. The submitted model files contain technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment). The ERG wishes to emphasise that all ERG analyses (except the R-mono comparison) are conditional upon the MAIC results for which uncertainty could not be quantified or incorporated in the economic model.

Although the tables below report pairwise comparisons only, R-CHOP and R-CVP could also be compared to R^2 in a full incremental analysis. However, as R-CHOP and R-CVP are by assumption equally effective, and R-CHOP is always the more costly strategy given the higher rate of ASCT performed in the R-CHOP patient population, it is not to be expected that there will be any shifts in the relative comparisons within the full incremental analysis. Therefore, in practice, the relevant comparison will be R^2 versus R-CVP. For R-mono, a full incremental analysis on the scenarios is not applicable, as a different set of scenarios was performed here, and, more importantly, because costs and QALYs in R^2 are different in this comparison.

Table 6.1: Deterministic ERG base-case for R^2 versus R-CHOP comparison: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use weibull for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			

ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,781
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£44,888
R-CHOP	[redacted]	[redacted]			

Table 6.2: Deterministic ERG base-case for R^2 versus R-CHOP comparison: Exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,651
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,581
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,138
R-CHOP	[redacted]	[redacted]			

Table 6.3: Deterministic ERG base-case for R^2 versus R-CHOP comparison: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			

Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use log-normal for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,245
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,531
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,177
R-CHOP	[redacted]	[redacted]			

Table 6.4: Deterministic ERG base-case for R^2 versus R-CHOP comparison: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use log-logistic for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,955
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,146
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,800
R-CHOP	[redacted]	[redacted]			

Table 6.5: Deterministic ERG base-case for R^2 versus R-CHOP comparison: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use gengamma for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£10,191
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,941
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,874
R-CHOP	[redacted]	[redacted]			

Table 6.6: Deterministic ERG base-case for R^2 versus R-CHOP comparison: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,909
R-CHOP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,002
R-CHOP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,475
R-CHOP	[redacted]	[redacted]			
Matter of judgement (4, use gompertz for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,765
R-CHOP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,866
R-CHOP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,935
R-CHOP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,019
R-CHOP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,229
R-CHOP	[redacted]	[redacted]			

Table 6.7: Deterministic ERG base-case for R^2 versus R-CVP comparison: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use weibull for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746

R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,404
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£59,810
R-CVP	[redacted]	[redacted]			

Table 6.8: Deterministic ERG base-case for R^2 versus R-CVP comparison: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,394
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,742
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,583
R-CVP	[redacted]	[redacted]			

Table 6.9: Deterministic ERG base-case for R^2 versus R-CVP comparison: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use log-normal for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,071
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,658
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,135
R-CVP	[redacted]	[redacted]			

Table 6.10: Deterministic ERG base-case for R^2 versus R-CVP comparison: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use log-logistic for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,886

R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,529
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£32,899
R-CVP	[redacted]	[redacted]			

Table 6.11: Deterministic ERG base-case for R^2 versus R-CVP comparison: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use gengamma for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,339
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,312
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,778
R-CVP	[redacted]	[redacted]			

Table 6.12: Deterministic ERG base-case for R^2 versus R-CVP comparison: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use gompertz for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,863
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,767
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£43,915
R-CVP	[redacted]	[redacted]			

Table 6.13: Deterministic ERG base-case for R^2 versus R-mono comparison: Weibull OS

Technologies	Total cost	Total QALYs	Incremental costs	Incremental QALY	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use weibull for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378

R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,659
R-mono	[redacted]	[redacted]			

Table 6.14: Deterministic ERG base-case for R^2 versus R-mono comparison: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,105
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,856
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,779
R-mono	[redacted]	[redacted]			

Table 6.15: Deterministic ERG base-case for R^2 versus R-mono comparison: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use log-normal for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,222
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,884
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,326
R-mono	[redacted]	[redacted]			

Table 6.16: Deterministic ERG base-case for R^2 versus R-mono comparison: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use log-logistic for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,661
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,366
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,027
R-mono	[redacted]	[redacted]			

Table 6.17: Deterministic ERG base-case for R^2 versus R-mono comparison: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use gengamma for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,993
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,457
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,462
R-mono	[redacted]	[redacted]			

Table 6.18: Deterministic ERG base-case for R^2 versus R-mono comparison: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,310
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use gompertz for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,208
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,625
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,15[illegible]
R-mono	[redacted]	[redacted]			

Table 6.19: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,781
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,163
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,630
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,107
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£39,668
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,482
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,036
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£32,035
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£667
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£56,826
R-CHOP	[redacted]	[redacted]			

Table 6.20: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,581
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,558
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,061
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,142
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,626
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,208
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,632
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,299
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£37,306
R-CHOP	[redacted]	[redacted]			

Table 6.21: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,531
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,373
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,218
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,535
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,583
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,712
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,059
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,297
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£32,101
R-CHOP	[redacted]	[redacted]			

Table 6.22: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,146
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,186
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,525
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,651
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,313
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,870
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,293
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,403
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£40,152
R-CHOP	[redacted]	[redacted]			

Table 6.23: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: generalizd gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£12,941
R-CHOP	[redacted]	[redacted]			

Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,660
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,821
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,136
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,944
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,797
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,097
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,853
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,541
R-CHOP	[redacted]	[redacted]			

Table 6.24: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,019
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,284
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,046
R-CHOP	[redacted]	[redacted]			

Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,220
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
R-CHOP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,337
R-CHOP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,833
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,190
R-CHOP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
R-CHOP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£50,728
R-CHOP	[redacted]	[redacted]			

Table 6.25: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,404
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£31,788
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,248
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,206
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£55,546
R-CVP	[redacted]	[redacted]			

Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£32,977
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£34,667
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£44,760
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,530
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£79,588
R-CVP	[redacted]	[redacted]			

Table 6.26: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,742
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,720
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,220
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,542
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,423
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,563
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,797
R-CVP	[redacted]	[redacted]			

Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£31,975
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£4,712,070
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£51,262
R-CVP	[redacted]	[redacted]			

Table 6.27: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,658
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,501
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,344
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,152
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,746
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,168
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£22,189
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,472
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£2,064,117
R-CVP	[redacted]	[redacted]			

Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£43,492
R-CVP	[redacted]	[redacted]			

Table 6.28: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,529
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,570
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,905
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,910
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,006
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,417
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,681
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£33,509
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,557
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£55,212
R-CVP	[redacted]	[redacted]			

Table 6.29: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,312
R-CVP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,031
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,190
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,097
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,512
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,595
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,470
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,893
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£2,064,117
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£36,886
R-CVP	[redacted]	[redacted]			

Table 6.30: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CVP: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,767
R-CVP	[redacted]	[redacted]			

Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,034
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,791
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,657
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,072
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£31,589
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£40,523
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£2,064,117
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£70,564
R-CVP	[redacted]	[redacted]			

Table 6.31: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALY	CER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,378
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,887
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£36,355
R-mono	[redacted]	[redacted]			

Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,138
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,090
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,514
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,313
R-mono	[redacted]	[redacted]			

Table 6.32: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,856
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,280
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,468
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,455
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,001
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,380
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,065
R-mono	[redacted]	[redacted]			

Table 6.33: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,884
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,346
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,920
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,774
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,878
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,193
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,778
R-mono	[redacted]	[redacted]			

Table 6.34: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,366
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,815
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28.463
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13.764
R-mono	[redacted]	[redacted]			

Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19.438
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16.086
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14.922
R-mono	[redacted]	[redacted]			

Table 6.35: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: generalised gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,457
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,015
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,926
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,539
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,076
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,220
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,957
R-mono	[redacted]	[redacted]			

Table 6.36: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,625
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,246
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£42,311
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,995
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,988
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,373
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,361
R-mono	[redacted]	[redacted]			

7. REFERENCES

[1] Riemsma R, Van Asselt ADI, Witlox W, Huertas Carrera V, Posadzki P, Armstrong N, et al. *Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma: a Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd, 2019

[2] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Addendum for the amended follicular lymphoma only population. Company evidence submission to National Institute for Health and Care Excellence. Single technology appraisal (STA)*: Celgene Ltd, 2019

[3] Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0086) March 2019. (Updated: August 2019) Data on file. 2019.

[4] Celgene Ltd. [redacted] clinical validation. 12 June 2019. Data on file. 2019.

[5] National Institute for Health and Care Excellence. *Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2019 [accessed 25.7.19] Available from: https://www.nice.org.uk/guidance/gid-ta10323/documents/final-scope

[6] 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.

[7] Andorsky D, Coleman M, Yacoub A, Melear JM, Brooks HD, Fanning SR, et al. Response rate to lenalidomide plus rituximab (R2) as independent of number of prior lines of therapy: interim analysis of initial phase of MAGNIFY phase IIIb study of R2 followed by maintenance in relapsed/refractory indolent NHL. *J Clin Oncol* 2018;36(15 Suppl):7516.

[8] Leonard JP, Jung SH, Johnson J, Pitcher BN, Bartlett NL, Blum KA, et al. Randomized trial of lenalidomide alone versus lenalidomide plus rituximab in patients with recurrent follicular lymphoma: CALGB 50401 (ALLIANCE). *J Clin Oncol* 2015;33(31):3635-40.

[9] Tuscano JM, Dutia M, Chee K, Brunson A, Reed-Pease C, Abedi M, et al. Lenalidomide plus rituximab can produce durable clinical responses in patients with relapsed or refractory, indolent non-Hodgkin lymphoma. *Br J Haematol* 2014;165(3):375-81.

[10] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Document B. Appendices. Single technology appraisal (STA)*: Celgene Ltd, 2019

[11] van Oers MH, 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.

[12] Sehn LH, Chua N, Mayer J, Dueck G, Trněný 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.

[13] Celgene Ltd, BresMed. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Cost-effectiveness model: R2 for previously treated follicular lymphoma (FL) and marginal zone lymphoma (MZL). v2.1 (ERG response) [Excel]*: BresMed, 2019

[14] Celgene Ltd. *Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Response to request for clarification from the ERG. August 2019*: Celgene Ltd, 2019

[15] Celgene Ltd. *Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]. Document B. Company evidence submission to National Institute for Health and Care Excellence. Single technology appraisal (STA)*: Celgene Ltd, 2019

[16] Celgene. *Draft summary of product characteristics. Revlimid® (Lenalidomide)*: Celgene, 2019

[17] Leonard JP, Trneny M, Izutsu K, et al. AUGMENT phase III study: Lenalidomide/rituximab (R2) improved efficacy over rituximab/placebo in relapsed/refractory follicular patients irrespective of POD24 status. Paper submitted to 15-ICML International Conference on Malignant Lymphoma; 18-22 Jun 2019; Lugano: Switzerland. 2019.

[18] Haematological Malignancy Research Network (HMRN). Clinical management and outcome of follicular lymphoma (FL) with a focus on relapsed/refractory disease (draft report). 17 July 2019. Haematological Malignancy Research Network, 2019.

[19] National Institute for Health and Care Excellence. *Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma. NICE technology appraisal guidance 137 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.2.19] Available from: https://www.nice.org.uk/guidance/ta137

[20] National Institute for Health and Care Excellence, Centre for Health Technology Evaluation. *Single Technology Appraisal. Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]. ERG report - factual accuracy check*. London: National Institute for Health and Care Excellence, 2019 [accessed 24.10.19]

[21] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

[22] Wild D, Walker M, Pettengell R, Lewis G. Pcn62 utility elicitation in patients with follicular lymphoma. *Value Health* 2006;9(6):A294.

[23] Pereira C, Negreiro F, Silva C. Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. PSY43. *Value Health* 2010;13(7):A468.

Appendix 1: Additional results from the AUGMENT trial

There is no change from the ERG report.[1]

Appendix 2: MAIC reporting checklist

There is no change from the ERG report.[1]

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma:

ERRATUM to
Addendum to the ERG report in response to company addendum for the amended follicular lymphoma only population

This document contains errata with respect to the addendum to the ERG report. The ERG noted that in the company model, the default for the time to next anti-lymphoma treatment (TTNLT) curve was set to generalized gamma, while in the company addendum the curve that was actually used was log-normal. This only affected the analyses for the R-mono comparison as in the other comparisons, the ERG actively changed the TTNLT curve. The ERG has therefore re-ran the R-mono analyses and provides here the corrected results in tables as well as in the text of the report.

The table below lists the page to be replaced in the original document and the nature of the change:

Page nr:	**Change:**
49-51	Text and Tables 5.10 and 5.11 deterministic and probabilistic results R-mono
57-60	Figures 5.17 to 5.23 CEACs for R-mono comparison
61	Conclusions: ICERs for R-mono comparison
70-73	Tables 6.13 to 6.18 ERG base-case results for R-mono comparison
83-87	Tables 6.31 to 6.36 ERG scenarios for R-mono comparison

which resulted in an ICER that ranged from £17,312 to £30,404. Finally, for the R-mono deterministic comparison, incremental costs varied from [redacted] to [redacted] and incremental QALYs from [redacted] to [redacted] with resulting ICERs ranging from £14,504 to £25,535.

The probabilistic ERG base-case (based on 1,000 iterations) for R^2 versus R-CHOP ranged from £16,874 to £44,888. For R^2 versus R-CVP, the ICER ranged from £23,135 to £59,810 and for R^2 versus R-mono, it ranged from £18,816 to £26,728. Compared with the deterministic base-case results, the ERG PSA resulted in higher ICERs, similar to what was seen in the company analyses. Particularly for the Weibull and Gompertz OS curves in the R-CHOP and R-CVP comparisons, the probabilistic ICER would sometimes be around twice the value of the deterministic ICER. For all the other OS curves, the differences between deterministic and probabilistic ICERs were more modest, although still considerable at times (see Table 5.11). The CEACs of all analyses are presented in Figures 5.6 to 5.23.

Table Error! No text of specified style in document.**.1: ERG pairwise deterministic base-case results**

Technologies	OS curve	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case for R^2 versus R-CHOP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£21,781
R-CHOP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£16,581
R-CHOP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£14,531
R-CHOP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£17,146
R-CHOP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£12,941
R-CHOP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£20,019
R-CHOP	Gompertz	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-CVP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£30,404
R-CVP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£22,742
R-CVP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£19,658
R-CVP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£23,529
R-CVP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£17,312
R-CVP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£27,767
R-CVP	Gompertz	[redacted]	[redacted]			
Deterministic ERG base-case for R^2 versus R-mono						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£21,341

R-mono	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£17,931
R-mono	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£16,951
R-mono	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£17,432
R-mono	Log-logistic	[redacted]	[redacted]			
R^2	Gengamma	[redacted]	[redacted]	[redacted]	[redacted]	£14,504
R-mono	Gengamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£25,535
R-mono	Gompertz	[redacted]	[redacted]			
ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						

Table Error! No text of specified style in document.**.2: ERG probabilistic base-case results**

Technologies	OS curve	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Probabilistic ERG base-case for R^2 versus R-CHOP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£44,888
R-CHOP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£17,138
R-CHOP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£17,177
R-CHOP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£20,800
R-CHOP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£16,874
R-CHOP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£30,229
R-CHOP	Gompertz	[redacted]	[redacted]			
Probabilistic ERG base-case for R^2 versus R-CVP						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£59,810
R-CVP	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£23,583
R-CVP	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£23,135
R-CVP	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£32,899
R-CVP	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£24,778
R-CVP	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£43,915
R-CVP	Gompertz	[redacted]	[redacted]			

Probabilistic ERG base-case for R^2 versus R-mono						
R^2	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£24,958
R-mono	Weibull	[redacted]	[redacted]			
R^2	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£18,816
R-mono	Exponential	[redacted]	[redacted]			
R^2	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£19,169
R-mono	Log-normal	[redacted]	[redacted]			
R^2	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£19,775
R-mono	Log-logistic	[redacted]	[redacted]			
R^2	Gen gamma	[redacted]	[redacted]	[redacted]	[redacted]	£25,394
R-mono	Gen gamma	[redacted]	[redacted]			
R^2	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£26,728
R-mono	Gompertz	[redacted]	[redacted]			
ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						

Figure Error! No text of specified style in document.**.1: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CHOP: Weibull OS**

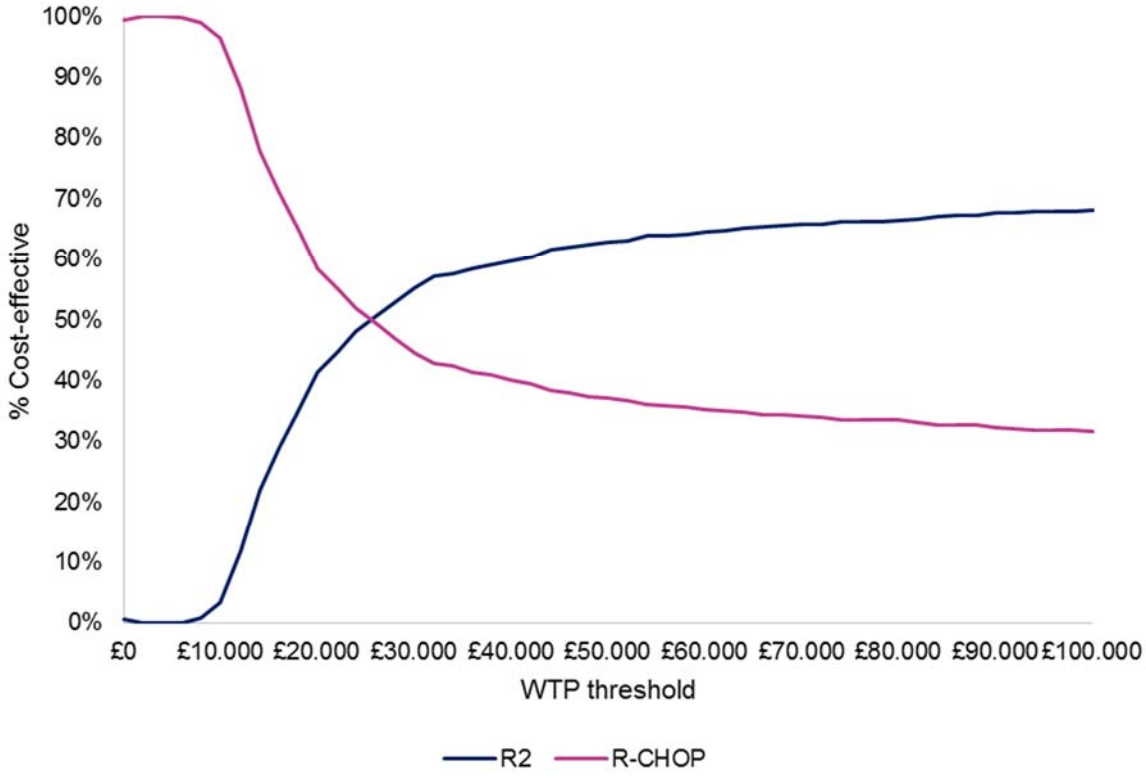


Figure Error! No text of specified style in document.**.2: ERG base-case cost effectiveness acceptability curve for R^2 versus R-CVP: Gompertz OS**

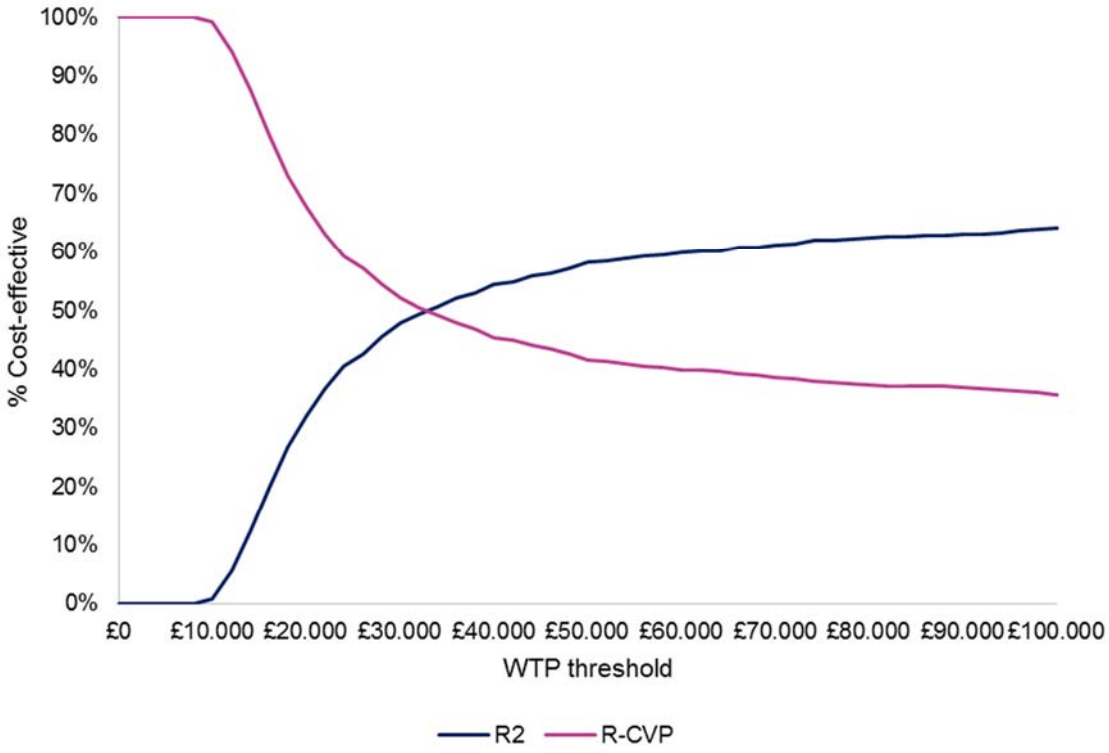


Figure Error! No text of specified style in document.**.3: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: Weibull OS**

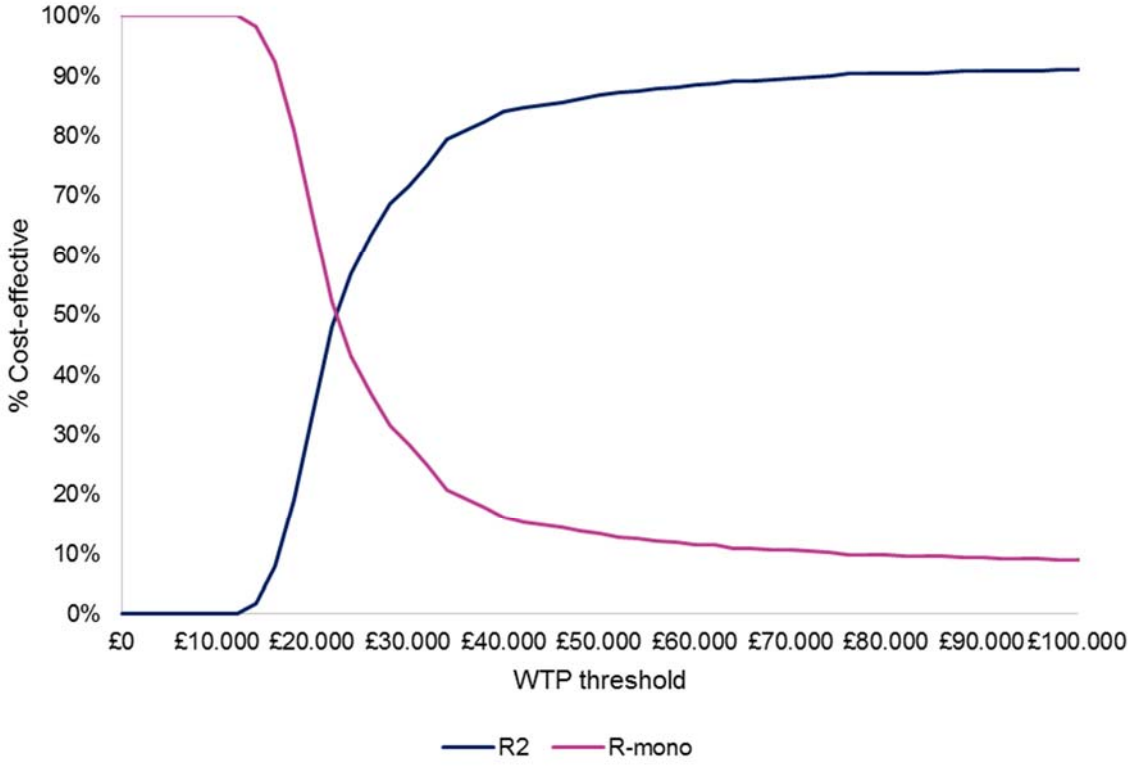


Figure Error! No text of specified style in document.**.4: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: exponential OS**

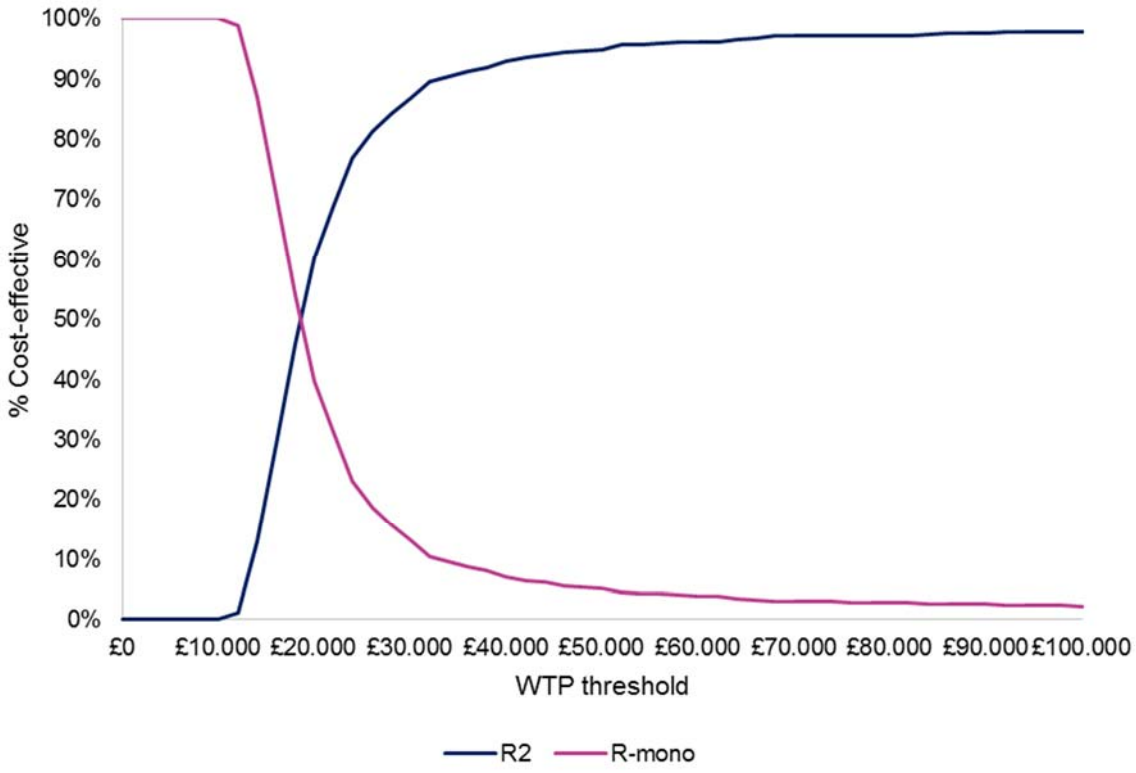


Figure Error! No text of specified style in document.**.5: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: log-normal OS**

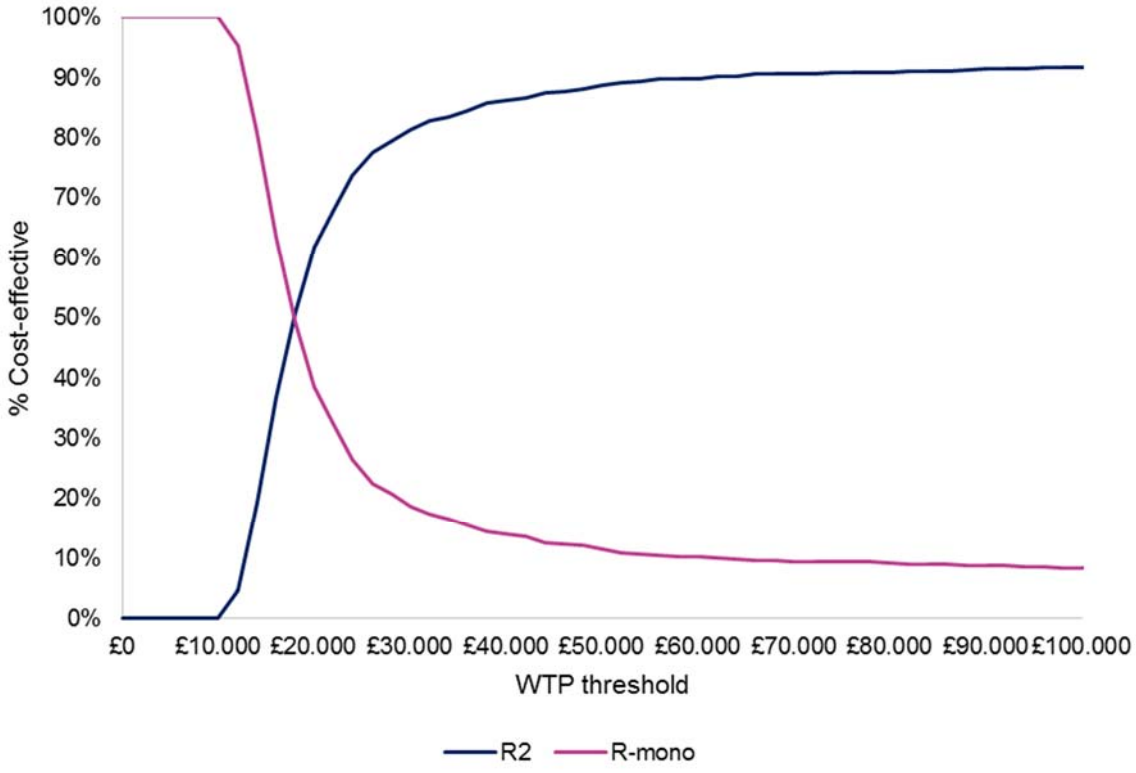


Figure Error! No text of specified style in document..**6: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: log-logistic OS**

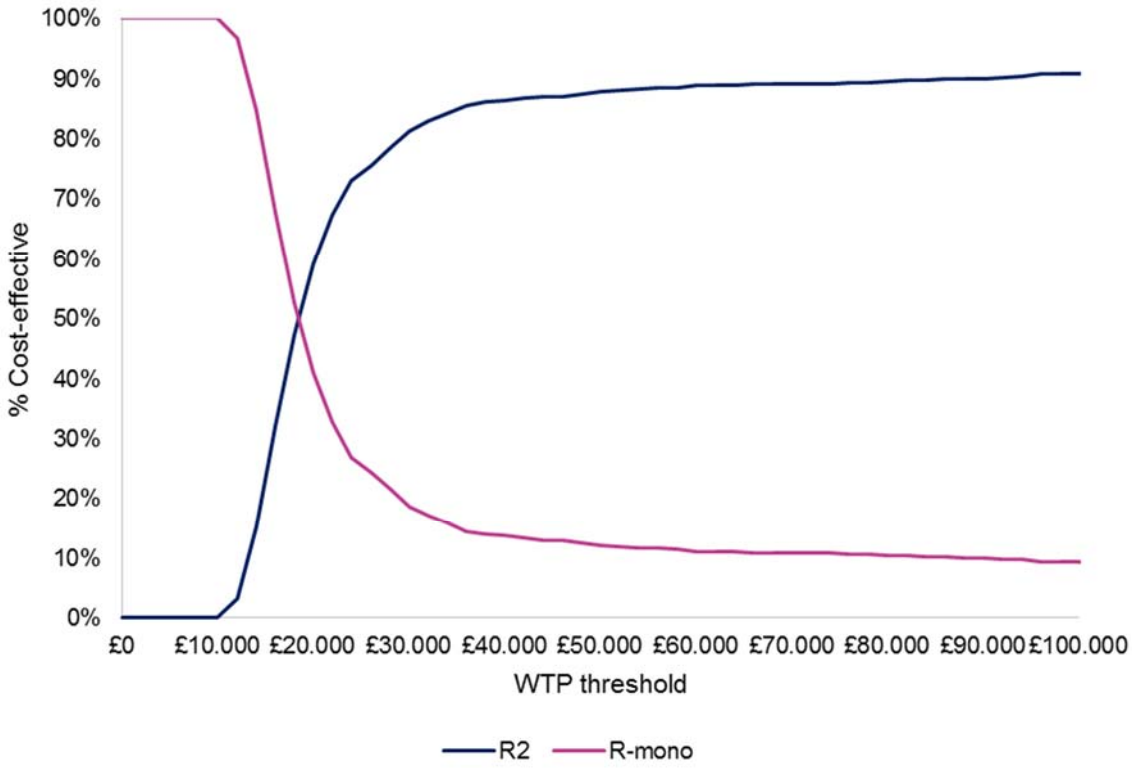


Figure Error! No text of specified style in document..**7: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: generalized gamma OS**

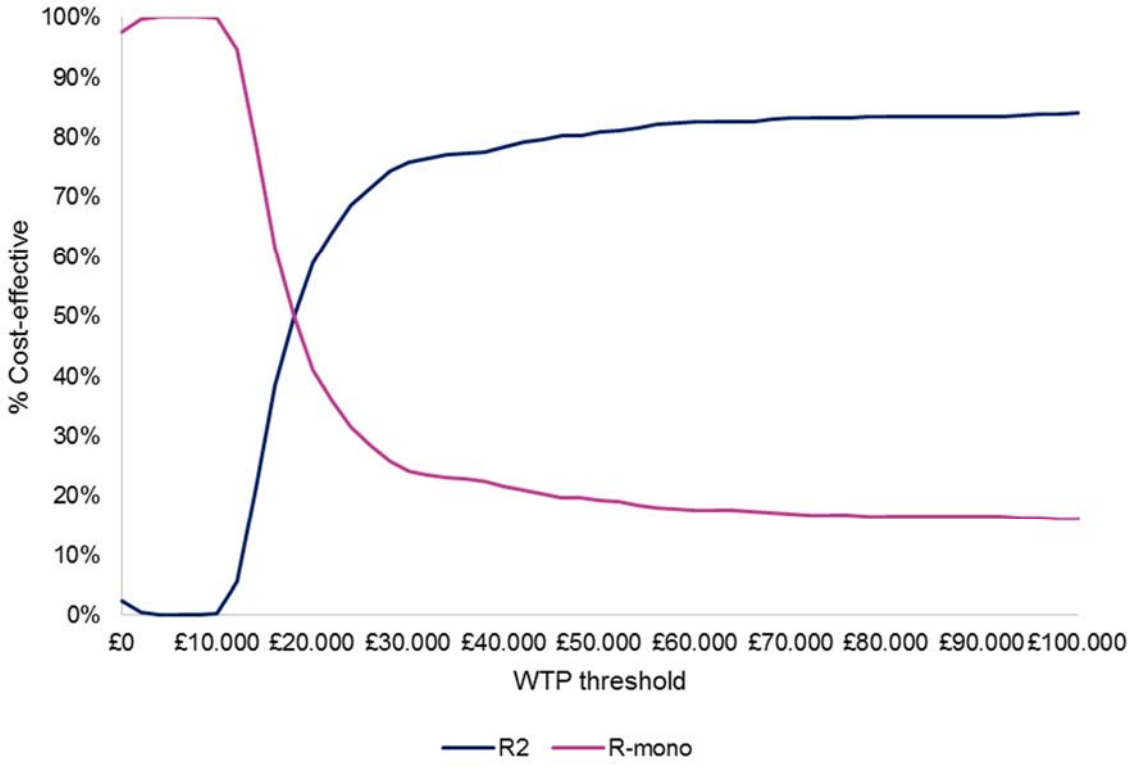
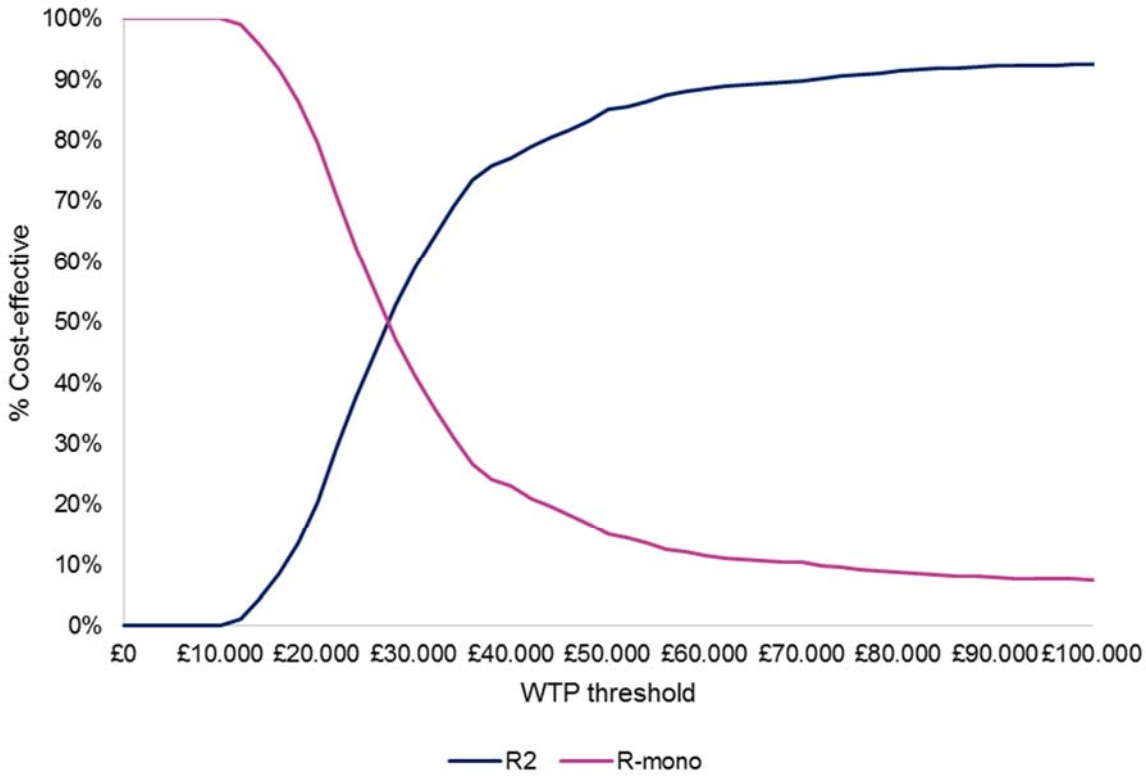


Figure Error! No text of specified style in document.**.8: ERG base-case cost effectiveness acceptability curve for R^2 versus R-mono: Gompertz OS**



5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These were all performed using the ERG base-case. Results are presented in Tables 6.19 to 6.36 in Section 6 of this report.

Exploratory analyses using the ERG base-case:

1. Alternative PFS distributions: use Weibull for PFS both arms (for the R-mono comparison, generalised gamma was used as the alternative PFS distribution) (section 5.2.6)
2. Alternative PFS distributions: use exponential For PFS R^2 and Weibull for PFS comparator (not applied to R-mono comparison) (section 5.2.6)
3. Treatment waning effect after three-year cut-off (section 5.2.6)
4. Treatment waning effect after seven-year cut-off (section 5.2.6)
5. Adverse events for comparator taken from Van Oers et al. (2006)[11] (Not applicable in R-mono comparison) (section 5.2.7)
6. Apply same subsequent treatment costs for R^2 as for R-CHOP/R-CVP (Not applicable in R-mono comparison) (section 5.2.9)
7. Alternative utilities taken from Wild et al. (2006)[22] 0.805 for PF, 0.736 for PP off treatment, and 0.62 for PP on treatment (section 5.2.8)
8. Source for R-CHOP efficacy taken from Van Oers et al. (Not applicable in R-mono comparison) (section 5.2.6)
9. Alternative utilities taken for PP states taken from Pereira et al. (2010)[23] 0.45 for both PP states. (section 5.2.8)

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

The main concern of the ERG in the original ERG report[1] was the questionable trustworthiness of R^2 efficacy resulting from the indirect comparison, which seemed to be inflated relative to the direct comparison data from AUGMENT. Although the ERG did not have the necessary data to quantify this uncertainty, it may have lowered the ICER substantially. This issue still applies to all the analyses presented here. The likely overestimation of utility values also still applies.

The ERG had concerns about the way survival curves were selected and validated. For the FL only analyses presented in the company addendum, overall survival as predicted by the parametric survival curves was very different from overall survival in the original submission. No clinical validation of these new OS curves was performed.

The ERG made various adjustments to the company base-case in the addendum.[2] The probabilistic ERG base-case for R^2 versus R-CHOP ranged from £16,874 to £44,888 per QALY gained (based on 1,000 iterations). For R^2 versus R-CVP, the ICER ranged from £23,135 to £59,810 and for R^2 versus R-mono, it ranged from £18,816 to £26,728.

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost-effectiveness estimates. For the R-CHOP/R-CVP comparisons, using R-CHOP and R-CVP efficacy from van Oers et al. would change the ICER substantially, but not always in the same direction. Alternative assumptions regarding lowered utilities in the PP health states and the time point at which treatment waning start could also change the ICER substantially, dependent on the OS curves chosen. In general, for the R-CHOP/R-CVP comparison it can be said that the model seems instable and results are highly dependent on the assumptions applied, with ICERs ranging between dominant and dominated. For the R-mono comparison, the ICERs are much less volatile, but still ranging between £11,539 and £42,448.

Of note, a full incremental analysis would result in R-CHOP being strictly dominated by definition (being equally effective and more costly than R-CVP), and the relevant ICER would therefore always be R^2 versus R-CVP. For R-mono, a full incremental analysis is not applicable, because costs and QALYs for R^2 are different in this comparison.

The main conclusion of the original ERG report[1] still applies, that is, even though the ERG base-case ICER for R-CHOP was below £20,000, the uncertainty around the cost effectiveness of R^2 is substantial, mainly caused by the possible bias introduced by the indirect treatment comparison, which could not be accounted for in the ERG analyses. In addition, specific to the FL only population analyses presented in the company addendum,[2] the uncertainty around the OS estimates and the lack of clinical validation of these estimates would warrant even more caution in the interpretation of results. The ICER for R-CVP is higher and suffers from the same uncertainty.

Table 6.3: Deterministic ERG base-case for R^2 versus R-CVP comparison: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,746
R-CVP	[redacted]	[redacted]			
Fixing violations (2, use pooled R-CHOP/R-CVP subs Tx insyead of mixed R-chemo)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,841
R-CVP	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,088
R-CVP	[redacted]	[redacted]			
Matter of judgement (4, use gompertz for OS in both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,863
R-CVP	[redacted]	[redacted]			
Matter of judgement (5, use log-logistic for PFS in R2 and Weibull for PFS comparator)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,991
R-CVP	[redacted]	[redacted]			
Matter of judgement (6, use log-logistic for TTNLT both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,844
R-CVP	[redacted]	[redacted]			
ERG base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£27,767
R-CVP	[redacted]	[redacted]			
ERG base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£43,915
R-CVP	[redacted]	[redacted]			

Table 6.4: Deterministic ERG base-case for R^2 versus R-mono comparison: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use weibull for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,958
R-mono	[redacted]	[redacted]			

Table 6.5: Deterministic ERG base-case for R^2 versus R-mono comparison: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use exponential for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,174
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,931
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,816
R-mono	[redacted]	[redacted]			

Table 6.6: Deterministic ERG base-case for R^2 versus R-mono comparison: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use log-normal for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,284
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,951
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,169
R-mono	[redacted]	[redacted]			

Table 6.7: Deterministic ERG base-case for R^2 versus R-mono comparison: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use log-logistic for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,722
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,432
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,775
R-mono	[redacted]	[redacted]			

Table 6.8: Deterministic ERG base-case for R^2 versus R-mono comparison: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use gengamma for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,037
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,504
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,394
R-mono	[redacted]	[redacted]			

Table 6.9: Deterministic ERG base-case for R^2 versus R-mono comparison: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,274
R-mono	[redacted]	[redacted]			
Fixing violations (3, cap utilities at the general population level)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Matter of judgement (4, use gompertz for OS both arms)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,126
R-mono	[redacted]	[redacted]			
Base-case (deterministic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,535
R-mono	[redacted]	[redacted]			
Base-case (probabilistic)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,728
R-mono	[redacted]	[redacted]			

Table 6.10: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-CHOP: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,781
R-CHOP	[redacted]	[redacted]			
Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£23,163
R-CHOP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,630
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,107
R-CHOP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£39,668
R-CHOP	[redacted]	[redacted]			

Use Weibull for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£29,034
R-CVP	[redacted]	[redacted]			
Use exponential For PFS R2 and Weibull for PFS comparator					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,791
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,657
R-CVP	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
R-CVP	[redacted]	[redacted]			
Adverse events for comparator taken from publication					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£30,072
R-CVP	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£31,589
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£40,523
R-CVP	[redacted]	[redacted]			
Source for R-CHOP/R-CVP efficacy from van Oers					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£2,064,117
R-CVP	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£70,564
R-CVP	[redacted]	[redacted]			

Table 6.11: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: Weibull OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,895
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£36,561
R-mono	[redacted]	[redacted]			

Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,066
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,098
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,477
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,281
R-mono	[redacted]	[redacted]			

Table 6.12: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: exponential OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,931
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,564
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,749
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,456
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£20,156
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,370
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,061
R-mono	[redacted]	[redacted]			

Table 6.13: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: log-normal OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,951
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,610
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£26,191
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,776
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,020
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£15,785
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,776
R-mono	[redacted]	[redacted]			

Table 6.14: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: log-logistic OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,432
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£17,078
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,786
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,763

R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£19,580
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,073
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,915
R-mono	[redacted]	[redacted]			

Table 6.15: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: generalized gamma OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,504
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,227
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£24,156
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£11,539
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£16,184
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£14,214
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£13,957
R-mono	[redacted]	[redacted]			

Table 6.16: Deterministic scenario analyses (conditional on ERG base-case) for R^2 versus R-mono: Gompertz OS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					

R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,535
R-mono	[redacted]	[redacted]			
Use Generalised gamma for PFS both arms					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£25,157
R-mono	[redacted]	[redacted]			
Treatment waning effect at 3 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£42,448
R-mono	[redacted]	[redacted]			
Treatment waning effect at 7 years					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,893
R-mono	[redacted]	[redacted]			
Apply same subsequent treatment costs					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£28,918
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Wild et al. (0.62)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£21,341
R-mono	[redacted]	[redacted]			
Alternative utilities for PP states from Pereira et al. (0.45)					
R^2	[redacted]	[redacted]	[redacted]	[redacted]	£18,333
R-mono	[redacted]	[redacted]			

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Monday 7 October 2019** 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.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Based on recent communication that MZL is no longer being considered within the licensing for R^2, all responses to the ERG report have been based on relevance to the FL only population and FL only cost-effectiveness results.

Issue 1 Relevance of a comparison versus R mono in a rituximab refractory population

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 1.1 on page 11, section 3.3 on page 28, section 4.1.5 on page 34, the report states "Secondly, even if the NICE committee accepts splitting the population in rituximab refractory patients and non-rituximab refractory patients, the CS should still have included a comparison with rituximab monotherapy for both populations as specified in the scope."	We propose that this statement is removed	A comparison of R^2 vs R mono in a rituximab-refractory population has not been provided because: (1) it would be clinically illogical to treat with rituximab monotherapy in a population that is deemed refractory to rituximab (2) We have not identified any rituximab monotherapy data in the rituximab refractory population	Not a factual error.

Issue 2 O-B is still a relevant comparator in the rituximab refractory population, despite being on the CDF

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 1.1 on page 11, section 3.3 on page 28, section 4.1.5 on page 34, the report states "Thirdly, the company included O-Benda as a comparator for rituximab refractory patients. However, in the response from NICE to comments on the draft scope, NICE clearly stated that "obinutuzumab in combination	We propose that this statement is amended to acknowledge the availability of O-B as a treatment option in the rituximab refractory population	Though we appreciate NICE's position on CDF drugs, this is the only NICE-recommended option for this patient group and clinical experts stated this is the favoured treatment choice for FL patients refractory to rituximab. Historically bendamustine monotherapy has been used as a treatment in this population, however at present it is not considered a comparator as	Not a factual error.

with bendamustine is only used as part of the Cancer Drugs Fund therefore it is not considered a relevant comparator for disease that is refractory to rituximab." (see NICE Response to comments on draft scope (page 4)). Therefore, we believe that the submission currently does not present any relevant evidence for R-refractory patients."		clinical experts believe O-Benda has largely replaced bendamustine monotherapy. We propose that the ERG acknowledge the relevance of O-Benda as a treatment available to patients in the UK so as a) to not ignore the predominant treatment option for this population in the UK and, b) to not mislead the reader as to the extent to which this treatment is applicable to the decision problem. In addition, the submission provides relevant evidence of the efficacy of R^2 in the rituximab refractory population through the presentation of data from MAGNIFY (which, unlike AUGMENT, does not exclude this patient population).	

Issue 3 R-CHOP and R-CVP patient characteristics

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 4.4, page 51: *"However, looking at the Advisory Board document provided by the company,[19] no such statement is included; therefore, it is not clear how this clinical feedback was obtained….while '[redacted]'.[19] This suggests that R-CHOP and R-CVP are generally considered for different types of patients, making a comparison of the effectiveness of the two drugs problematic."*	We propose this statement is amended to acknowledge that the feedback regarding the similar efficacy of R-CHOP and	The company would like to clarify that this feedback was obtained through post advisory board consultation. The company also feels that this is an over simplified interpretation of the advisory board feedback. Patient age and fitness are only two	Not a factual error.

	R-CVP was obtained post-advisory board.	considerations of multiple factors that contribute to treatment decision making. There is distinct overlap of patients who may be eligible for R-CHOP or R-CVP due to varying interpretations of and thresholds for "fitness" and age across the country, i.e. the same patient could be offered R-CHOP by one physician and R-CVP by another.	
Section 4.4, page 52: *"The one covariate that was consistently related to outcome was age, which suggests that R-CVP will be more often considered for elderly patients and R-CHOP will be more often considered for younger patients; which means that the drugs are generally considered for different populations, making a comparison problematic."* Section 5.2.6, page 72: *"Although the ERG appreciates that R-CHOP and R-CVP data were pooled to obtain a larger sample size, it is still small and the pooling may have introduced additional bias as the KM curves from the HMRN report[65] show a rather consistent difference in favour of R-CHOP, which may be a result of the fact that the target population for R-CHOP is the younger and fitter group, enhancing efficacy"*	It is proposed that a statement is included to clarify that age (and other reported baseline characteristics) are balanced across R-CVP and R-CHOP patients in the HMRN data.	Similar baseline characteristics (in particular age) of the R-CVP and R-CHOP patients in the HMRN data supports pooling of the data. Age is reasonably well balanced across R-CVP treated patients and R-CHOP treated patients in the HMRN data. R-CVP patients [redacted] median age: [redacted] R-CHOP patients [redacted] median age [redacted]. Further, all other baseline characteristics are reasonably balanced across the R-CHOP patients and	Not a factual error.

		R-CVP patients in the HMRN data. This reflects our understanding of variations in clinical practice, i.e., thresholds for age and fitness can vary across the country and between treating physicians.	

Issue 4 Covariates included within the MAIC versus HMRN

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Misreporting of covariates available for the MAIC in the context of the indirect comparison of R^2 with R-CHOP/R-CVP based on HMRN data Section 4.4.2, page 57 Table 4.12 - several baseline characteristics (refractory to last therapy, serum beta-2 microglobulin high, bone marrow involvement, diameter of largest node >6 cm, haemoglobin <12 dL/L, time from last treatment, ECOG performance status, presence of B-symptoms) were commented as "*Not included in the MAIC*"	It is proposed that the comments in Table 4.12 are changed to "*Not collected in or reported by HMRN*". In addition, the accompanying notes in Table 4.12 *"Notes: Adjusted N is the sum of absolute weights…."* is for the comparison to Van Oers and not to HMRN and should be removed. It incorrectly states that % previous rituximab exposure was not matched for vs HMRN. However, this was possible with the HMRN data, providing a key advantage of the comparison to HMRN data over the comparison to Van Oers). The text on pages 57 and 58 will also need updating to reflect these changes.	The characteristics listed as "not included in MAIC" (with the exception of bone marrow involvement) were either not collected by HMRN at the relapsed/refractory baseline and were only available at diagnosis (high LDH, serum beta-2 microglobulin high, diameter of largest node >6 cm, haemoglobin <12 dL/L, ECOG performance status, presence of B-symptoms) or were not included in the report provided by HMRN (refractory to last therapy, time from last treatment) and therefore could not be considered for inclusion in the MAIC. Bone marrow involvement was initially not included in the MAIC but was included at the ERG clarification	Not a factual error.

		stage. It is important to clarify that all available prognostic factors and treatment effect modifiers were considered for inclusion in the MAIC.	

Issue 5 Validation of OS curves for R^2 versus R-CHOP/CVP

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 1.5, page 16 *"The choice of OS curve was mainly based on a previous STA (TA137: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma)."* Section 5.2.6, page 70, 73 *"The company did not explain why the exponential distribution was not used, but stated that AIC/BIC for Weibull suggested a reasonable fit, and Weibull was also used in TA137."* *"For OS, the company's argument for choosing the Weibull distribution over the better fitting exponential distribution was that the AIC/BIC for Weibull suggested*	We propose the statements are changed to: *"The choice of OS curve for R-CHOP/CVP was mainly based on clinical opinion and the use of a previous STA (TA137: Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma)."* *"The company did not explain why the exponential distribution was not used, but stated that AIC/BIC for Weibull suggested a reasonable fit, and Weibull was also used in TA137. In response to clarification question B9 the company also confirmed that post-advisory board clinical validation was sought for the extrapolations of R-CHOP/CVP which confirmed that Weibull was the most appropriate"* *"For OS, the company's argument for choosing the Weibull distribution over the better fitting*	In response to clarification question B9, the company state that: *"The clinical expert thought the bottom 3 curves (exponential, Weibull and log-logistic) were more plausible than the top 3 (generalized gamma, Gompertz, and log-normal). The bottom 3 curves estimated OS at 20 years between 30% - 21%. They felt that any of these could be plausible and suggested to choose the middle curve (Weibull), as the 20-year OS estimate of 27% seems reasonable and wouldn't expect anyone to challenge this."* This clarification should be clear in the statements made by the ERG as this offers an additional justification for the choice of Weibull distribution for the company's base	Not a factual error.

a reasonable fit, and Weibull was also used in TA137[59] on R-mono."	*exponential distribution was that the AIC/BIC for Weibull suggested a reasonable fit, and Weibull was also used in TA137[59] on R-mono. In response to clarification question B9 the company also confirmed that post-advisory board clinical validation was sought for the extrapolations of R-CHOP/CVP which confirmed that Weibull was the most appropriate"*	case.	
Section 1.5, page 17: *"It is not clear whether all assumptions and extrapolations (notably for PFS, OS and TTNLT for patients treated with R-CHOP/R-CVP) were validated by experts."* Section 5.2.12, page 87: *"The advisory board meeting report supported some model approaches and assumptions, but not all: for instance, the model structure including the on- and off-treatment division was not corroborated, and neither was the choice of distributions for R-CHOP/R-CVP OS and PFS."* *"Furthermore, it is not clear whether these extrapolations have been validated by experts as the expert meeting minutes only contained a statement regarding*	We propose the statements be amended to include: *"In response to clarification questions, the company confirmed that OS for R-CHOP/CVP extrapolations were validated in a follow-up consultation with advisors subsequent to the advisory board."* *"The advisory board meeting report supported some model approaches and assumptions, but not all: for instance, the model structure including the on- and off-treatment division was not corroborated, and neither was the choice of distributions for R-CHOP/R-CVP OS and PFS which were subsequently validated in a post-advisory board meeting."* *"The expert meeting minutes contained a statement regarding (the comparison with) R-mono. The company subsequently validated the R-CHOP/CVP OS extrapolations in follow-up consultation with advisors subsequent to the advisory board, confirmed in response to clarification question B9a."*	In response to clarification question B9, the company confirmed that in follow-up consultation with advisors subsequent to the advisory board, extrapolations for R-CHOP/CVP were discussed.	Not a factual error.

(the comparison with) R-mono."			

Issue 6 Subsequent ASCT and R-maintenance

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 1.5 page 16: *"Also, the ERG considered it to be inconsistent that AEs related to subsequent ASCT and R-mono therapy were only taken into account for R-CHOP and R-CVP and so this was corrected for in the ERG base-case."* Section 5.2.7, page 76: *"For patients in the R-CHOP/R-CVP arm undergoing ASCT and R-mono as subsequent therapy, AEs related to these treatments were accounted for and costed in the model. Also, a small utility decrement was applied for these AEs."* Table 5.11, page 89 Section 5.3, page 91: *"AEs related to subsequent treatments not accounted for in R^2 arm (section 5.2.7). The ERG*	The cost for AEs and disutility for ASCT and R-maintenance has been misinterpreted by the ERG and therefore have provided incorrect statements and ICERs relating to their updated base case and resulting sensitivity analysis. AEs relating to ASCT and R-maintenance were only applied to treatments where patients had subsequent ASCT and R-maintenance following induction therapy (pre-progression) and were not applied to ASCT or R-monotherapy received as a subsequent treatment post-progression. Therefore, as no ASCT or R-maintenance therapy was included post R^2 induction therapy, no further AEs or utility decrements were applied in the model for R^2. The revision conducted by the ERG for their base case is incorrect as it accounts only for costs and utility decrements for subsequent ASCT and R-mono following R^2 and not for further ASCT and R-mono received as subsequent therapy to the comparators. The company request the ERG revise any statements regarding this misinterpretation and revise their base case and subsequent	AEs and utility decrements for subsequent therapies are not included in the model for any treatment. Only costs and decrements associated with treatments received pre-progression have been accounted for.	The ERG agrees to this. In the addendum for the FL-only population, the ERG analyses have been updated to no longer include this change.

included costs and utility decrement for AEs related to ASCT and rituximab subsequent treatment in R^2 arm like in the comparator arm." Table 5.12, page 92 Table 5.13, page 93 Table 5.14, page 93 Figure 5.4, page 94 Figure 5.5 & 5.6, page 95 Figure 5.7, page 96 Section 5.4 Tables 6.1, 6.2, 6.4, 6.5 and 6.6	sensitivity analysis accordingly. Removing the additional AE cost and utility decrement from the R^2 arm in the ERG FL only scenario produces the following revised deterministic ICERs (Tables 6.4-6.6, pages 101-103,): R^2 vs R-CHOP: £16,581 R^2 vs R-CVP: £22,742 R^2 vs R-mono: £17,856		
Section 5.2.9, page 84: *"The ERG was unable to find any report of actual incidence of ASCT performed post R^2 in AUGMENT, but would have liked to see a scenario using observed frequencies, as clinical practice may sometimes contrast with protocols and clinical opinion."* Table 5.11, page 90	The company would like to clarify that within AUGMENT, 2 patients received ASCT post R^2. These are already accounted for within the subsequent treatment costs as these were received as 3rd subsequent treatments and not directly after R^2 induction. This was within the response to ERG clarification question B14: *"Two patients in the AUGMENT trial did receive stem-cell transplant subsequently, however this was administered as a 3rd subsequent regimen and thus included within the costs for subsequent treatment assigned to the R^2 arm."* To suggest that excluding consolidation ASCT in R^2 is bias in favor of R^2 is incorrect. Firstly, no patients within AUGMENT received ASCT	To aid clarity of the matter of subsequent ASCT rates applied in the model.	Not a factual error. From the text in the CS it appeared that the 0% was in fact based on AUGMENT protocol and clinical opinion.

	directly post R^2 induction. It was therefore not assumed to be zero, but correctly modeled as zero, and the costs are associated with the efficacy. If ASCT post induction was to be added to the R^2 arm, this would also need to be reflected within the efficacy. Secondly, clinician feedback states that R^2 is not expected to be an induction therapy for ASCT, unlike R-CHOP.		

Issue 7 Subsequent treatments

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 1.5, page 17: *"The ERG questioned the company's choice to include subsequent treatments as a one-off cost, which may not reflect clinical practice as patients may receive more than one subsequent treatment."* Section 5.2.9, page 83: *"Subsequent treatments were included in the model as a one-off cost to those patients entering the PP on treatment health state. The ERG is concerned that this assumption does not reflect clinical practice as patients may receive more than one subsequent treatment and subsequent treatment costs in the economic*	We propose the statements are changed to: *"Subsequent treatments were included in the model as a one-off cost to those patients entering the PP on treatment health state. The company costed for all incidences of subsequent treatments from the data sources."* *"Subsequent treatments were included in the model as a one-off cost to those patients entering the PP on treatment health state. The company costed for all incidences of subsequent treatments from the data sources."*	This is a misunderstanding. The subsequent treatments included in the model account for patients receiving more than one subsequent treatment (i.e. the costs are based on all incidences of subsequent treatments and not just patients first subsequent treatment).	The ERG agrees that costs are based on all incidences of subsequent treatments and thus account for patients receiving more than one subsequent treatment. However, for R^2 these incidences were based on the AUGMENT trial with a much shorter follow-up, and the model thus did not account for subsequent treatment costs in the R2 arm over the whole time span patients remained in the PP on treatment health state. Amended into: 'The ERG questioned the company's choice to include subsequent treatments as a one-off cost to those

model are therefore likely to be underestimated."			patients entering the PP on treatment health state. The company costed for observed incidences of subsequent treatments from the data sources, which for R^2 had a much shorter follow-up than for R-CHOP/R-CVP and therefore may not be reflective of clinical practice.' And 'Subsequent treatments were included in the model as a one-off cost to those patients entering the PP on treatment health state. The company costed for observed incidences of subsequent treatments from the data sources, which for R^2 had a much shorter follow-up than for R-CHOP/R-CVP. The ERG is concerned that because of the limited follow-up in AUGMENT as compared to HMRN, this assumption does not reflect clinical practice and subsequent treatment costs for R^2 in the

			economic model are therefore likely to be underestimated.'

Issue 8 Utility values from Pereira et al.

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 5.2.8, page 76: *"The company stated in the CS that the mean utility value for the PF state was generally consistent with those reported in the three studies selected from the SLR, with the exception of the lower PF utility value of Pereira et al."*	We propose the statement is changed to: *"The company stated in the CS that the mean utility value for the PF state was generally consistent with those reported in the three studies selected from the SLR, with the exception of the lower PF utility value of Pereira et al. which had a smaller sample size (n=21) and did not report the methods or patient characteristics"*	The company request the ERG acknowledge the evaluation undertaken for the Pereira et al study within the submission and clearly present reasons why the company did not consider this to be a relevant study for use within the economic model. This is stated in the company submission, Section B.3.4 (page 177) and Table 43 (page 178) *"Given the smaller sample size, inconsistencies and lack of reporting on methods and patients, Pereira et al. was not considered for use in the economic model."* Additionally, within the response to clarification question B10b: *"It is worth noting that the reasons the utilities from the study of Pereira et al. were not considered appropriate for this submission are detailed in Section B.3.4. of Document B, including:* • *Small sample size (n=21)* • *Lack of reporting on methods and patient characteristics* • *Inconsistencies in the*	Not a factual error.

		difference between PF and PP off-treatment utilities between the Pereira et al. study and AUGMENT, Wild et al. and GADOLIN."	

Issue 9 Re-iterating existence of clinical evidence for R^2 in a rituximab-refractory population

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 1.1 on page 11, section 3.3 on page 28 and section 4.1.5 on page 34, the report states "Therefore, we believe that the submission currently does not present any relevant evidence for R-refractory patients."	We propose rewording to "*Therefore, we believe that the submission currently presents no relevant cost-effectiveness evidence for R-refractory patients."*	Even if the ERG doesn't consider O-B a relevant comparator, relevant evidence for the treatment effect of R^2 in a rituximab-refractory population has been provided in the submission. MAGNIFY is used as a supportive study to provide evidence of R^2 efficacy in the rituximab refractory population, a population excluded from the pivotal trial AUGMENT. MAGNIFY provides supportive data that R^2 has favourable efficacy in both rituximab refractory and non-rituximab refractory populations.	Not a factual error.

Issue 10 Incorrect statements on patient population of AUGMENT

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 1.2 on page 12, and in section 4.6 on page 60, the report states "In conclusion, the CS	We suggest rewording the second sentence to *"All patients in this trial are relapsed/refractory, although rituximab-refractory patients were*	This is an incorrect statement because R-refractory patients were excluded from this study.	Corrected.

included one relevant study, for the comparison of R2 versus R-monotherapy: the AUGMENT trial. All patients in this trial were R-refractory." In Section 4.1.5 on pages 34 and 35, the report states "Methods and results of the indirect comparison for the rituximab refractory population, R^2 versus R-CHOP and R-CVP, are discussed in Section 4.4 of this report. Methods and results of the indirect comparison for the non-rituximab refractory population, R^2 versus O-Benda, will be ignored as this is not a relevant comparator according to NICE.[28]"	*excluded."* The section on pages 34 and 35 should read, *"Methods and results of the indirect comparison for the* ***non- rituximab refractory*** *population, R^2 versus R-CHOP and R-CVP, are discussed in Section 4.4 of this report. Methods and results of the indirect comparison for the* ***rituximab refractory*** *population, R^2 versus O-Benda, will be ignored as this is not a relevant comparator according to NICE.[28]"*	The ERG incorrectly states a population as being refractory when they are in fact non-refractory.	

Issue 11 Extent of expert opinion used in the company submission

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 2.2 on page 21, the report states "The CS describes the following sources that were used in the company's interpretation of the positioning of R2 in the treatment pathway for FL (see Figure 2.1): an advisory board conducted by Celgene in March 2019, involving six UK clinical experts in NHL and two	We suggest rewording to "*The CS describes the following sources that were used in the company's interpretation of the positioning of R2 in the treatment pathway for FL (see Figure 2.1): an advisory board conducted by Celgene in March 2019, involving six UK clinical experts in NHL and two health economics experts, ad-hoc follow up with advisors,…"*	Additional expert opinion was elicited in addition to the advisory board mentioned.	Corrected.

health economics experts,…."			

Issue 12 Positioning of R^2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 2.2 on page 21, the report states "Rituximab in combination with lenalidomide (R2) is an option for both, R-refractory patients and non-R-refractory patients in second-line."	We suggest rewording to *'Rituximab in combination with lenalidomide (R2) is an option for both R-refractory patients and non-R-refractory patients in second-line and beyond'*	The anticipated positioning for R2 is in second line and beyond, not second-line only.	Not a factual error.

Issue 13 Potential misunderstanding on evidence available for indirect comparison

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In section 3.4, on page 28, the report states "The company was not able to find any evidence providing a common comparator linking R2 with any of the comparators of interest (apart from rituximab monotherapy, which was dismissed by the company)."	We suggest rewording to '*The company was not able to find any direct evidence providing a common comparator linking R2 with any of the comparators of interest.'*	The current text is misleading and suggests there was a common comparator link found for R2 and R mono so that an indirect comparison could be conducted between these. Further, analyses versus rituximab monotherapy were provided in response to clarification questions posed by the ERG and therefore the use of the term 'dismissed' is misleading.	Not a factual error.

Issue 14 Minor changes

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 11, Table 3.1 (page 24) and page 27, the report states, "…although the Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated on [redacted], and marketing authorisation is expected in [redacted]."	This should be reworded "…although the Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated on [redacted], and marketing authorisation is expected in [redacted]."	Regulatory dates have shifted slightly.	The text was correct at the time we wrote the report.
Section 4.1.1, page 31 states: "*It is not clear if MEDLINE In-Process, Ahead of Print, and Daily Update were searched*"	The Medline database searched was: MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily The date range this covers is from 1946 to the day it is searched.	Providing clarity	Thank you for this additional information. Text not changed as it was unclear at the time of writing the report.
In section 1.2, on page 12, section 4.2.2 on page 35, and section 4.6 on page 60, the report states "The trial did not include any sites or patients from the UK."	We suggest rewording to "*The trial included 4 sites in the UK, although no patients were recruited by these sites."*	It is not correct that there were no sites in the UK.	Corrected as follows: "The trial did not include any patients from the UK."
Section 4.2.4 (page 42) states: "*Analyses were performed using both FDA and European Medicines Agency (EMA) censoring rules for PFS but only the EMA censoring rule analyses for the ITT population were presented in the CS.*"	We propose the statements are changed to clarify that PFS analyses for the ITT population using FDA censoring rules are presented in the CS Appendix	Correction. Note, these analyses are presented in Appendix N.1: "*Progression-free survival based on data form IRC review using FDA censoring*"	Corrected.
Section 4.4.1, Table 4.9, page 52 Misreporting of variables as "not	High LDH and Bone marrow involvement were not reported in Van Oers	These data were not included in the MAIC due to not being reported	Not a factual error.

included in MAIC" instead of "no data reported in Van Oers"		in the comparator study.	
Incorrect source referenced in Section 4.4.2, Table 4.13, page 59: "Source: CS, Appendix D2, Table 15, page 36."	Should be updated to: "Source: CS, Appendix D3, Table 29, page 57."	The source currently referenced is for the RCHOP comparison vs published literature (Van Oers) and not vs HMRN.	Corrected
Section 4.4.2, page 59: "The comparison of R^2 versus R-CHOP was adjusted for the variables listed in Table 4.13…"	Should read: "The comparison of R^2 versus R-CHOP/R-CVP was adjusted for the variables listed in Table 4.13…"	R-CVP was incorrectly missed out	Corrected
Misunderstanding – Section 5.1.4, page 64. *"They then stated that 'more details of how these evaluations have informed the de novo analysis are discussed in Section B.3.2.'.[1] However, Section B.3.2. of the CS does not contain any information on the use of these evaluations."*	The company would like to clarify that the statement made by the company within the ERG comment was in relation to the SLR evaluations as a whole and not specific to the evaluations that had a UK perspective. Within the 'Model Structures' Section of the company submission (page 126), the company discusses how the evaluations found in the SLR have informed the model structure.	Further clarification to aid the ERG's understanding of the use of the SLR reports.	Thank you for this clarification. Text not changed as it was unclear at the time of writing the report.
Incorrect reference – Section 5.2.3 page 69: *""The AUGMENT trial was used as the source for utilities, subsequent treatments and disease monitoring for R2 as well as the comparators.*"	We propose the statement is changed to: '*The AUGMENT trial was used as the source for utilities for R2 as well as the comparators, and as the source of subsequent treatments for R2*.'	Disease monitoring was not sourced from AUGMENT for any treatment (R2 or the comparators). As described in Section B.3.5 of the company submission, disease monitoring resource use were based on those presented in previous FL NICE submissions and ESMO guidelines.	Corrected as suggested

Misreporting of scenarios – Section 5.2.11 Section 5.2.11, page 86: *"The company conducted several scenario analyses. The results for R2 versus R-CHOP showed ICERs ranging between £8,174 and £14,891 per QALY gained, excluding the scenarios assessing different time horizons."* *"The three most influential scenarios that decreased the ICER were a 0.0% discount rate for QALYs (£8,174), using a log-normal distribution for R2 ToT (£8,312) and using lenalidomide trial RDI (£9,128)."* *""The three most influential scenarios that increased the ICER were a 6.0% discount rate for QALYs (20,636), applying the comparator hazard to R2 arms after three years (£20,471) and using a Gompertz distribution for R2 PFS (£14,891)."*	The ERG has missed a scenario relating to R2 ToT (using the exponential distribution) which has a smaller ICER than quoted (£4,398). This scenario should be grouped within the three most influential scenarios. Please revise these statements to include this scenario. The ICER quoted for using a Gompertz distribution for R2 PFS is in comparison to R-CHOP and should be changed for the correct ICER in comparison to R-CVP (£20,413).	The ERG had missed a scenario regarding R2 ToT which was one of the smallest ICERs reported and mis-reported an ICER.	Thank you for spotting these mistakes. The exponential distribution for R^2 ToT was added to the three most influential scenarios and the ICER quoted for using a Gompertz distribution for R^2 PFS was corrected as suggested.
Misinformation – Section 5.2.11, page 86: *"For R2 versus R-mono, several scenarios exploring different distributions for ToT, OS, PFS and*	The company provided the flexibility within the model to explore several distributions relating to the ToT, OS, PFS and TTNLT for the R2 versus R-mono comparison. The company would like to request the ERG to remove this statement or	Further clarification required from the ERG.	The model allows for these scenarios – although they were not reported by the company in the response to clarification. The ERG

TTNLT were not available."	clarify what is meant by 'not available'.		removed this statement.
Misreporting of ICERs – Section 6, page 102 Table 6.5.	Several scenarios relating to the ERG base case within Table 6.5 have been incorrectly reported. The company request the ERG to revise these results for; - Use of Weibull for PFS, - Exponential PFS for R2 and Weibull for R-CVP - Treatment waning effect at 3 and 7 years - Adverse events from publication.	Corrections required from the ERG.	The ICERs of these particular scenarios were indeed misreported, and have been corrected.
Appendix 1, Table A1.2, page 113: Patient numbers from the overall ITT population, and not the MZL specific population are stated. R^2 (n=178) R-mono (n=180)	This should be updated to: *R^2 (n=31)* *R-mono (n=32)*	Correction required as MZL specific numbers were not reported.	Corrected

Issue 15 AIC/CIC marking

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Section 5.2.9, page 82 *"In the economic model, the company applied ASCT to [redacted] of patients in R-CHOP."*	*"In the economic model, the company applied ASCT to [redacted] of patients in R-CHOP."*	Text contains the results of unpublished HMRN data used to inform the cost-effectiveness. Celgene will inform NICE once scientific publication plans by HMRN have been finalised. However, publication dates are not expected on completion of UK HTA submissions for this indication.	This should indeed have been marked AIC, apologies. Amended as requested.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft technical report

Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma

This document is the draft technical report for this appraisal. It has been prepared by the technical team with input from the lead team and chair of the appraisal committee.

The technical report and stakeholder's responses to it are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the appraisal committee meeting.

The technical report includes:

- topic background based on the company's submission
- a commentary on the evidence received and written statements
- technical judgements on the evidence by the technical team
- reflections on NICE's structured decision-making framework.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

Issue date: November 2019

1. Summary of the draft technical report

1.1 In summary, the technical team considered the following:

Issue 1 The matched adjusted indirect comparison (MAIC) is highly uncertain

Issue 2 The model structure may be inappropriate

Issue 3 Survival extrapolations are highly uncertain

Issue 4 Utility scores appear inflated in comparison to population norms

1.2 The technical team recognised that the following uncertainties would remain in the analyses and could not be resolved:

- The clinical trial evidence in AUGMENT is immature and median overall survival has not been reached

1.3 Taking these aspects into account, the technical team's preferred assumptions result in an incremental cost-effectiveness ratio (ICER) of £30,404 per QALY gained for lenalidomide with rituximab vs. R-CVP, and £14,457 per QALY gained for lenalidomide with rituximab vs, rituximab monotherapy. This estimate does not include the commercial arrangement for obinutuzumab, which is given as a subsequent treatment, because this is confidential and cannot be reported here. Given that only 2.0% and 2.7% of patients (in the lenalidomide with rituximab and rituximab monotherapy arms respectively) receive obinutuzumab in subsequent treatment, including this commercial arrangement would produce ICERs slightly higher than those reported here.

1.4 Based on the modelling assumptions, the intervention is not likely to meet the end-of-life criteria (see table 3).

1.5 The technology is unlikely to be considered innovative (see table 3).

1.6 No equality issues were identified.

2. Topic background

2.1 Disease background: Follicular lymphoma (FL)

- The lymphatic (white blood cell) system is responsible for fighting infection or disease in the body.
- Follicular lymphoma (FL) is an incurable disease of the lymph nodes within the lymphatic system. FL is an indolent (slow growing) 'non-Hodgkin's' lymphoma, meaning that they are typically more common, harder to detect, and diagnosed at more advanced stages than Hodgkin's lymphomas.
- There are 3,579 new indolent non-Hodgkin's lymphoma cases in the UK every year. FL accounts for 18% of these.
- FL is associated with lymph node enlargement (potentially leading to restricted movement, cosmetic disfigurement and pain), bone marrow failure (resulting in cytopenias), and other symptoms (night sweats, fever and weight loss).

2.2 Lenalidomide with rituximab

Anticipated marketing authorisation	[redacted]
Mechanism of action	The combination immunotherapy of lenalidomide and rituximab acts by complementary mechanisms. These include direct tumour destruction in FL, and immune-mediated activities such as activation of natural killer cells, and immune synapse formation resulting in increased antibody-dependent cellular cytotoxicity in vitro.
Administration	The recommended starting dose of lenalidomide is 20 mg, orally once daily on days 1 to 21 of repeated 28-day cycles for up to 12 cycles of treatment. The recommended starting dose of rituximab is 375 mg/m^2 IV every week in cycle 1 (days 1, 8, 15, and 22) and day 1 of every 28-day cycle for cycles 2 through 5.
Price	Lenalidomide is available as a 21-capsule pack. The cost per pack (excluding VAT; company submission) is £4,168.50 (20

	mg). A confidential patient access scheme (PAS) has been arranged with NHS England, with a simple discount in place. The list price of rituximab is £349.25 per 2 x 100 mg vials and £873.15 per 1 x 500 mg vial (MabThera®). The prices of biosimilar rituximab also used in the economic analyses are £314.33 per 2 x 100 mg vials and £785.84 per 1 x 500 mg vial. Assuming the starting dose of 20 mg lenalidomide and the AUGMENT mean patient body surface area of 1.85 m^2, the per 28-day cycle costs are £4,168.50 (list price) or [redacted] (with PAS) for lenalidomide and £4,845.98 (cycle 1) and £1,211.50 (cycles 2–5) for rituximab.

2.3 Patient and professional views

Lymphoma Action:

- The psychological impact of an 'incurable' disease affects carers as well as patients. Some carers worry that their loved one's lymphoma has relapsed whenever they are ill.

NCRI/RCP/RCR/ACP:

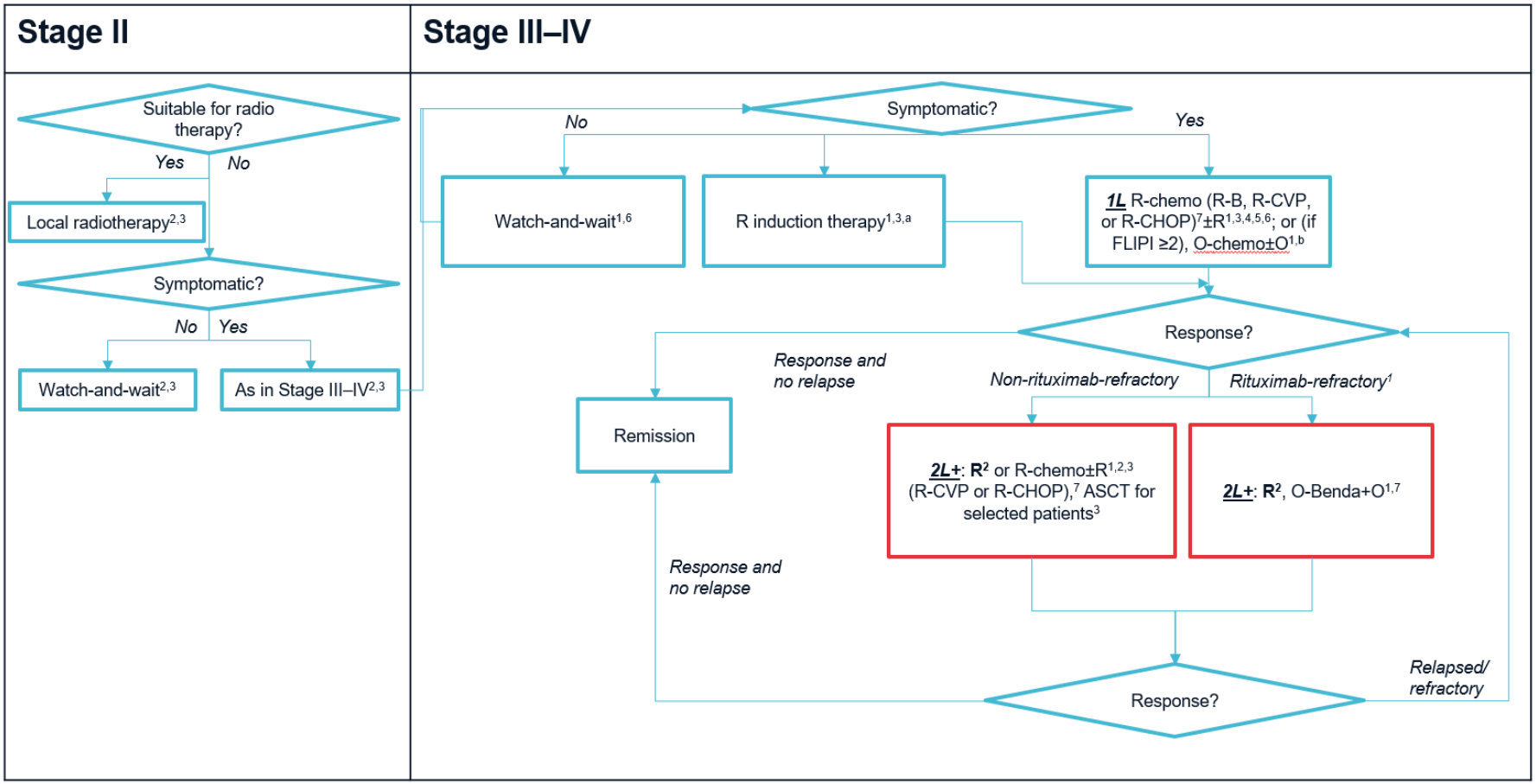
- Lenalidomide plus rituximab represents a major treatment breakthrough for indolent NHL as there are currently no approved targeted therapies other than CD20 antibodies.
- Lenalidomide is suitable for older, less fit patients who would not tolerate further chemoimmunotherapy. This is pertinent given that the median age of patients with newly diagnosed FL is about 60.
- Lenalidomide plus rituximab is generally well tolerated with a toxicity profile that seems fully justified by its clinical efficacy in a setting where effective treatment options are limited.

The Christie NHS Foundation Trust:

- The new technology addresses important areas of unmet need including overcoming treatment resistance in patients with early relapse or heavily pre-treated chemotherapy-refractory disease.
- The NHS is set-up to deliver the new technology with no significant resource implications.

2.4 Treatment pathway

2.4.1 Current clinical pathway for treatment of follicular lymphoma, with proposed positioning of lenalidomide with rituximab (figure 1 of company evidence submission).



2.5 Clinical evidence and results

2.5.1. Rituximab-refractory clinical evidence

AUGMENT: Phase III, multicentre, double-blind, randomized study.

	Follicular lymphoma subgroup	
Outcome	**Lenalidomide with rituximab (N=147)**	**Rituximab monotherapy (N=148)**
Population	All patients with non-rituximab-refractory follicular lymphoma	
Median overall survival	[redacted]	[redacted]
	[redacted]	
Median PFS (months)	[redacted]	[redacted]
6 month PFS rate (95% CI)	[redacted]	[redacted]
1-year PFS rate (95% CI)	[redacted]	[redacted]
2-year PFS rate (95% CI)	[redacted]	[redacted]
PFS Hazard Ratio (95% CI)	[redacted]	

HR, hazard ratio; NE, not estimable; PFS, progression-free survival; OS, overall survival.
[A]: Median survival can only be calculated when cohort survival reaches 50%, which it has not in the AUGMENT study.
[B]: calculated using a Cox proportional hazards model for the available data.

2.5.2. Non-rituximab-refractory clinical evidence

In the non-rituximab-refractory population, R-CHOP/R-CVP (rituximab with different chemotherapy agents) are the most relevant comparators for lenalidomide with rituximab in clinical practice. R-CHOP/R-CVP have been assumed to be equally effective in clinical practice (see issue 1), and have therefore been pooled together as a single comparator. Given equal effectiveness, and that R-CVP is cheaper than R-CHOP, R-CHOP is strictly dominated by R-CVP, and R-CVP has been used as the comparator for lenalidomide with rituximab in cost-effectiveness analyses.

In the absence of direct comparative evidence of lenalidomide with rituximab compared with R-CHOP/R-CVP, a matched-adjusted indirect comparison (MAIC) was used to pair the 'lenalidomide with rituximab' arm of AUGMENT to individuals with R-CHOP/R-CVP from the Haematological Malignancy Research Network (HMRN) registry. The idea of the MAIC is to match the patient characteristics of both arms as closely as possible, in order to create the adequate conditions for an accurate comparison of patient outcomes. If patients are not matched adequately, then outcomes will not be comparable due to the presence of several confounders.

To pair the individuals in the lenalidomide with rituximab and R-CHOP/R-CVP arms, the criteria in the table below was used for matching patient characteristics.

Population	Original AUGMENT FL (n=147)	Matched AUGMENT FL (n=[redacted])	HMRN (R-CHOP/R-CVP) (n=[redacted])
Description	All patients with non-rituximab-refractory follicular lymphoma		
Patient characteristics			
% Prior rituximab	[redacted]	[redacted]	[redacted]
% Age ≥60yrs	[redacted]	[redacted]	[redacted]
% Ann Arbor stage III-IV	[redacted]	[redacted]	[redacted]
% Nodal sites ≤4	[redacted]	[redacted]	[redacted]
% 1 prior lines of therapy	[redacted]	[redacted]	[redacted]
% 2 prior lines of therapy	[redacted]	[redacted]	[redacted]
% Early relapse	[redacted]	[redacted]	[redacted]
% Bone marrow involved	[redacted]	[redacted]	[redacted]

Source: Company addendum, Table 13 and Table 44
FL, follicular lymphoma; N, sample size; HMRN, Haematological Malignancy Research Network

The results of the MAIC (median survival, hazard ratios) are presented in the table below.

Population	Lenalidomide with rituximab (n=[redacted])	R-CHOP/R-CVP (n=[redacted])
Description	All patients with non-rituximab-refractory follicular lymphoma	
Median overall survival (months)	[redacted]	[redacted]
OS Hazard Ratio (95% CI)	[redacted]	
Median PFS (months)	[redacted]	[redacted]
PFS Hazard Ratio (95% CI)	[redacted]	

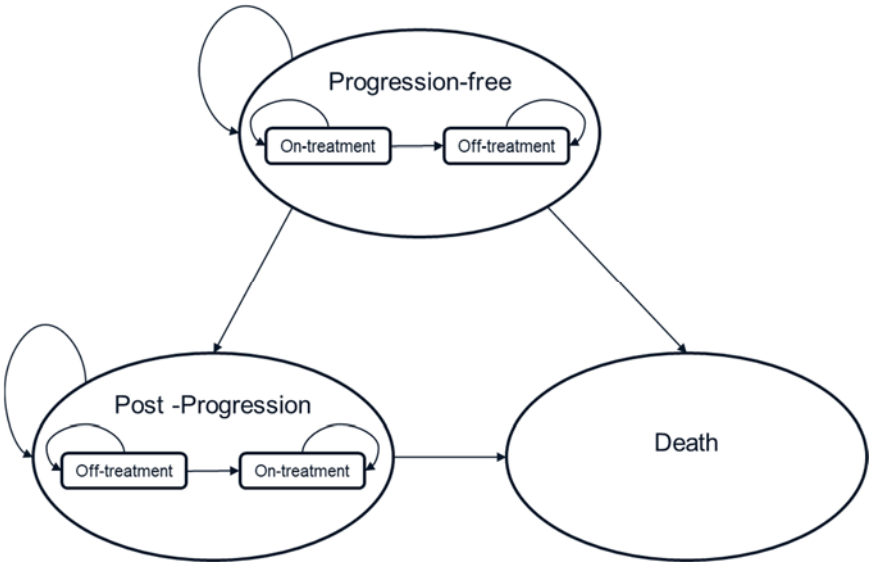
CI, confidence interval; HR, hazard ratio; NE, not estimable; NR, not reported; OS, overall survival; PFS, progression-free survival.
[A]: Median survival can only be calculated when cohort survival reaches 50%, which it has not in the AUGMENT study.
[B]: calculated using a Cox proportional hazards model for the available data (table 13). Confidence interval not reported.
[C]: Taken from the company evidence submission addendum (table 14).

2.6 Model structure

- Partitioned survival model with 3 health states: post-progression, progression-free, and death. 28-day cycle length (with half-cycle corrections).
- Comparator:
  - If non-rituximab-refractory: R-CHOP and R-CVP
  - If rituximab-refractory: rituximab monotherapy
- Clinical effectiveness data from:
  - If non-rituximab-refractory: AUGMENT and HMRN registry
  - If rituximab-refractory: AUGMENT

- Extrapolations based on the most clinically plausible distributions with consideration of best statistical fit. Note: extrapolations are based on a previous patient population assumed to be unchanged (see issue 3).
- ERG base case extrapolations (non-rituximab-refractory):
  - Overall survival: Weibull (both arms)
  - Progression-free survival: log-logistic (lenalidomide with rituximab) and Weibull (R-CHOP/R-CVP)
  - Time to next anti-lymphoma treatment: log-logistic (both arms)
- ERG base case extrapolations (rituximab-refractory):
  - Overall survival: generalised gamma (both arms)
  - Progression-free survival: log-logistic (both arms)
  - Time to next anti-lymphoma treatment: generalised gamma (both arms)
- Treatment assumed to wain after 5 years.
- Health-related quality of life weights derived from AUGMENT EQ-5D-3L data (excludes adverse reaction disutilities).
- 40-year time horizon, discount rates of 3.5% for costs and benefits.

Figure 1: Cost-effectiveness model structure (figure 21 of company evidence submission)



2.7 Key model assumptions

Model input and cross references	Source/assumption	Company justification/reason
Efficacy	Treatment effect is assumed until 5 years.	5-years was chosen based on the previous appraisals in the same patient population (TA137 and TA472).
	R-CHOP and R-CVP are assumed to have the same efficacy.	Clinical experts, consulted by the company, believe that similar outcomes would be expected between the two treatments in the relapsed/refractory setting. HMRN data shows that endpoints for OS and PFS are similar.
	After the maximum follow-up for PFS in AUGMENT, lenalidomide with rituximab	PFS KM curves for lenalidomide with rituximab vs R-CHOP/R-CVP appear to diverge from initiation

Model input and cross references	Source/assumption	Company justification/reason
	is assumed to have the same PFS hazard per cycle as R-CHOP/R-CVP.	but converge as time progresses. The same is seen for the PFS KM curves for lenalidomide with rituximab vs obinutuzumab with bendamustine. This suggests the relative treatment effect is non-constant. To best reflect this, the observed lenalidomide with rituximab data was used to maximum follow-up, beyond which the comparator hazard was applied to extrapolate lenalidomide with rituximab.
	Adverse Events (AEs) were assumed to be 0% for comparators if they were not reported in the literature.	No other evidence to suggest what incidence would be compared to lenalidomide with rituximab, therefore conservative assumption used.
Utilities	Equal health state utilities were assumed between treatment comparisons.	The utility regression model using data from AUGMENT did not show a significant difference in utility between the lenalidomide with rituximab and rituximab-mono treatment arms. Data was also not available to compare all treatments based upon treatment effects. Literature data are used in scenario analysis.
Dosing	All patients receive subcutaneous injection for R-maintenance.	Clinical opinion suggests a large proportion of patients would have R-maintenance subcutaneously. A lower proportion of patients receiving R-maintenance subcutaneously was tested in scenario analysis.
	Pharmacy preparation costs and NHS transport was assumed to apply to all administrations.	These are in line with ERG preferred assumptions in a previous appraisal in the same disease area (TA472).
Subsequent treatments	The distribution of subsequent treatments for the R-refractory population	There is no data for the distribution of treatments for R-refractory patients. Clinician opinion

Model input and cross references	Source/assumption	Company justification/reason
	was assumed to be the same as the non-R-refractory population.	suggests that the distribution of treatments for these two populations in practice are similar.
Key: AE, adverse event; FL, follicular lymphoma; HMRN, Haematological Malignancy Research Network; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; PLD, patient-level data; R-refractory, rituximab refractory; ToT, time on treatment; TTNLT, time to next lymphoma treatment.		

2.8 Overview of how quality-adjusted life years accrue in the model

To calculate average life years in the model, firstly, a parametric curve is fitted to the cohort level survival data available from AUGMENT and HMRN. Life years for lenalidomide with rituximab and comparators are calculated by taking the average (mean) area underneath the overall survival parametric survival curves.

However, to calculate a quality-adjusted life year (QALY) using partitioned survival models, two survival curves are required for average time to event estimation: progression-free survival (time until the disease is classified as progressed), and overall survival (composite of progression-free survival and post-progression survival).

This is because separate health-related quality of life (i.e. utility) values are applied to time in post-progression disease or progression-free disease. Time spent in each health state is weighted by the respective utility value, and then summed together to estimate the number of QALYs that accrue over the time horizon.

3. Key issues for consideration

Issue 1 – The MAIC is highly uncertain

Questions for engagement	1. Is there sufficient evidence to suggest that R-CHOP and R-CVP could be assumed equally efficacious in clinical practice? 2. Can the MAIC be considered appropriate given the extent of covariates that were not included for matching?
Background/description of issue	**The company** has assumed in their model that R-CHOP and R-CVP are equally efficacious. This is due to small patient numbers for R-CHOP and R-CVP individually, and a judgment of similar efficacy between treatments from the company based on Kaplan-Meier plots and Cox proportional hazard models for overall survival, progression-free survival and time to next anti-lymphoma treatment. R-CHOP and R-CVP data have been pooled. The company has omitted several covariates that would have made the sample size of the matched comparison too small, or for reasons not stated. Such covariates include: • Refractory to last therapy status • Follicular Lymphoma International Prognostic Index (FLIPI) risk group. (not collected in HMRN) • FLIPI2+ components: o serum beata-2 microglobulin high o Diameter of largest node o Haemoglobin levels • Time from last treatment • Eastern Cooperative Oncology Group (ECOG) performance status • Presence of B symptoms

	This has prevented all preferable matching criteria from being included in the adjusted matched population, including some covariates described as 'highest priority' by consulted clinical experts for the company (refractory to last therapy status). **The ERG** is concerned that in clinical practice, treatments are generally considered for different populations (R-CVP for older patients and R-CHOP for younger patients) and that the effectiveness of R-CHOP and R-CVP is therefore difficult to compare. The ERG believes that it is unclear why the company has excluded several covariates, and that the trustworthiness of the MAIC is uncertain due to the omission of key covariates. **The clinical experts** do not think it appropriate to combine R-CHOP and R-CVP populations. They believe that, based on clinical trial evidence, R-CHOP has a longer time to treatment failure than R-CVP (despite similar response rates), and that PFS is longer for R-CHOP than R-CVP.
Why this issue is important	Inadequate matching of populations for the MAIC has caused structural uncertainty that cannot be adequately explored in sensitivity analysis. The pooling of patients treated with R-CHOP and R-CVP may be inappropriate. The effect of these issues on the ICER is unclear.
Technical team preliminary judgement and rationale	The technical team agree with the ERG that the results of the MAIC are uncertain. A justification for the removal of each covariate is required from the company to determine if any uncertainty in the MAIC could be resolved.

Issue 2 – The model structure may be inappropriate

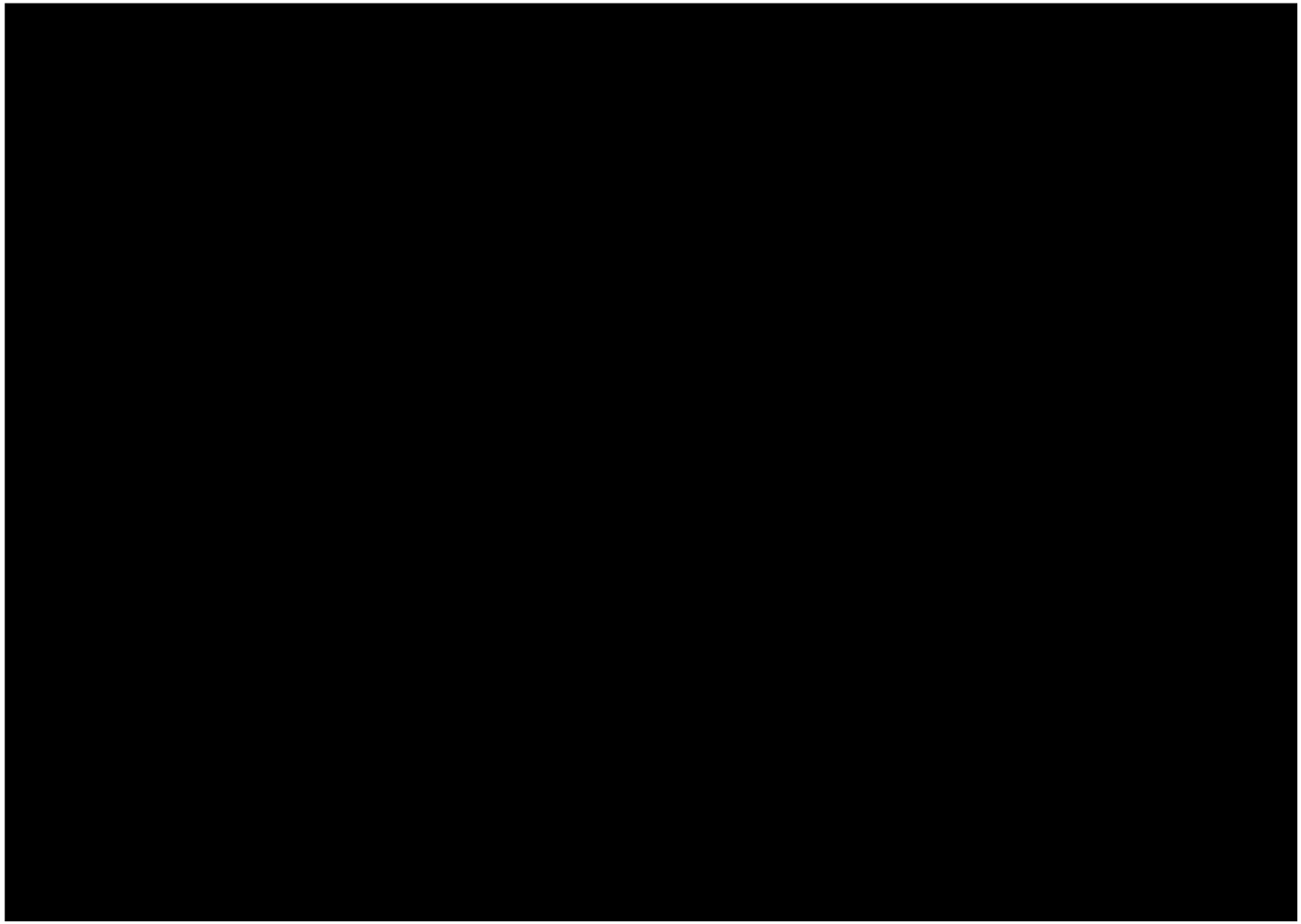
Questions for engagement	3. Would re-analysis of the data, through use of a Markov model, be useful to reduce uncertainty?
Background/description of issue	State transition models incorporate intermediate endpoint data (movements between health states) when predicting what future survival may look like. Partitioned survival models do not, as transition probabilities are not required for the calculation of average survival. The failure to take account of information on intermediate endpoints in the partitioned survival analysis approach may increase the uncertainty associated with the extrapolations generated using this method. Because of this, extrapolations may be inaccurate, and average length of life associated with a treatment may be under or overestimated. It is therefore advised by the NICE decision support unit (technical support document 19, recommendation 11) that if a partitioned survival model is used then a state transition model should also be constructed to cross-examine what future survival may look like. **The ERG** believes that a state transition model should have been used alongside the partitioned survival model to assess the plausibility of extrapolations in the model. This was requested from the company but refused. **The company** argued that because of the weight of the limitations in the state transition model approach, combined with the data available for this decision problem, constructing a state transition model is not applicable for this submission. The company identified the following limitations in the company submission and clarification response to the ERG: • The data required to inform an STM are lacking. The relevant comparators for this submission are not included in the head-to-head study with R2 and so the data available for informing PPS for these treatments are reduced to available published data or alternative sources.

	• Data required for the state-transition model such as post time to next anti-lymphoma treatment survival, and deaths from progression-free off treatment are not specifically reported. • The ERG for TA472 (O-bendamustine for treating FL refractory to rituximab) considered the company's approach to modelling using an STM unreliable due to the discrepancies between the predicted and observed data. They consequently amended the model to a PSM, acknowledging the limitations and immaturity of the data. • The use of unrandomized end points to model transitions such as post-progression survival is highly prone to bias due to the selection effects and informative censoring.
Why this issue is important	The direction of change and magnitude of impact on the ICER is unknown.
Technical team preliminary judgement and rationale	Results of a state transition model would have been valuable for cross-comparison of the model extrapolations, and to reduce structural uncertainty associated with the partitioned survival model.

Issue 3 – Kaplan-Meier extrapolations are highly uncertain

Questions for engagement	4. Upon visual inspection, which parametric distributions appear to be most suitable for the extrapolation of overall survival?
Background/description of issue	The accuracy of the matched cohort that forms the data for R-CHOP/R-CVP survival curves is considered to be highly uncertain by the ERG (see issue 1). Partitioned survival analysis has been performed without pairwise comparison of extrapolations from a state transition model (see issue 2). Both of these issues increase the uncertainty of the survival curve extrapolations. Additionally, clinical experts have not been elicited to choose the most suitable distributions for survival extrapolation using the new FL-only analyses, and previously chosen distributions based on a pooled patient population (patients with follicular lymphoma or marginal zone lymphoma) have been retained. A rationale for assuming that the most suitable distributions for extrapolation are the same, between the FL-only and pooled population, has not been provided. Survival estimates between the original pooled population and the follicular lymphoma-only population appear to be very different. For instance, the Weibull distribution (the distribution used in the company base-case) in the FL+MZL population displays an estimated OS of [redacted]% and [redacted]% at 10 and 20 years respectively. In the FL-only population, these percentages were [redacted]% and [redacted]%. These differences are even more apparent as the FL only population is supposed to have a better prognosis as they are younger and fitter. **The ERG** considers the difference in OS estimates between the FL-only and pooled FL+MZL populations to be implausible, and that since the company did not offer an explanation for this phenomenon, it contributes to the uncertainty around the model results. **The company** has not provided an explanation for the omission of new expert validation for the survival curves. **The clinical experts** note that the overall survival curves for lenalidomide with rituximab, R-CVP and rituximab monotherapy will probably look very similar. This is because FL has a long natural history and continues to respond to subsequent lines of therapy, and most patients get the same treatments over the course of the disease, just in a different order. There might be a difference in survival if these

	treatments were the only treatment given to a patient, but that is not the case for the overall population. The clinical experts believe that generalised gamma is the most appropriate distribution for the R-CVP arm given that median survival for FL is typically greater than 20 years.

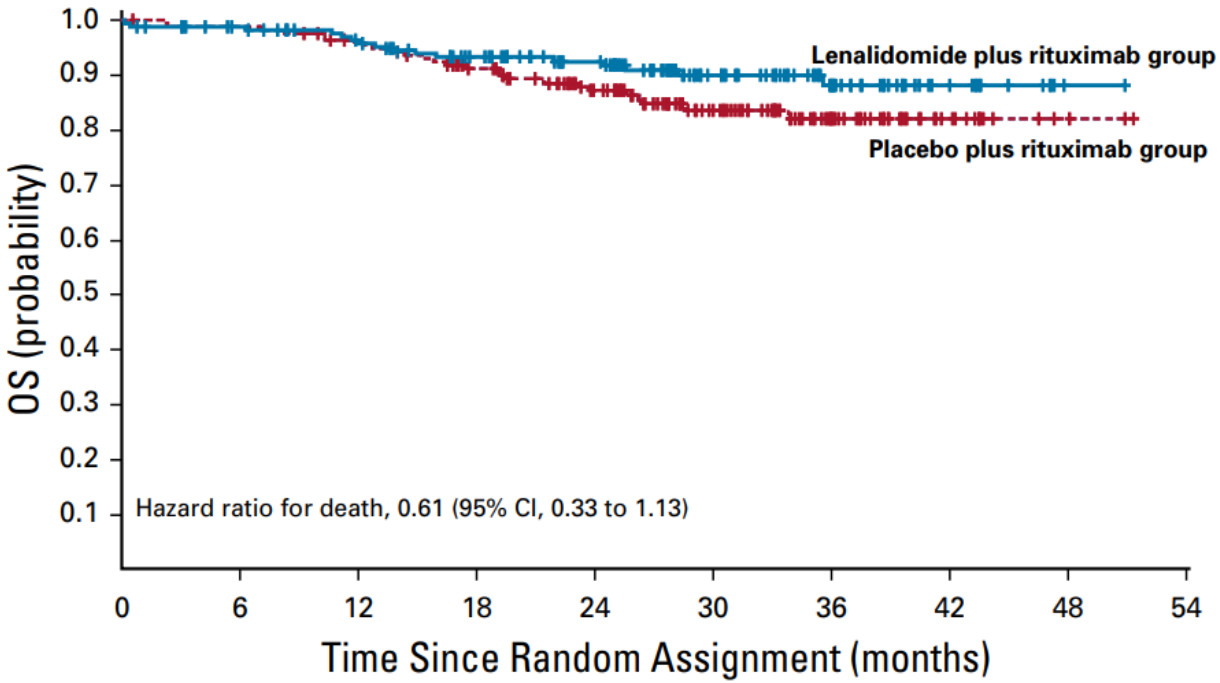
	Lenalidomide with rituximab (FL-only, matched population, overall survival): 

	R-CVP (FL-only, matched population, overall survival): 

	Lenalidomide with rituximab (FL-only, unmatched population, overall survival): 

	R-monotherapy (FL-only, unmatched population, overall survival): 

AUGMENT Kaplan-Meier graph. Lenalidomide with rituximab; rituximab monotherapy (FL and marginal zone lymphoma population, overall survival):



No. at risk:										
Lenalidomide + rituximab	178	167	155	143	122	80	44	15	1	0
Placebo + rituximab	180	176	167	145	116	79	40	14	3	0

Why this issue is important	The choice of extrapolation has a large effect on the ICERs: Using the technical team preferred assumptions, there are a range of deterministic ICERs for lenalidomide with rituximab vs R-CVP depending on the distribution chosen: (£17,312 to £30,404 per QALY gained) Using the technical team preferred assumptions, there are a range of deterministic ICERs for lenalidomide with rituximab vs rituximab-monotherapy depending on the distribution chosen: (£14,457 to £25,625 per QALY gained)
Technical team preliminary judgement and rationale	There is substantial uncertainty surrounding the extrapolations. If clinical opinion is uncertain then the most conservative distributions for extrapolation should be used.

Issue 4 – Utility scores appear inflated in comparison to population norms

Questions for engagement	5. Would you expect health-related quality of life to be higher for the progression-free and post-progression health states compared with the general population (0.8)? 6. Do decrements of -0.026 and -0.056 for utility seem feasible for post-progression compared to progression-free survival?
Background/description of issue	An average individual living in the UK, aged between 55-64, has a health related quality of life of 0.8 (Kind et al., 1999). The utility values for progression-free disease (PFS), post-progression (PP) on treatment and PP off treatment are 0.86, 0.84 and 0.81 respectively in the AUGMENT study. The utility difference between the PF health state and the PP off treatment and PP on treatment health states were -0.026 and -0.056 respectively in the lenalidomide with rituximab versus R-CHOP/R-CVP comparison and respectively -0.026 and -0.055 in the lenalidomide with rituximab versus rituximab-mono comparison. This seems modest given the difference in utility value between these health states reported in the literature, which show differences up to -0.27. **The ERG** believes that utility scores higher than those of the general population seem quite unlikely in patients with treated FL. The ERG has capped the utility values in its base-case analysis to match those of the general population. The ERG also judges that a larger utility difference between PF and PP health states may be more plausible and explored this in a scenario analysis using utility values of Wild et al. (0.62) and Pereira et al. (0.45) for PP health states. **The company** has provided scenario analyses using utility values for health states previously used in technical appraisals (Wild et al., Pereira et al.). The results of these scenario analyses are reported below.
Why this issue is important	Using the technical team preferred assumptions, utility values from Wild et al. result in an ICER of £44,760 and £14,220 per QALY gained for lenalidomide with rituximab vs R-CVP and lenalidomide with rituximab vs rituximab monotherapy respectively.

	Using the technical team preferred assumptions, utility values from Pereira et al. result in an ICER of £79,588 and £13,957 per QALY gained for lenalidomide with rituximab vs R-CVP and lenalidomide with rituximab vs rituximab monotherapy respectively.
Technical team preliminary judgement and rationale	The utility values for progression-free and post-progressed follicular lymphoma should not exceed the utility values of the general population.

4. Issues for information

Tables 1 to 4 are provided to stakeholders for information only and not included in the technical report comments table provided.

Table 1: Technical team preferred assumptions and impact on the cost-effectiveness estimate (lenalidomide with rituximab, R-CVP, FL-only population)

Alteration	Technical team rationale	ICER	Change from base case
Company base case	–	£23,746	–
Subsequent treatment rates are for R-CVP/R-CHOP	Subsequent treatments are either R-CHOP or R-CVP rather than rituximab-chemo. Rituximab-chemo is likely to be an overestimate.	£24,841	+£1,095
Cap utilities at general population	Health state utility values should not exceed those of the general population	£26,088	+£2,342
For both OS arms - Weibull distribution for extrapolation	Technical team matter of judgment. Most conservative assumption given current lack of clinical opinion. Also used by the company, hence no change from the base case.	£23,746	–
For PFS arms - Log-logistic and Weibull distributions for extrapolations	ERG matter of judgement. Log-logistic is used for lenalidomide with rituximab as the hazard appears to be non-constant from the log-cumulative hazard plot. Weibull is used as the comparator based on best statistical fit.	£27,991	+£4,245

Alteration	Technical team rationale	ICER	Change from base case
For both time to next anti-lymphoma treatment arms - Log logistic distributions for extrapolations	ERG matter of judgement. Log-normal curves appear to be suboptimal	£23,844	+£98
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate	**–**	**£30,404**	**+£6,658**

Note: The ERG considers all ICERs to be highly uncertain by due to the underlying uncertainty in the MAIC.

Table 2: Technical team preferred assumptions and impact on the cost-effectiveness estimate (lenalidomide with rituximab, rituximab monotherapy, FL-only population)

Alteration	Technical team rationale	ICER*	Change from base case
Company base case	–	£20,274	–
Cap utilities at general population	Health state utility values should not exceed those of the general population	£21,378	+£1,104
For both OS arms - generalised gamma for extrapolation	Best statistical fit of the lenalidomide with rituximab arm	£13,993	-£6,281
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate	–	**£14,457**	**-£5,817**

*Including lenalidomide PAS

Table 3: Other issues for information

Issue	Comments
Comparators	The company's rituximab refractory comparator (obinutuzumab with bendamustine) is out of NICE scope since it is recommended for use within the Cancer Drugs Fund (CDF) as an option for treating adults with follicular lymphoma. Because of its position in the CDF however it cannot be considered a relevant comparator for lenalidomide with rituximab. Rituximab monotherapy has therefore been used as the comparator for R-refractory individuals.
Population	The current indication is proposed to be follicular lymphoma only.
Cancer Drugs Fund	The company has not proposed a case for lenalidomide with rituximab being considered for funding through the Cancer Drugs Fund. The uncertainty in the analyses arises primarily from the MAIC, which uses relatively small patient numbers available in the HMRN to match the AUGMENT FL-only population. The technical team note that there is the plausible potential for lenalidomide to be a cost-effective use of NHS resources.
Innovation	The company considers the drug to be innovative. However, the technical team considers that all relevant benefits associated with the drug are adequately captured in the model.
Eno of life criteria	The drug does not meet the end of life criteria. Patients in the AUGMENT trial do not have a life expectancy of less than 24 months.

Authors

Stephen O'Brien

Appraisal committee chair

Joel Russell

Technical lead

Sally Doss

Technical adviser

Jasdeep Hayre

Associate director

With input from the lead team:

John Hampson

Clinical lead

Prithwiraj Das

Cost Lead

Stella O'Brien

Lay lead

NICE National Institute for Health and Care Excellence

Technical engagement response form

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm** on **Monday 16 December 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	Celgene
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	none

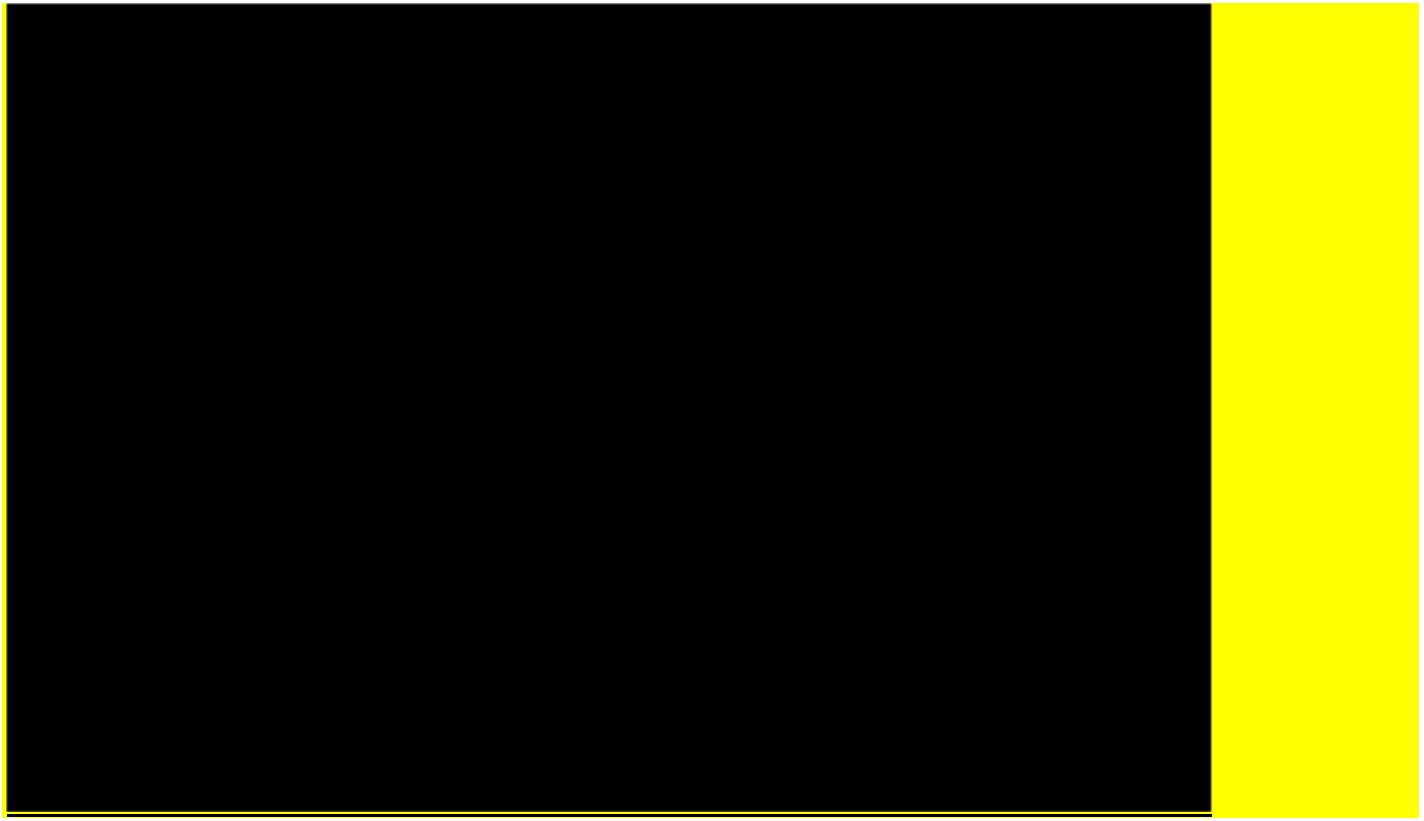
Questions for engagement

Issue 1: The MAIC is highly uncertain	
1. Is there sufficient evidence to suggest that R-CHOP and R-CVP could be assumed equally efficacious in clinical practice?	There is an absence of published evidence comparing R-CHOP and R-CVP in relapsed/refractory FL. While there is comparative data published for FL patients treated with R-CHOP and R-CVP in 1st line, interpretation is complicated by study design and evolving practice. For example, the FOLL-05 study included a comparison of R-CHOP and R-CVP in 1st line FL, reporting superior outcomes for R-CHOP.[1] However, contrary to current practice in the UK, patients were not permitted R-maintenance (which was in fact considered treatment failure for this study). By contrast, recently published retrospective data from Korea demonstrated no significant difference in outcomes between treatment with R-CHOP or R-CVP with all patients receiving R-maintenance.[2] Caution is anyway required in translating 1st line observations to the relapsed/refractory setting. Due to this absence, the assumption of similar efficacy for R-CHOP and R-CVP in relapsed/refractory FL was therefore not based on published literature but derived from the real-world evidence generated by HMRN, and UK clinical opinion, which suggest that it is reasonable to assume similar efficacy for R-CHOP and R-CVP in relapsed/refractory FL.[3] In response to ERG clarification question A14, KM plots of OS, PFS and TTNLT were presented, along with the Cox proportional hazard (PH) model outputs for the pooled R-CHOP and R-CVP data when including treatment, age, prior lines of therapy, early relapse, stage, nodal sites and prior rituximab as covariates. The

KM plots for OS, PFS and TTNLT are overlapping (Figure 1 - Figure **3**), and the results of the Cox PH model (Table 1 - Table **3**) suggests that treatment is not a statistically significant covariate for either endpoint.

Additionally, clinical expert opinion, gained post ad-board meeting and during a clinical validation exercise performed to support this response to the technical report, has endorsed the clinical acceptability of assuming similar efficacy for R-CHOP and R-CVP in treating relapsed/refractory FL.

Figure 1: Kaplan-Meier plot of overall survival for R-CHOP and R-CVP



Key: OS, overall survival; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Table 1: Cox PH Model outputs for overall survival with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
Prior lines of therapy	[redacted]	[redacted]	[redacted]
Early relapse from diagnosis	[redacted]	[redacted]	[redacted]
Stage	[redacted]	[redacted]	[redacted]
Nodal sites	[redacted]	[redacted]	[redacted]
Prior rituximab	[redacted]	[redacted]	[redacted]

Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Figure 2: Kaplan-Meier plot of progression-free survival for R-CHOP and R-CVP

Key: PFS, progression-free survival; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Table 2: Cox PH Model outputs for progression-free survival with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment	[redacted]	[redacted]	[redacted]
Age	[redacted]	[redacted]	[redacted]
Prior lines of therapy	[redacted]	[redacted]	[redacted]

Early relapse from diagnosis	[redacted]	[redacted]	[redacted]
Stage	[redacted]	[redacted]	[redacted]
Nodal sites	[redacted]	[redacted]	[redacted]
Prior rituximab	[redacted]	[redacted]	[redacted]

Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.

Figure 3: Kaplan-Meier plot of time to next anti-lymphoma treatment for R-CHOP and R-CVP

Key: R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone; TTNLT, time-to-next antilymphoma therapy.

Table 3: Cox PH Model outputs for time to next antilymphoma treatment or death with pooled R-CHOP R-CVP data

Variable	Coefficient	Standard error	P-value
Treatment			
Age			

Prior lines of therapy	[redacted]	[redacted]	[redacted]	
Early relapse from diagnosis	[redacted]	[redacted]	[redacted]	
Stage	[redacted]	[redacted]	[redacted]	
Nodal sites	[redacted]	[redacted]	[redacted]	
Prior rituximab	[redacted]	[redacted]	[redacted]	
Key: PH, proportional hazards; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CVP, rituximab plus cyclophosphamide vincristine prednisolone.				

2. Can the MAIC be considered appropriate given the extent of covariates that were not included for matching?

Celgene acknowledges that there were several baseline characteristics that were identified by clinicians as being potentially prognostic variables/effect modifiers that were not adjusted for in the original MAIC analyses based on the availability of HMRN data. In particular, the original analyses did not adjust for refractory to last therapy and FLIPI risk group (low vs. intermediate vs. high) which were identified as highest priority covariates by UK clinical experts. FLIPI risk group was not included in the MAIC analyses as the high LDH component was not collected by HMRN at baseline, however, the other three FLIPI components (age, Ann Arbor stage and nodal sites) were adjusted for in the MAIC.

Refractory to last therapy was not included in the original MAIC analyses as it was not reported within the HMRN report. However, post the technical engagement discussion, Celgene clarified with HMRN whether it was possible to collect refractory to last therapy and confirmed the availability of all covariates listed as prognostic and/or treatment effect modifiers. HMRN were able to provide the percentage of subjects who were refractory to last therapy and confirmed all other variables not included in the analyses were either not collected or not collected at the relapsed/refractory baseline (only at diagnosis). Table 4 confirms the availability of each variable.

Table 4: Variables considered for the MAIC and availability

Characteristic identified as prognostic and/or a treatment effect modifier	Covariates included in the MAIC based on HMRN	Reasons for exclusion
*Previous exposure to rituximab	✓	

	*Age (mean, or median if mean no reported, or % >60 years if neither reported)	✓	
	*Refractory to last therapy	✓[a]	Not included in the HMRN report, however, in response to the Technical Engagement step, Celgene have accessed "refractory to last therapy" data and have re-run the MAIC
	*Prior lines of therapy 1 vs. 2 vs. >2 (or mean/median if categories not reported)	✓	
	*FLIPI risk group (low vs. intermediate vs. high)	×	Not included given LDH is only collected at diagnosis by HMRN. The other three FLIPI components (age, Ann Arbor stage and nodal sites) were adjusted for.
	Ann Arbor Stage (III-IV)	✓	
	Nodal sites (>4)	✓	
	High LDH	×	Only collected at diagnosis by HMRN (not at baseline)
	Serum beta-2 microglobulin high	×	Only collected at diagnosis by HMRN (not at baseline)
	Bone marrow involvement	✓[a]	ERG clarification response included bone marrow involvement as scenario analysis but showed little difference to outcomes. This is now included in the MAIC.
	Diameter of largest node >6 cm	×	Diameter of nodes not recorded by HMRN
	Haemoglobin <12 dL/L	×	Only collected at diagnosis by HMRN (not at baseline)
	FLIPI2 risk group (low vs. intermediate vs. high)	×	Bone marrow involvement is the only FLIPI2 component collected at baseline, thus the FLIPI2 risk group cannot be determined at the relapsed/refractory baseline
	Time from last treatment	×	Time from last treatment is captured through matching for high risk early relapse/POD24 (a characteristic that has been included in the MAIC)

POD24	✓	
ECOG performance status (0–1 vs. 2+)	×	Only collected at diagnosis by HMRN (not at baseline)
Presence of B-symptoms	×	Only collected at diagnosis by HMRN (not at baseline)
Notes: * Considered high priority by UK clinicians [a]Covariates that were not matched for in the original submission but have been for the updated base-case		

The MAIC analyses were re-run including the percentage of patients who were refractory to last therapy. There were a similar percentage of subjects in AUGMENT (FL patients only) and HMRN who were refractory to last therapy at baseline (17.7% vs [redacted]). Table 5 below presents the results when including refractory to last therapy within the MAIC which shows very little difference in the hazard ratios (HR).

Table 5: MAIC results when including refractory to last therapy

R2 (AUGMENT) versus R-CHOP/CVP (HMRN)	HR and 95% CI from MAIC excluding refractory to last therapy[a] (ESS = 58.88)	HR and 95% CI from MAIC including refractory to last therapy[b] (ESS = 58.91)
OS	[redacted]	[redacted]
PFS	[redacted]	[redacted]
TTNLT	[redacted]	[redacted]
Note: [a]Variables included in the adjustment were: age, prior lines of therapy, prior rituximab therapy, POD 24, Ann Arbor stage, number of nodal sites and bone marrow involvement. [b]Variables included in the adjustment were: age, prior lines of therapy, prior rituximab therapy, POD 24, Ann Arbor stage, number of nodal sites, bone marrow involvement and refractory to last therapy.		

These updated MAIC results are included in the company's revised base-case cost-effectiveness results (presented in Appendix 1). To note, the majority of the high priority covariates have been matched for (minus FLIPI risk group, for which 3 out of 4 components have been matched).

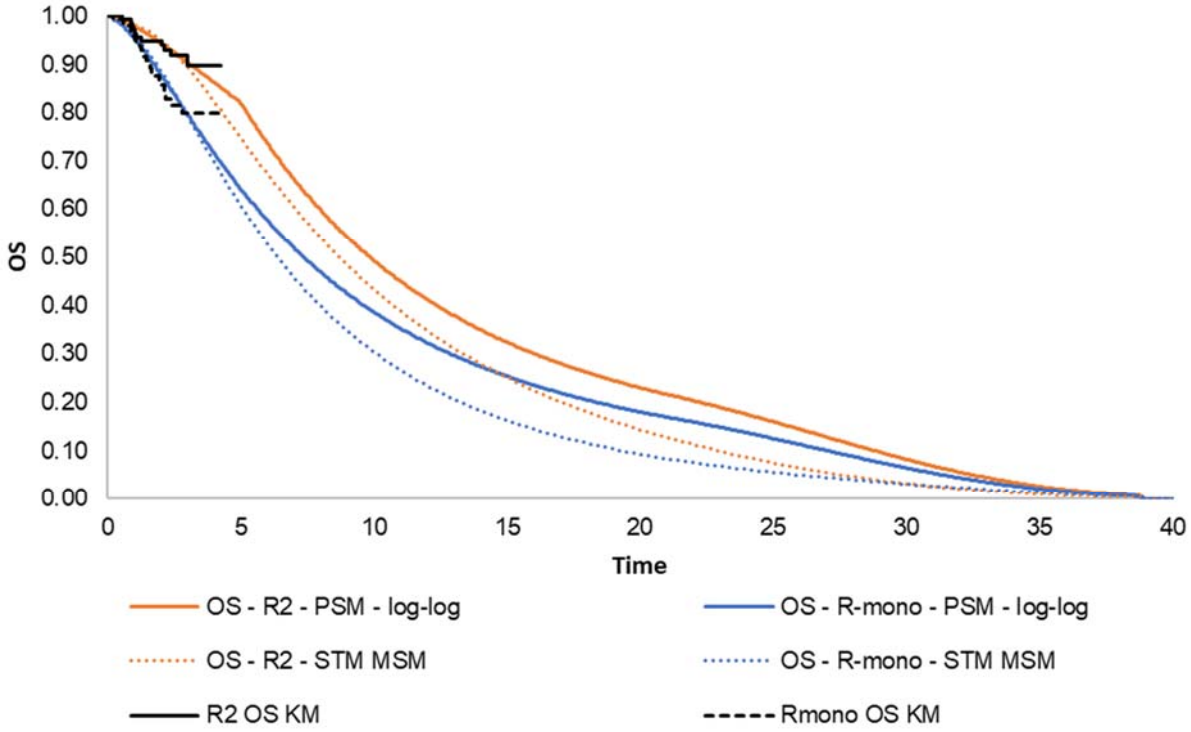
Issue 2: The model structure may be inappropriate

3. Would re-analysis of the data, through use of a Markov model, be useful to reduce uncertainty?	As discussed in response to clarification question B2, the company had multiple concerns with providing a state-transition model due to some of the general limitations of a state-transition approach and concerns specific to the method's suitability for this appraisal (in particular, the application of the method for the comparison versus R-CHOP/CVP using HMRN data). These concerns were raised during the clarification stage, in discussion with the ERG, however, the ERG suggested that the state-transition model should still be provided for the within trial comparison (vs R-monotherapy). Therefore, Celgene have agreed to provide a state-transition model in order to assist in the validation of the company's base case extrapolations using the AUGMENT data for R^2 versus R-mono. Details of the state-transition model methods have been provided separately in an addendum as requested by NICE. Incremental cost-effectiveness results from the state-transition model have been compared to the results from the partitioned survival model (using the three health-state structure) and demonstrate that the outcomes are similar between the two modelling approaches (Table 6). Figure 4 presents the overall survival estimates from the state-transition model in comparison to the partitioned survival model extrapolations using the log-logistic distribution. The state-transition model predicts slightly more pessimistic results in comparison to the partitioned survival model but the difference between the two treatment arms remains consistent between the two approaches. **Table 6: Results of the PSM versus STM**

Technology	Incremental			ICER (£/QALYs)
	Costs	LYG	QALYs	
PSM (3HS)	[redacted]	[redacted]	[redacted]	£17,300
STM	[redacted]	[redacted]	[redacted]	£16,160

Key: HS, health state; ICER, incremental cost-effectiveness ratio; LYG, life years gained; mono, monotherapy; PSM, partitioned survival model; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab; STM, state-transition model.

Figure 4: Overall survival: PSM vs STM



Key: KM, Kaplan-Meier; PSM, partitioned survival model; STM, state-transition model.

In conclusion, the incremental OS remains consistent whether modelled using the state-transition model or the partitioned survival model, and the associated change in the ICER is immaterial.

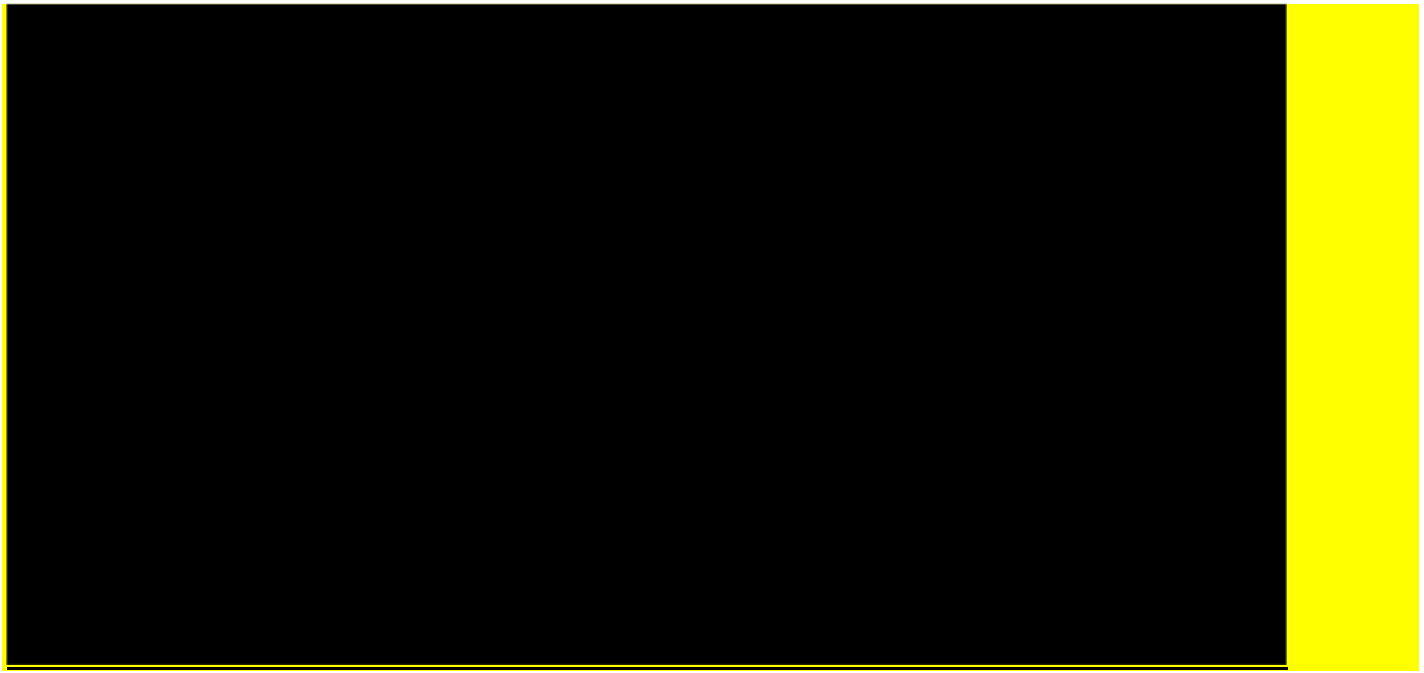
Issue 3: Kaplan-Meier extrapolations are highly uncertain

4. Upon visual inspection, which parametric distributions appear to be most suitable for extrapolation?	**OVERALL SURVIVAL** Celgene acknowledge the uncertainty with the overall survival extrapolations, as is common with most economic evaluations where extrapolation is required. Since the technical engagement call, Celgene have consulted additional clinicians to validate the extrapolations. Clinicians were asked to comment on the extrapolated OS outcomes for the pooled R-CHOP/CVP patients from HMRN and the R-monotherapy patients from AUGMENT, based on survival estimates at years 5, 10 and 20. Clinicians were then asked to comment on the expected difference for R^2 vs R-CHOP/RCVP or R-monotherapy at each of these timepoints. ***R^2 vs R-CHOP/CVP*** **R-CHOP/CVP**: The clinician stated that the Weibull and exponential distributions for R-CHOP/CVP were most plausible based on associated 20-year survival estimates ([redacted] and [redacted] respectively). These choices are further validated by comments made by the clinical expert during the technical engagement call, that 46% survival at 10-years could be considered a reasonable estimate for this population (the 10-year survival estimates for the Weibull and exponential distributions are [redacted] and [redacted] respectively). Based on this feedback, Celgene consider both exponential and Weibull distributions to be the most clinically appropriate to estimate OS of R-CHOP/CVP. **R^2**: Clinical opinion states that the treatment effect of R^2 versus RCHOP/RCVP would start to reduce at about 5-10 years, and over the long-term there is expectation that the survival curves will begin to merge closer together. The model already applies a treatment waning effect at 5 years, after which the hazard of death for R^2 patients is the same as the hazard from the comparator arm. Given the lack of information to

suggest otherwise, and as suggested in NICE TSD 14[4] (for situations in which treatment arms are modelled independently), the same distribution has been used for R^2 as for the comparator.

Figure 5 and Figure 6 present OS for R^2 vs R-CHOP/CVP using the exponential and Weibull distribution, respectively. These curves include treatment waning applied to R^2 at 5 years and both curves being adjusted for general population mortality.

Figure 5: Final OS curves for R^2 versus R-CHOP/CVP – Exponential distribution



Key: KM, Kaplan-Meier, OS, overall survival.

Figure 6: Final OS curves for R^2 versus R-CHOP/CVP – Weibull distribution

Key: KM, Kaplan-Meier, OS, overall survival.

Conclusion: In conclusion, the ERG and the NICE technical team have chosen the Weibull distribution as the most appropriate for OS extrapolations. Further UK clinical validation is supportive of the Weibull, and in addition, states the exponential as another plausible distribution for OS extrapolation. Therefore, the company advocate for the use of the Weibull or the exponential distributions for OS extrapolations in the R^2 vs RCHOP/RCVP comparison.

R^2 vs R-monotherapy

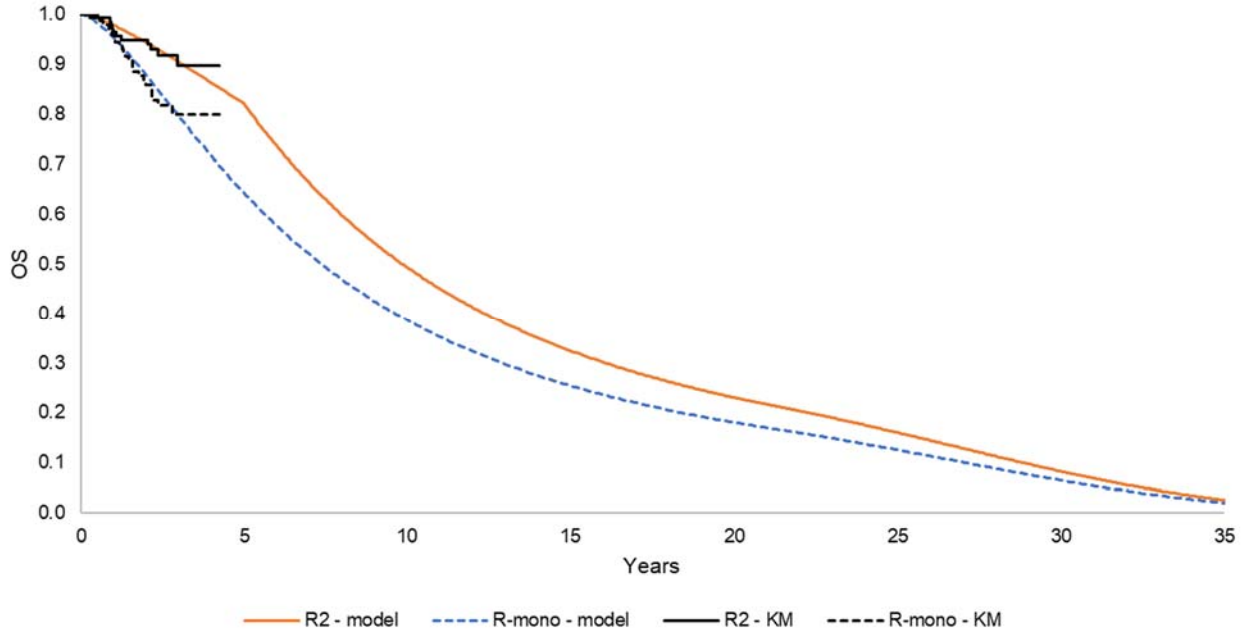
R-monotherapy: For the R-monotherapy distributions, the log-logistic was considered the most clinically plausible based on the 20-year estimate (18%). Generalised gamma was considered too optimistic

estimating survival at 35% at 20-years. Based on this feedback, the log-logistic has been selected as the most clinically plausible curve.

R^2: The same distribution has been chosen for R^2 as for the comparator (R-monotherapy), for similar reasons to those stated above.

Figure 7 presents the final OS curves using the log-logistic distribution (these curves include treatment waning applied to R^2 at 5 years and both curves being adjusted for general population mortality).

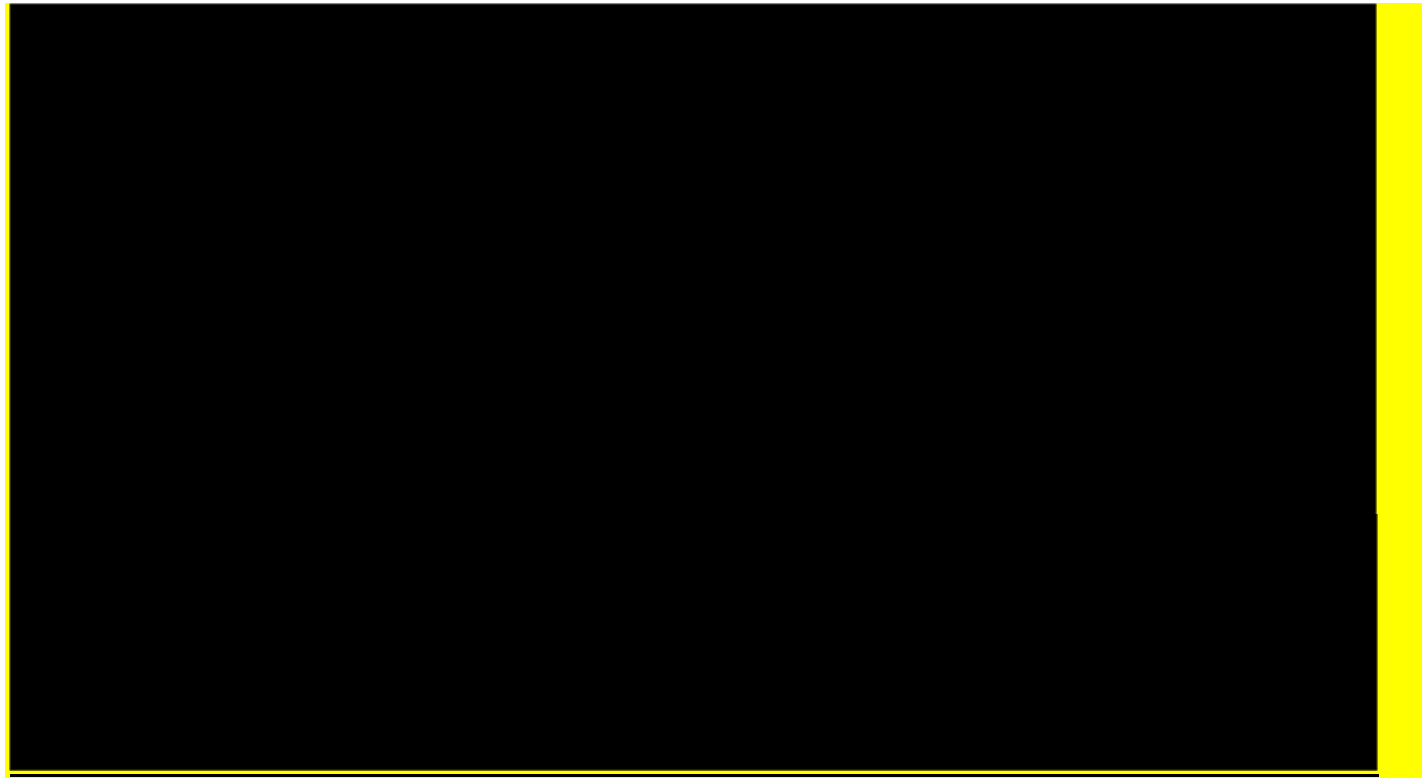
Figure 7: Final OS curves for R^2 versus R-mono – log-logistic distribution



Key: KM, Kaplan-Meier, OS, overall survival.

Conclusion: The ERG and NICE technical team have chosen the generalised gamma distributions for OS extrapolations for R^2 and R-monotherapy. However, given expectation that this overestimates survival, for

	both arms, Celgene advocate for the use of the log-logistic distribution for R^2 and R-monotherapy (supported by UK clinical validation). **PROGRESSION-FREE SURVIVAL** For the comparison of R^2 and R-CHOP/CVP the ERG and NICE technical team chose the log-logistic distribution for R^2 and the Weibull distribution for R-CHOP/CVP to estimate PFS, based on AIC and BIC. Celgene have concerns with these choices because the PFS curves cross at 3 years, meaning patients who receive R^2 end up with worse PFS than those receiving R-CHOP/CVP, over time (see Figure 8). Clinical opinion regarding the relative treatment effect of R^2 vs R-CHOP/CVP concluded that "the curves were more likely to merge than cross" and "crossing was implausible" (a view also expressed by the clinical expert on the technical engagement call). Given that all the parametric distributions for R^2 cross the Weibull distribution for R-CHOP/CVP by 4 years (See Figure 9), none of these options are deemed appropriate for estimating PFS for R^2. This issue occurs because the R^2 parametric extrapolation is informed by insufficient follow-up to accurately capture long-term trends in the PFS hazard. In relation, the follow-up for R^2 from AUGMENT is 3.9 years in comparison to 11.6 years for R-CHOP/R-CVP from HMRN. The HMRN data therefore provides an additional 7.7 years of follow-up which is material evidence on long-term PFS for a relapsed/refractory follicular lymphoma patient population. To address this issue, the company base-case used the KM data from AUGMENT for R^2 until the maximum follow-up (46.7 months (3.9 years)), beyond which the comparator hazard was applied to extrapolate (referred to as "KM + comparator hazard approach"). The ERG was concerned that this approach overestimates PFS for R^2 because the associated extrapolation predicts greater PFS for R^2 relative to the standard parametric extrapolations (see figure 5.2 in ERG report). However, this concern is invalid because it relies on comparisons with parametric extrapolations for R^2, which have been shown to contradict findings of the clinical validation (due to the curves crossing - see above). In contrast, the KM + comparator hazard approach utilises the observed data from AUGMENT followed by an extrapolation approach which assumes equivalent hazards in both arms based on the longer follow-up from the HMRN dataset (additional 7.7 years vs. AUGMENT) and yields extrapolations which align with clinical expert opinion (i.e. over the long-term, PFS will begin to converge).

	In conclusion, Celgene does not believe the parametric distributions selected by the ERG and the NICE Technical team for R^2 are valid for decision making and advocate the KM + comparator hazard approach for the reasons described above. **Figure 8: Final PFS curves for R^2 versus R-CHOP/CVP – log-logistic distribution (R^2), Weibull distribution (R-CHOP/CVP)** 

Key: KM, Kaplan-Meier, PFS, progress-free survival.

Figure 9: Shows PFS parametric curves for R2 against the ERG preferred Weibull parametric curve (red) for R-CVP

Key: KM, Kaplan-Meier, PFS, progress-free survival.

An additional concern raised by the ERG was the timepoint from which extrapolation began (the base case used the maximum follow up from AUGMENT, 3.9 years), for which they noted that the number of patients approaches zero near the tail. Celgene acknowledge that the time point from which to start extrapolations is an assumption, so have provided scenarios with 2 additional timepoints. 700 days, was chosen based on

	the numbers at risk for R^2 being between 25-30 which is considered a reasonable sample size[5] (see Figure 10) and another time point was chosen in between 700 days and the end of the KM (2.9 years). Altering the timepoint had minimal impact on the ICER and doesn't translate to decision uncertainty. Using the Weibull distribution for OS and the company's other preferred assumptions (see Appendix 1), the ICER goes from £26,444 using the time point at the end of the KM to £26,039 using 700 days and £26,987 at 2.9 years. **Figure 10: PFS KM for R^2 versus R-CHOP/CVP**

	 **Key:** PFS, progress-free survival.
Issue 4 – Utility scores appear inflated in comparison to population norms	
5. Would you expect health-related quality of life to be higher for the	Celgene agree that the health-related quality of life of FL patients should not be greater than the general population and agree to use the capped general population values in the base case. In the original company

progression-free and post-progression health states compared with the general population (0.8)?	base case, Celgene used the utility values directly from the AUGMENT trial in order to reflect the most relevant patient population and be consistent with the NICE reference case. These resulted in utilities of 0.867, 0.841 and 0.806 for the progression-free, progressed (off treatment) and progressed (on-treatment) health states, respectively. Using the UK general population utility estimates from the Ara and Brazier model[6], and the average starting age of 65 years, the general population utility value was estimated to be 0.803. Therefore, applying the general population cap in the model currently changes all model health states to 0.803 which doesn't reflect the decrement incurred from moving to another health state. To resolve this issue, Celgene have included another option which caps the progression-free health state at general population utility but keeps the percentage increments between health states as per the AUGMENT trial utilities. These result in utility values of 0.803, 0.780 and 0.747 for the progression-free, progressed (off treatment) and progressed (on-treatment) health states, respectively. In the comparison to R^2 versus R-mono, the average age of patients was 61, resulting in utility values when general population cap was applied of 0.821, 0.796 and 0.762 for the progression-free, progressed (off treatment) and progressed (on-treatment) health states, respectively.
6. Do decrements of -0.026 and -0.056 for utility seem feasible for post-progression compared to progression-free survival?	Celgene agree that the decrement between progression-free and progressed (off treatment) is small, however clinical opinion is that the quality of life for patients is not greatly affected by clinical progression only (i.e. no further treatment required yet). During discussions with the clinical expert at the technical engagement meeting, the expert agreed that patients who progress clinically and remain asymptomatic, may remain quite well and that progression in the absence of symptoms does not have a huge impact on quality of life. In some cases, treatment is not required for several years. Therefore, Celgene feel that the small utility decrements between the progression-free and post progression (off treatment) health state are clinically plausible. Celgene acknowledge that there is still uncertainty with this decrement and have options within the model to use the post-progression utility values from Wild et al[7]. These scenarios are presented below in Table 7. The NICE technical report states *"The company has provided scenario analyses using utility values for health states previously used in technical appraisals (Wild et al., Pereira et al.)."* Celgene would like to clarify that we have not provided a scenario using Pereira et al. (which reports a progressed utility value of 0.458), and this scenario was provided by the ERG. Celgene have dismissed this source as a plausible scenario given the very small patient population (n=22) and the lack of clarity of the methods used to derive this value or the patients involved. In addition, the value of 0.45 was considered far lower than any other reported

values, doesn't capture the slow progressing nature of the disease and is not reflective of clinical opinion that these patients are generally similar to the general population.

Table 7: Utility scenarios

Scenario	ICER vs R-CVP		ICER vs R-mono
	Exponential OS	Weibull OS	
Company's revised base case (see Appendix 1)	£20,156	£26,444	£17,233
Utility post progression (on treatment) from Wild et al (0.62)	£19,934	£23,974	£16,165
Utility post progression (off treatment) and post progression (on treatment) from Wild et al (0.736 and 0.62 respectively)	£20,704	£25,167	£16,069

Additional Questions raised for the company during the engagement T/C

- **Is splitting the decision problem by rituximab-refractory status appropriate?**

The clinical expert said that r-refractory status is 'artificial clinically', and that splitting the decision problem by r-refractory status may not be appropriate. He also mentioned that autologous stem cell treatment (ASCT) is a relevant comparator for people that have been previously refractory to rituximab, especially if they are young and healthy. The proposed clinical pathway and relevant comparators for r-refractory

Celgene would like to address the feedback provided by the ERG and NICE technical team regarding data sources and comparators for the rituximab-refractory Follicular Lymphoma (FL) population and provide a

patients in the model do not currently include ASCT. Does the company have further comments on these issues?	simplified scenario considering a single relapsed/refractory FL population for this submission, in line with the anticipated licence. The company submission originally split the relapsed/refractory FL population into two: • Non-rituximab refractory (comprising all patients not defined as rituximab-refractory) • Rituximab-refractory (patients who did not respond (at least a PR) to rituximab or R-chemo therapy and/or time to disease progression <6 months after last rituximab dose) This was due to: • NICE TA472 guidance recommending obinutuzumab-bendamustine (O-Benda) for use within the Cancer Drugs Fund (CDF) "for treating follicular lymphoma refractory to rituximab" specifically. • Exclusion of rituximab-refractory patients from the pivotal AUGMENT trial • Availability of data from the induction phase of the MAGNIFY study showing favourable R^2 efficacy in both rituximab refractory and non-rituximab refractory FL patient populations **ERG/NICE Technical Team commentary on rituximab-refractory FL** • The concept of a rituximab-refractory population within the setting of previously treated FL was not within the original NICE scope. • O-Benda is a regimen accessed via the CDF, and thus out of scope. • As only data from the pre-randomisation induction phase of the MAGNIFY study was used in the submission, it was considered a single-arm study and thus rejected as a source of R^2 data. • AUGMENT, the pivotal RCT, has been identified as the only R^2 data source (given the rejection of MAGNIFY), thus the only source from which to derive data for rituximab-refractory analyses. • The NICE Technical Team selected rituximab monotherapy as the default comparator for the rituximab-refractory population, having rejected O-Benda (in accordance with the NICE position to

exclude regimens obtained via the CDF as comparators[9]) and due to the inclusion of rituximab monotherapy as the comparator arm in the AUGMENT study.

Company-identified inconsistencies with the ERG/NICE Technical Team commentary

- Per the eligibility criteria of the AUGMENT study, rituximab-refractory patients were specifically excluded, thus AUGMENT does not provide data which applies directly to the rituximab-refractory population.
- Rituximab monotherapy does not offer a clinically logical intervention to treat a population defined as rituximab-refractory.

Technical Engagement TC Clinical Discussions (company minutes)

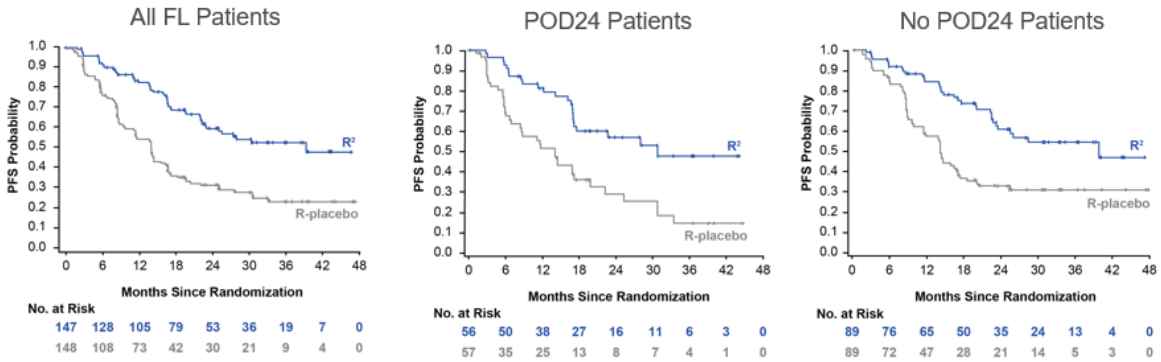
- The concept of rituximab-refractory disease is clinically somewhat artificial; a more clinically valid consideration is the early relapse or immunochemotherapy-refractory population, comprising patients who relapse early following initial treatment with a CD-20 antibody (rituximab or obinutuzumab) combined with chemotherapy. Typically, early relapse is defined as disease progression within 2 years of initial therapy (or, as an alternative definition, within 2 years of initial diagnosis) (POD24).
- In the scenario where O-Benda is not considered a relevant comparator (due to not being routinely funded in the NHS), treatment options relevant for cost-effectiveness analysis are reduced to those available to the general relapsed/refractory FL population. However, as an early relapsing group of patients, further immunochemotherapy is a potentially futile intervention other than as induction therapy for the very small number of patients that are eligible for ASCT (i.e. younger/fitter).

Proposed Reconciliation:

- Celgene recognise NICE's preference to exclude interventions only available via the CDF from the decision problem. Additionally, based on the clinical opinion listed above, a valid alternative to the consideration of the rituximab-refractory population is to consider the population of patients who relapse early following initial treatment with a CD-20 antibody combined with chemotherapy. Based on clinical opinion, the most relevant comparator for these patients is immunochemotherapy (e.g. RCHOP and RCVP).
- AUGMENT and HMRN provide data for patients that experience early relapse (as defined as disease progression within 2 years of initial therapy i.e. POD24).[10]

- Figure 11 presents data from the AUGMENT study, for FL patients, stratified by early-relapse status (i.e. POD24). Table 8 presents the median PFS and HR for FL patients, according to early-relapse (i.e. POD24) status.
- These data demonstrate that the treatment effect of R^2 is maintained in the early relapse population (a population considered to have worse prognosis).

Figure 11: AUGMENT FL patients stratified by early-relapse status



Key: FL, follicular lymphoma; PFS, progression-free survival; POD24, early relapsed; R, rituximab

Table 8: AUGMENT FL patients, median PFS according to early-relapse status

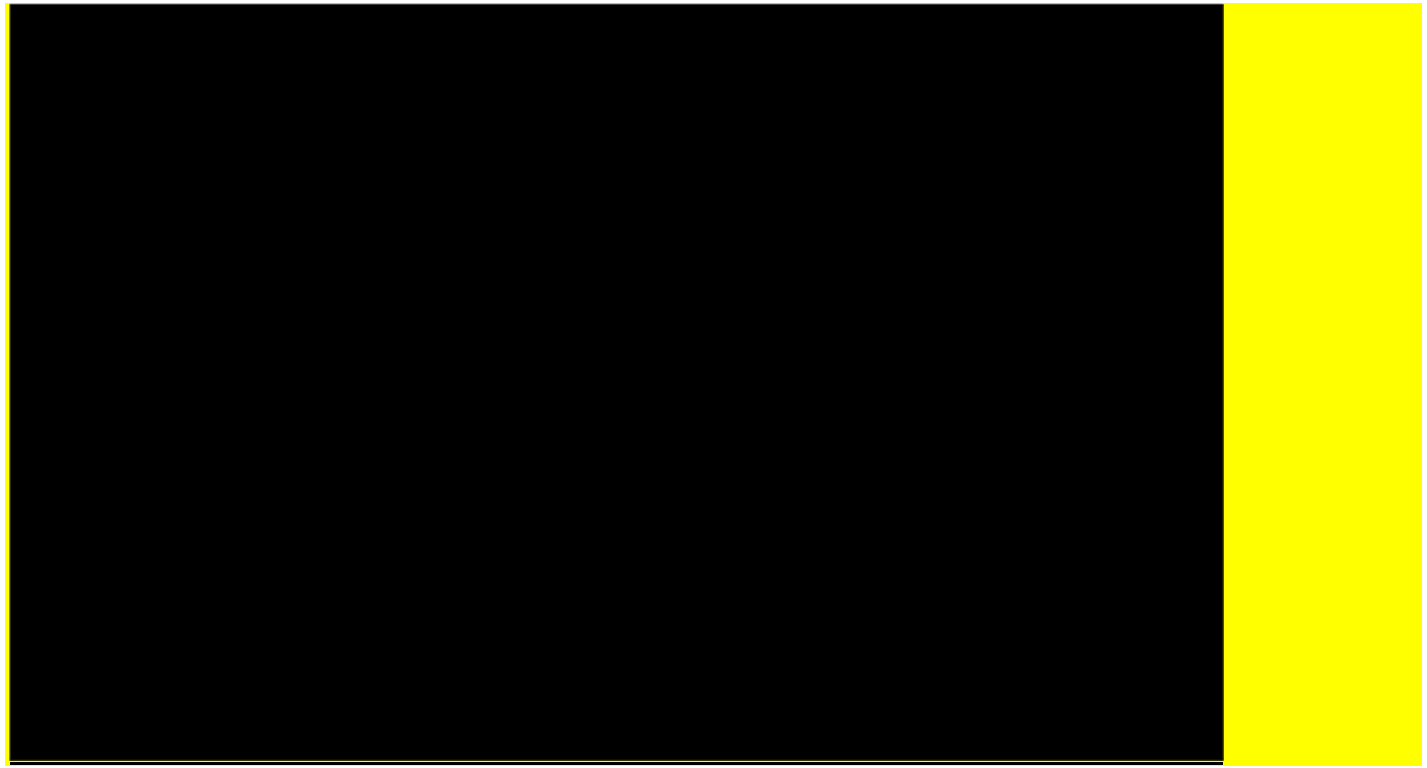
Median PFS, mo (95% CI) *(n R^2/n R-placebo)*	All FL Patients *(n = 147/148)*	POD24 *(n = 56/57)*	No POD24 *(n = 89/89)*
R^2	39.4 (23.1-NR)	30.4 (16.8-NR)	39.4 (22.9-NR)
R-placebo	13.9 (11.2-16.0)	13.8 (6.7-16.9)	13.9 (11.2-16.6)
HR (95% CI)	0.40 (0.29-0.56)	0.41 (0.24-0.68)	0.43 (0.28-0.65)
P value	< 0.0001	0.0004	< 0.0001

Key: FL, follicular lymphoma; PFS, progression-free survival; POD24, early relapsed; R, rituximab

- This proposed reconciliation addresses the ERG/NICE technical team feedback as it:

	o does not rely on comparators funded via the CDF which, while relevant in clinical practice, do not comply with NICE's position statement o ensures population and comparators is aligned with clinical expert opinion provided during the technical engagement, removing separate consideration of a rituximab-refractory population. o uses the AUGMENT study as the main source of data for R^2 for a single group of relapsed/refractory FL patients, with clinical reassurance for its general applicability derived from inclusion of both early and later relapsing patients with favourable efficacy for both groups (see above) • Therefore, the cost-effectiveness analyses provided, based on AUGMENT and HMRN comparing R^2 to R-CHOP/R-CVP and R-mono are representative of the overall relapsed/refractory FL population and address the points raised in the ERG/NICE technical team commentary.
• **Should R-CVP and R-CHOP be pooled?** It has been stated that both R-CVP and R-CHOP are used in separate populations. R-CHOP is also stated to have better PFS than R-CVP, but Stephen O'Brien quoted a meta-analysis that concluded the opposite, adding to the uncertainty as to whether or not pooling is appropriate. Does the company have further comments on this issue?	Further comments on this issue have been raised in Issue 1. Data available for R-CHOP and R-CVP is only in the first-line setting and any inferences made from this to the relapsed/refractory population has been cautioned against by clinicians. Celgene maintain that it is appropriate to pool R-CHOP and R-CVP from HMRN in the relapsed/refractory population given the similarities in the OS, PFS and TTNLT KM outputs and the clinical expectation that it is reasonable to assume similar efficacy.

• **Can the MAIC be improved by requesting data from HMRN?** The company state that all of the covariates included in the HMRN report have been matched on. Stephen O'Brien notes, and the technical team agrees, that the HMRN may have more data available at request (rather than those included in the report alone), and that the company should explore if a better match can be provided. Can a better match be provided?	An additional covariate has been included in the MAIC ("refractory to last therapy") and has been included in the response to Issue 3 and the updated company base-case in Appendix 1. Explanation has been provided for other variables that have not been matched for in Table 4.
Questions posed for the company following the engagement T/C	
Why does altering the treatment waning effect change the ICER in an illogical way? In the non-r-refractory population of the company and ERG model, reducing the treatment waning effect from 5 years decreases the ICER and increasing the treatment waning from 5 years increases the ICER. This is illogical, as we would assume that a longer treatment effect would improve patient outcomes and reduce the ICER, and vice-versa. This pattern is not observed in the original FL-MZL	The treatment waning effect adjusts the PFS, TTNLT and OS curves such that the hazard for R^2 is the same as the comparator from that time point onwards. Therefore, from that time point, the hazard of progressing, receiving next treatment or dying is the same between both treatment arms. No treatment effect is assumed thereafter. The company accept that usually, the treatment waning effect reduces the ICER by increasing the time point the treatment waning is applied and increases the ICER when decreasing the time point. In this case, the opposite occurs due to the resulting hazards when extrapolated using the stratified OS Weibull distributions for R^2 and R-CVP. Figure 12 presents the hazards resulting from the OS Weibull parametric curves which show that the hazard estimated for R^2 is higher than R-CHOP/CVP after nearly 3 years. In the model the adjustment is applied at 5-years resulting in the R^2 curve then following the R-CVP hazard, if the adjustment was not applied at 5-years, the OS curves using the Weibull distributions would cross (see Figure 13). Therefore, applying the treatment waning cut-off at a later time point results in a higher ICER due to the estimated R^2 curve being closer to or below the R-CHOP/CVP curve. It is important to double check the plausibility of any approach taken to extrapolating hazards over an extended period of time. During the technical engagement discussions, the clinical expert agreed that crossing of curves was implausible and this coincides with the clinical feedback Celgene has received. Therefore, applying the adjustment at 5 years ensures that the crossing of curves does not occur. These

non-r-refractory population, or in the r-refractory population. Can the company please explain this phenomenon?	adjustments create R^2 extrapolated outcomes which are clinically plausible, showing an initial benefit in R^2 and the curves slowly merging together after 5-10 years. **Figure 12: OS hazard plots for R^2 versus R-CHOP/CVP –Weibull distribution** 

Key: KM, Kaplan-Meier, OS, overall survival.

Figure 13: OS curves for R^2 versus R-CHOP/CVP – Weibull distribution

Key: KM, Kaplan-Meier, OS, overall survival.

The exponential distribution for OS was also considered plausible by clinical experts (see response to Issue 3), for which applying the treatment waning effect does not change the ICER in an illogical way. Changing the treatment waning effect time point to 10 years reduces the base-case ICER from £20,156 (see company's revised base case Appendix 1) to £14,690.

References

1. Luminari S, Ferrari A, Manni M, et al. Long-Term Results of the FOLL05 Trial Comparing R-CVP Versus R-CHOP Versus R-FM for the Initial Treatment of Patients With Advanced-Stage Symptomatic Follicular Lymphoma. *J Clin Oncol*. 2018; 36(7):689-96.
2. Kim HK, Kang W, Sinn DH, et al. Real world data on follicular lymphoma patients treated by rituximab-containing immunochemotherapy and rituximab maintenance. *Korean J Intern Med*. 2019.
3. HMRN. Clinical Management and Outcome of Follicular Lymphoma (FL) with a Focus on Relapsed/Refractory Disease. 2019. Data on File
4. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluaions Alongside Clinical Trials - Extrapolation With Patient-Level Data. 2011. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed: 10 May 2019.
5. Robert V. Hogg and Elliot A. Tanis. *Probability and Statistical*, 7th Edition ed. 2005.
6. Ara R and Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health*. 2010; 13(5):509-18.
7. Wild D, Walker M, Pettengell R and Lewis G. PCN62 Utility elicitation in patients with follicular lymphoma. *Value in Health*. 2006; 9(6):A294.
8. Pereira C, Negreiro F and Silva C. PSY43 Economic analysis of rituximab in combination with cyclophosphamide, vincristine and prednisolone in the treatment of patients with advanced follicular lymphoma in Portugal. *Value in Health*. 2010; 13(7):A468.
9. National Institute for Health and Care Excellence (NICE). Position statement: consideration of products recommended for use in the Cancer Drugs Fund as comparators, or in a treatment sequence, in the appraisal of a new cancer product. 2019. (Updated: January 2019) Available at: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/cancer-drugs-fund/CDF-comparator-position-statement.pdf. Accessed: 5 May 2019.
10. Leonard JP, Trneny M, Izutsu K, et al. AUGMENT phase III study: Lenalidomide/rituximab (R2) improved efficacy over rituximab/placebo in relapsed/refractory follicular patients irrespective of POD24 status (ABSTRACT 069). 15-ICML International Conference on Malignant Lymphoma. Lugano, Switzerland. 18–22 June, 2019 2019.

Appendix 1: Company's revised base case

Based on the feedback and discussions during the technical engagement phase, Celgene have made changes to their approach and have a revised based case. Table 9 and Table 10 present the company's new base case in comparison to R-CVP and R-mono, respectively.

Table 9: Company's revised base case – R^2 versus R-CVP

Revision	Company rationale	ICER	Change
Technical team preferred assumptions	-	£30,404	-
1. Efficacy data from matching including refractory to last therapy	Refractory to last therapy was considered an important covariate to match on by clinicians. The company have included every covariate available. (see Issue 1, question 2)	£31,521	+£1,117
2. Revise the OS extrapolations based on clinical opinion (exponential or Weibull)	Clinical opinion chose the exponential and Weibull as the most plausible extrapolations for R-CHOP/CVP (see Issue 3, question 4).	£22,742 (exponential) No change (Weibull)	-£7,662 £0.00
3. Revert back to the KM PFS plus comparator hazard approach for R^2 with time point at the end of the KM.	ERG and technical team preferences caused the PFS curves to cross which was considered implausible by clinical opinion. Using the KM+comparator hazard approach means that the initial difference between the treatments is preserved with the same hazard applied at the end of the KM (see Issue 3, question 4).	£26,040	-£4,364
4. Capping utilities at general population but preserving the increments between the health states	To avoid all the health states starting with the same value and acknowledge there are decrements between the health states.	£31,094	+£690
Final company revised ICERs (1+2+3+4)	-	**£20,156 (exponential OS)** **£26,444 (Weibull OS)**	**-£10,248** **-£3,960**

Table 10: Company's revised base case – R^2 versus R-mono

Revision	Company rationale	ICER	Change
Technical team preferred assumptions	-	£14,504*	-
1. Revise the OS extrapolations based on clinical opinion (log-logistic)	Clinical opinion chose log-logistic as the most plausible extrapolation for R-mono (see Issue 3, question 4).	£17,432	+£2.928
2. Capping utilities at general population but preserving the increments between the health states	To avoid all the health states starting with the same value and acknowledge there are decrements between the health states.	£14,466	-£38
Final company revised ICERs (1+2)	-	**£17,233**	**+£2,729**
Note: *The ICER reported in the draft technical engagement report was incorrect (TTNLT curves were set to generalised gamma when these should be log-normal). This ICER is the corrected value.			

Lenalidomide with rituximab for treated follicular lymphoma and marginal zone lymphoma [ID1374]

Addendum for the state-transition model

1. Purpose of the addendum

During clarification questions the evidence review group (ERG) asked the company to provide a state-transition model (STM) as a scenario in order to validate the extrapolations generated by the partitioned survival model (PSM). The company had multiple concerns with providing a STM due to some of the general limitations of a state-transition approach and concerns specific to the method's suitability for this appraisal. In short:

1. Applying the STM approach for the comparison versus R-CHOP/CVP exacerbates the current difficulties given the data for these interventions is 'real-world evidence' and patients' disease progression is therefore not assessed as regularly (deriving eventual OS estimates from intermediary events related to disease progression is therefore more dubious)
2. Applying the STM approach for the comparison versus R-mono is less fraught with methodological concerns, however, given the extremely low UK usage of R-mono, may be somewhat irrelevant (multiple modelling methods are being brought to bare on a comparator of dubious relevance to the UK decision problem)

During the clarification stage discussion with the ERG, the company voiced these concerns however the ERG suggested that the STM should still be provided for the within trial comparison (versus R-mono). The company still had concerns with this approach given the irrelevance of this comparator and therefore chose not to provide the scenario in response to clarification questions (see response to clarification question B2).

During the technical engagement stage, the NICE technical team stated that the results of the STM would have been valuable for cross-comparison of the model extrapolations, and to reduce structural uncertainty associated with the PSM. The company would like to reiterate its concerns regarding the use of STM in this context.

The STM has to make use of non-randomised end points which is highly prone to bias due to selection effects and informative censoring.[1] As stated in response to

clarification question B2, within the FL setting, the timing of relapse can influence the next treatment option and impact on response to the next treatment affecting overall survival.[2] Given that the data available from AUGMENT are immature, the use of such data to inform post-progression or post next therapy outcomes in the STM could induce bias. In addition, extrapolating outcomes from a group of patients who no longer have comparable characteristics and based on patients who are progressing early could also induce bias. In contrast, this is not an issue when using the PSM which models OS directly from the point of randomisation and all patients contribute to the function used to fit the curve. The OS hazard currently observed for the R^2 arm in the PSM, reflects those patients who die quickest, hence the distributions selected for the base case are likely to be conservative.

Despite the concerns stated above, Celgene have provided an STM for the within-trial comparison, comparing R^2 vs R-mono, as a scenario for validation purposes.

2. Methods

The STM has been produced using the multi-state modelling approach with the statistical software R (*mstate* package version 0.210).[3]

Multi-state models are the preferred approach for state-transition models as described by NICE TSD 19.[4] Multi-state models take into account the individual transitions between each of the health states, in this analysis, the timing of these transitions have also been accounted for using state-arrival extended multi-state models.[5]

The current company partitioned-survival model structure encompasses three health states, two of which are split into two sub health states. Given the number of possible transitions between these five health states and the data required to inform each transition, the use of all five health states would make the STM over complicated and unreliable (for example, the transition between progression-free (off treatment) to progressed (on treatment) only has 5 events from R-mono and 2 events from R^2 when considering the ITT population, therefore possibly smaller events when considering the FL only population). Therefore, the STM only considers

three health states; before next anti-lymphoma treatment (NALT), on NALT and death.

NALT has been chosen to define the health states over progression as this is considered more relevant to the patient's health-related quality of life and costs incurred (due to reasons stated in the original submission Section B2.6). This was also considered more reflective of 'progression' in clinical practice given most patients would only be defined as progressed when showing symptoms, at which point patients would receive the next treatment. This is in line with the PSM which incorporated TTNALT as well as PFS.

Individual transitions between each of these health states were estimated from the AUGMENT study for each treatment:

- Before NALT → NALT (time to NALT censoring death)
- Before NALT → death (time to death censoring NALT)
- NALT → death

Stratified parametric survival curves were fit to each transition and the selection of the most appropriate distribution were in line with the previous approaches used in the original submission using guidance from NICE TSD 14.[4] Transitions to death were adjusted for general population mortality if they were estimated to be less than the age- and gender-match general population mortality taken from the ONS life tables.

Estimated overall survival from the state transition model was compared to the PSM using the three-health state scenario assuming that PFS=TTNLT (as per response to clarification question B6). Visual inspection of the extrapolated outcomes between each approach were assessed as well as looking at the incremental differences in survival and cost-effectiveness results.

3. Clinical data

3.1. Before NALT to NALT

Table 1 presents the Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics for each of the distributions for each treatment. Log-normal is shown to be the best statistically fitting.

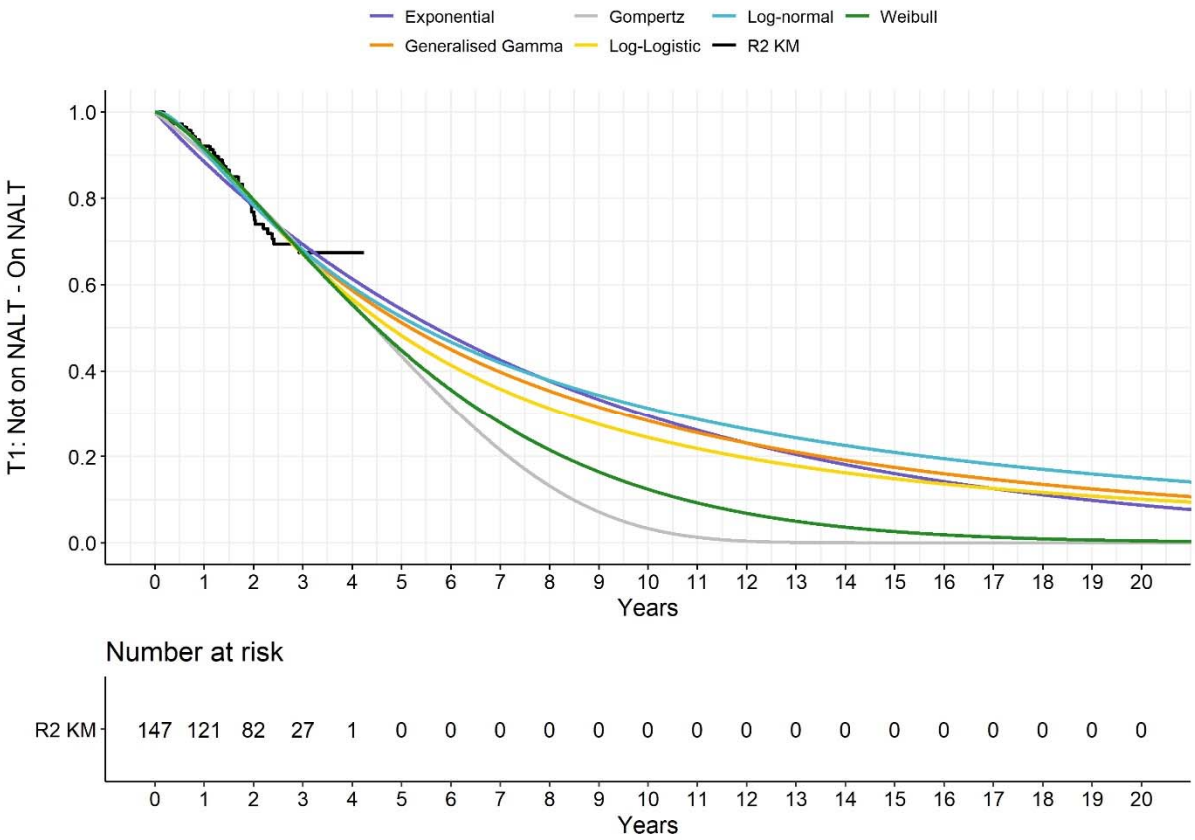
Figure 1 and Figure 2 present the parametric distributions which all show a reasonable fit to the data. Based on these, the log-normal has been selected given this has the best fit to the actual data and is consistent with what has been chosen for the TTNLT extrapolation in the PSM with similar extrapolated outcomes.

Table 1: Before NALT to NALT – AIC/BIC statistics

Distribution	**R^2**		**R-mono**	
	AIC	**BIC**	**AIC**	**BIC**
Exponential	231.62	234.61	318.52	321.52
Weibull	229.43	235.41	318.34	324.33
Gompertz	232.13	238.12	320.51	326.51
Log-Logistic	228.45	234.43	313.37	319.36
Log-normal	**228.25**	**234.23**	**310.41**	**316.40**
Generalised Gamma	230.20	239.17	311.14	320.13

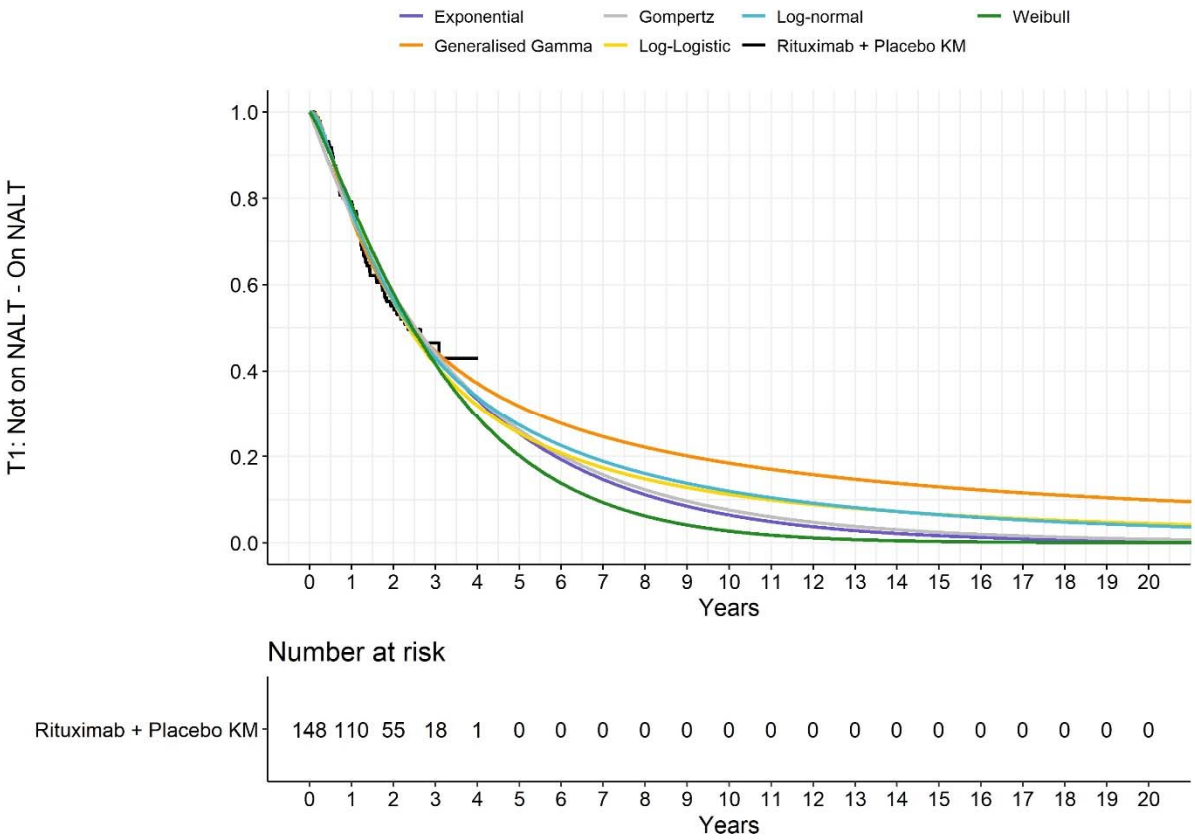
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; NALT, next anti-lymphoma treatment

Figure 1: Before NALT to NALT –Parametric curves: R^2



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

Figure 2: Before NALT to NALT –Parametric curves: R-mono



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

3.2. Before NALT to death

Table 2 presents the AIC and BIC statistics for each of the distributions for each treatment. Exponential is shown to be the best statistically fitting.

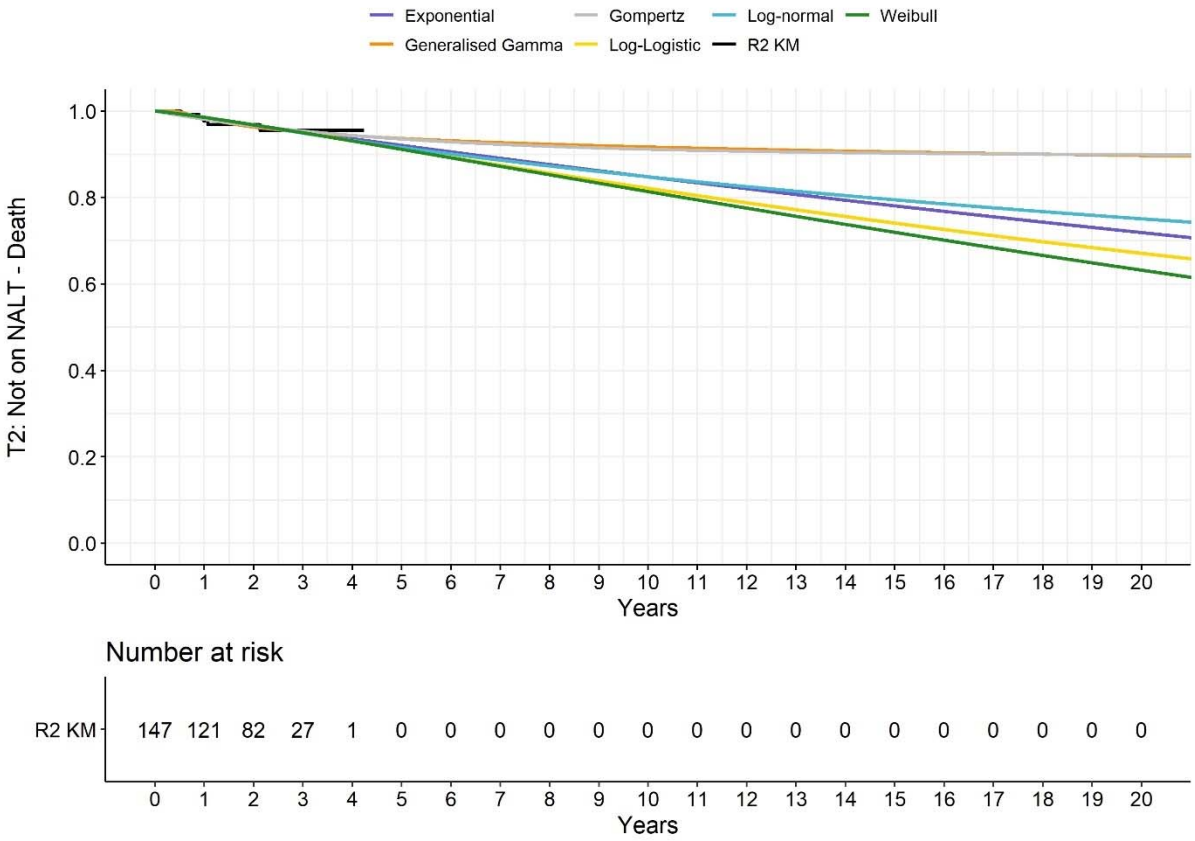
Figure 3 and Figure 4 present the parametric distributions which all show a reasonable fit to the data. Given the small number of events to inform this transition (5 for R^2 and 10 for R-mono), it is assumed that the transition to death before next anti-lymphoma treatment is constant over time, and therefore the exponential distribution has been selected.

Table 2: Before NALT to death– AIC/BIC statistics

Distribution	R^2		R-mono	
	AIC	BIC	AIC	BIC
Exponential	**53.04**	**56.04**	**86.50**	**89.50**
Weibull	54.93	60.91	87.15	93.14
Gompertz	54.92	60.90	87.80	93.80
Log-Logistic	54.88	60.86	87.09	93.08
Log-normal	54.28	60.26	87.00	92.99
Generalised Gamma	53.62	62.59	88.98	97.97

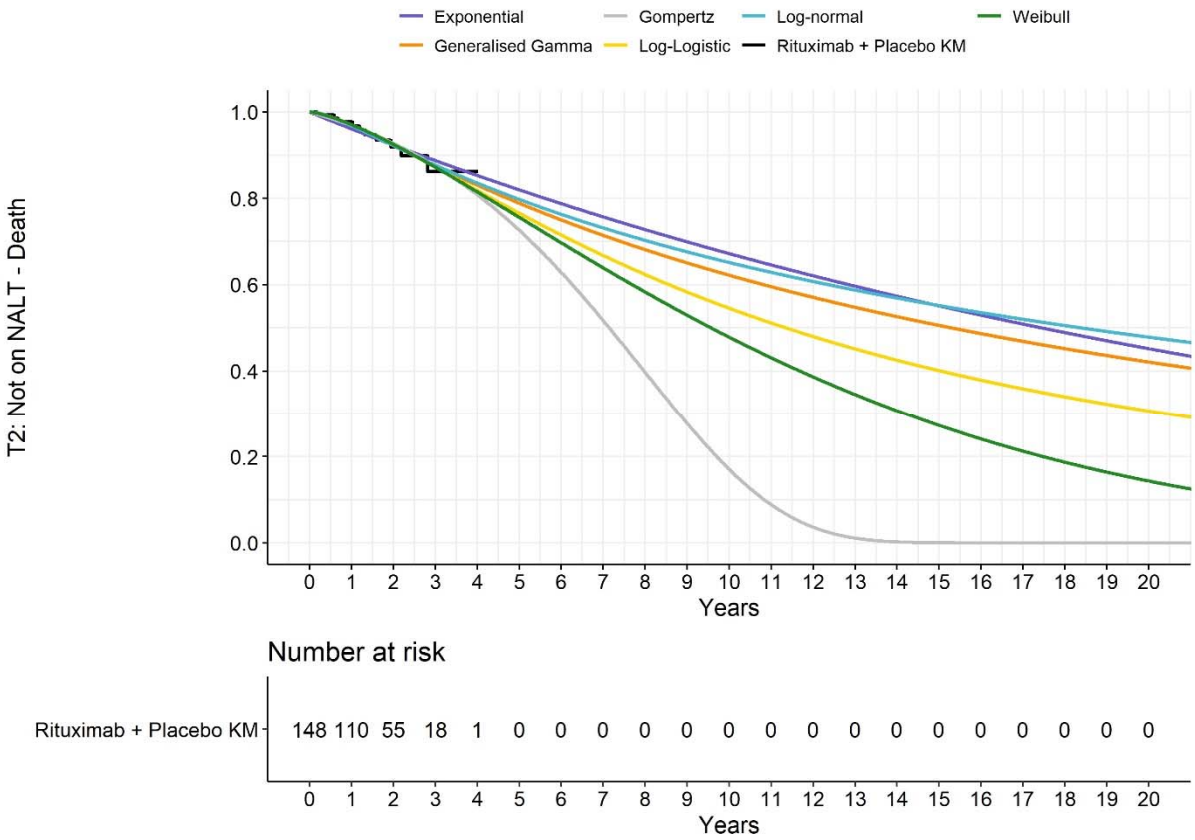
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; NALT, next anti-lymphoma treatment

Figure 3: Before NALT to death –Parametric curves: R^2



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

Figure 4: Before NALT to death –Parametric curves: R-mono



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

3.3. NALT to death

Table 3 presents the AIC and BIC statistics for each of the distributions for each treatment. Generalised gamma is shown to be the best fitting distribution for R^2. Generalised gamma and log-normal are shown to be the best distributions for R-mono according to AIC and BIC, respectively. Figure 6 and Figure 7 present the parametric distributions which all show a reasonable fit to the data.

To validate the plausibility of the extrapolation, we compared the distributions to overall survival at third-line for all patients that received systemic therapy in the HMRN dataset. Though not a like-for-like comparison due to differences between patient populations, this data indicates that 10-year OS for FL patients at third line is between [redacted](see Figure 5), which gives an indication of the upper bound of the survival estimates from a 3L+ population in AUGMENT. In comparison, the

generalised gamma predicts approximately 45% of patient alive after receiving their next anti-lymphoma treatment at 10 years, hence likely overpredicts survival for both treatment arms. Log-normal, which is the 2nd best fitting for R^2 and best fitting according to BIC for R-mono, gives reasonable 10-year survival estimates compared to the HMRN data (i.e. below the estimated survival from a third line only population), and has therefore been selected for these transitions for both treatment arms. The exponential distribution, which is 3rd best fitting based on BIC, also generates reasonable 10-year estimates.

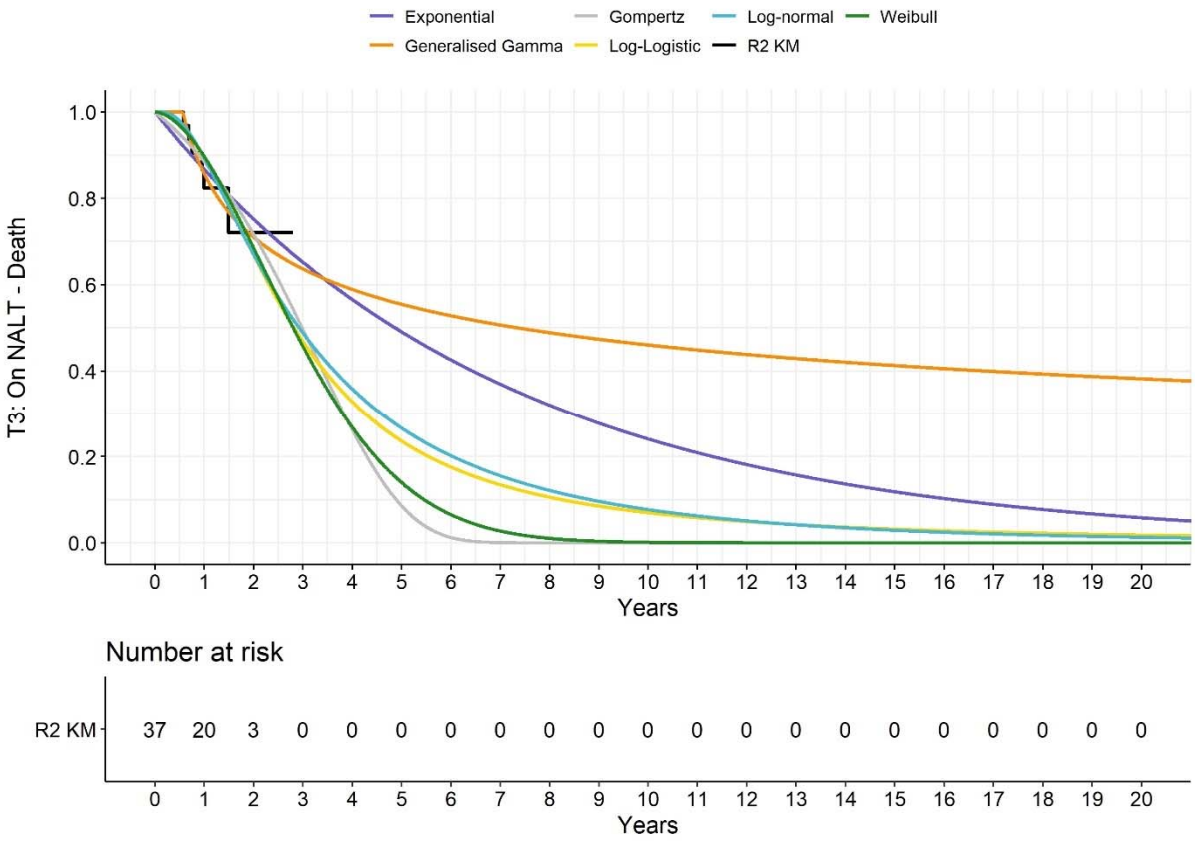
Table 3: NALT to death– AIC/BIC statistics

Distribution	**R^2**		**R-mono**	
	AIC	**BIC**	**AIC**	**BIC**
Exponential	37.39	39.00	81.88	84.11
Weibull	36.95	40.17	81.69	86.15
Gompertz	38.71	41.93	83.70	88.17
Log-Logistic	36.39	39.61	80.64	85.11
Log-normal	35.36	38.58	78.81	**83.27**
Generalised Gamma	**32.84**	**37.68**	**77.08**	83.78

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; NALT, next anti-lymphoma treatment

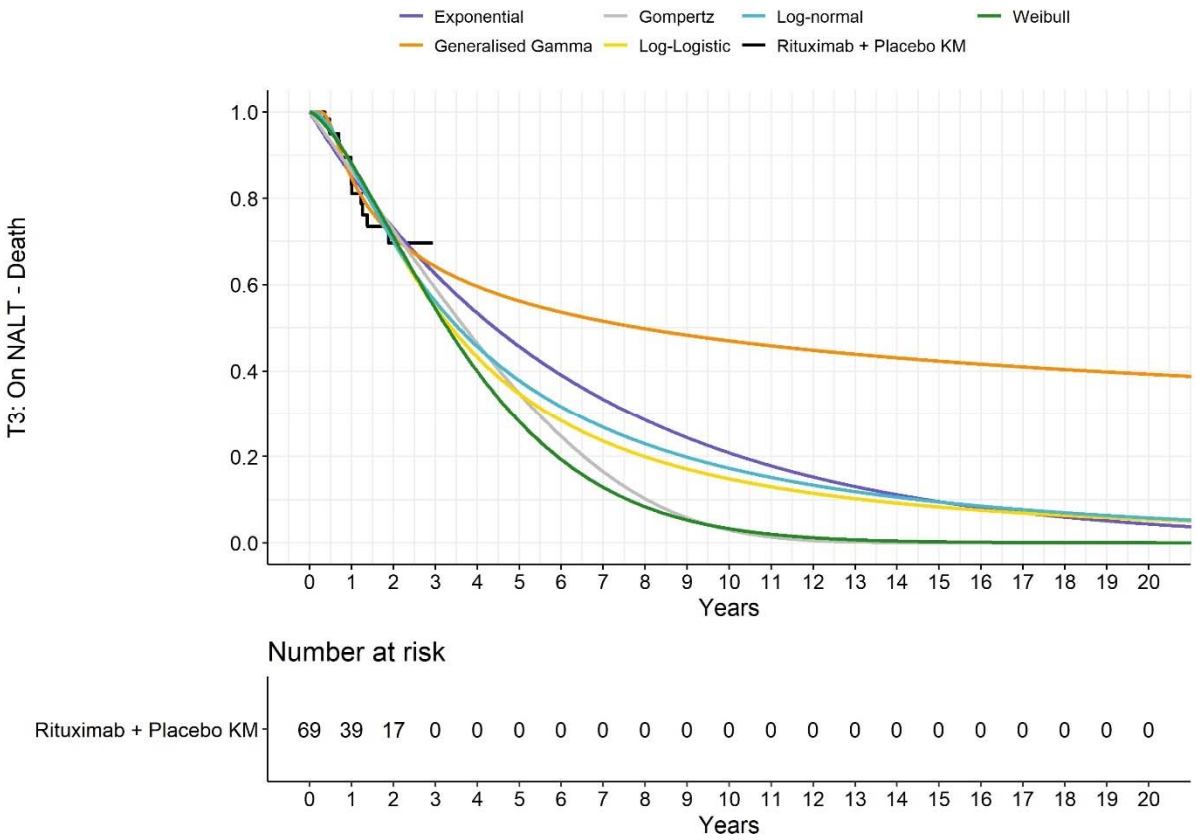
Figure 5: Overall survival FL patients from third line[6]

Figure 6: NALT to death –Parametric curves: R^2



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

Figure 7: NALT to death –Parametric curves: R-mono



Key: KM, Kaplan-Meier; NALT, next anti-lymphoma treatment

4. Health related quality of life

For the STM, the utilities are applied based on the three-health state model (as per the response to clarification question B6). These have additionally been re-analysed using the FL only data from AUGMENT. Based on discussions during the technical engagement call, and preference by the ERG and NICE technical team, the utilities have been capped at the general population data to ensure that they are never greater than the age-adjusted general population norms. Additionally, to ensure that the health state utility values don't predict the same value after general population capping, the company have also maintained the percentage utility increments between the health states estimated from the AUGMENT regression model (see company response to technical engagement Issue 4). Table 4 presents the utility values from the AUGMENT FL only regression model for the three health states and

the resulting utilities after capping for general population whilst maintaining the % utility increments between health states.

Table 4: Utility values for the STM

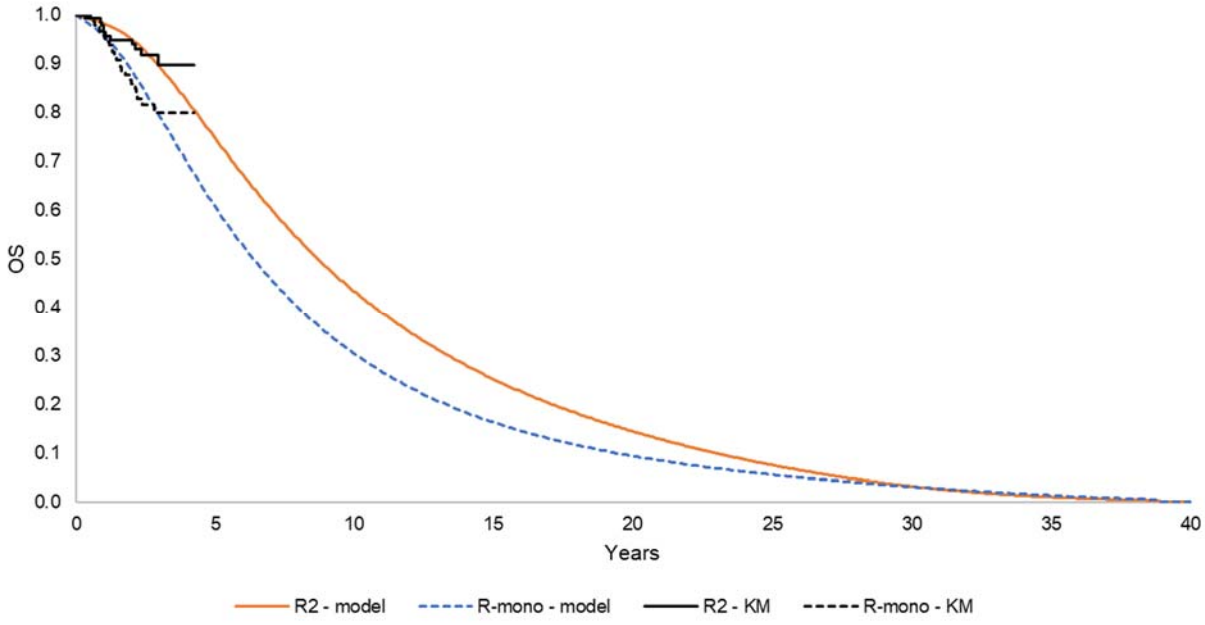
Health state	AUGMENT regression model	Capped at general population*
Before NALT	0.843	0.821
Post NALT	0.791	0.770
Key: NALT, next anti-lymphoma treatment **Note:** *0.821 reflects the age matched general population utility value. 0.770 reflects the utility decrement between health states predicted from the AUGMENT regression model: 0.821 – ((0.843-0.791)/0.843)		

5. Results

Figure 8 presents the OS projections from the STM. These demonstrate a reasonable fit to the AUGMENT OS KM data, based on visual inspection.

Table 5 present the cost-effectiveness results from the STM showing that R^2 is cost-effective versus R-mono at the £30,000 willingness to pay threshold.

Figure 8: Overall survival projections from the STM



Key: KM, Kaplan-Meier; OS, overall survival

Table 5: Results of the STM

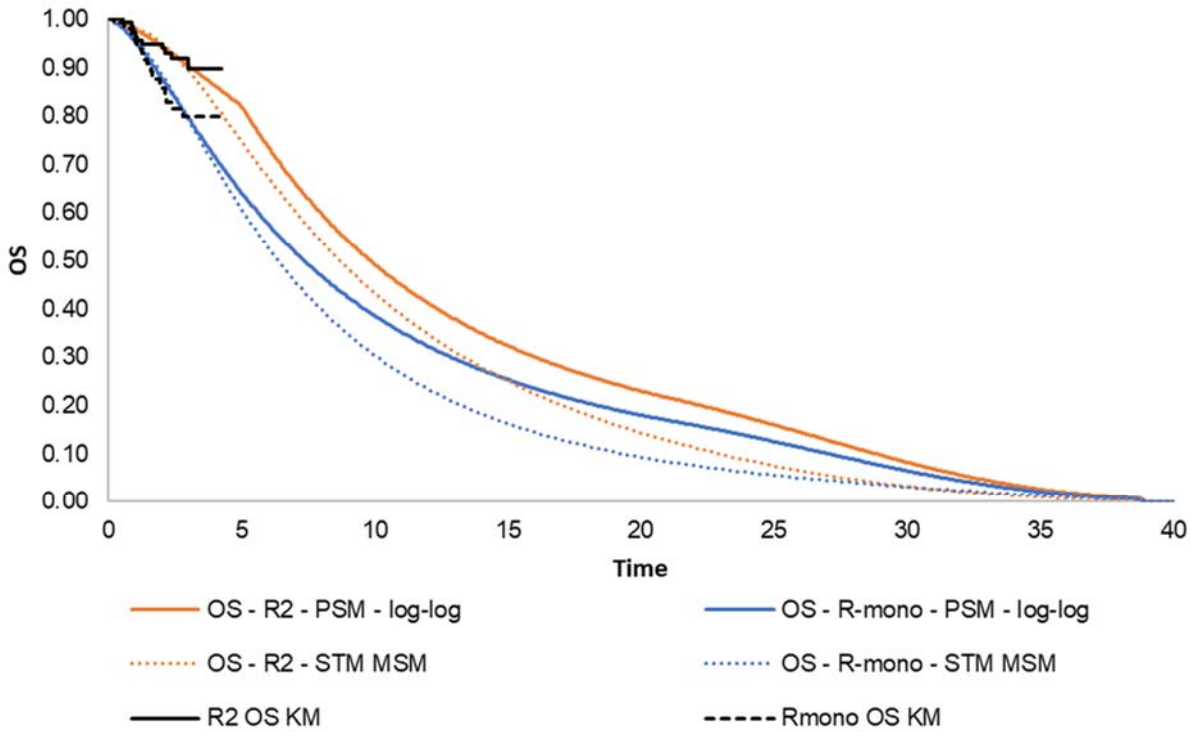
Technology	Total			Incremental			ICER (£/QALYs)
	Costs	LYG	QALYs	Costs	LYG	QALYs	
R-mono	[redacted]						
R^2							£16,160

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; mono, monotherapy; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab.

6. Validation of the PSM

Figure 9 presents the overall survival estimates from the STM in comparison to the PSM extrapolations using the log-logistic distribution. The STM predicts slightly more pessimistic results in comparison to the PSM but the difference between the two treatment arms appears consistent between the two approaches. Table 6 shows that the survival prediction increments from the STM are similar to the PSM.

Figure 9: Overall survival: PSM vs STM



Key: KM, Kaplan-Meier; PSM, partitioned survival model; STM, State-transition model.

Table 6: Survival predictions at certain time points for each approach

Proportion alive time point	PSM*			STM		
	R^2	R-mono	Increment	R^2	R-mono	Increment
5 years	81%	64%	17%	74%	60%	14%
10 years	49%	38%	11%	43%	30%	7%
15 years	32%	25%	7%	25%	16%	9%
20 years	23%	18%	5%	14%	9%	5%
25 years	16%	12%	4%	7%	6%	1%
30 years	8%	6%	2%	3%	3%	0%

Key: LYs, life-years; NALT, next anti-lymphoma treatment
Note: *Using the three-health state model structure

Table 7 present the cost-effectiveness results of the STM in comparison to the results from the PSM using the three-health state scenario. The results show similar incremental costs, LYs and QALYs between the two approaches and give similar ICERs (£17,300 for the PSM and £16,160 for the STM).

Table 7: Results of the PSM versus STM

Technology	Incremental			ICER (£/QALYs)
	Costs	LYG	QALYs	
PSM (3HS)	[redacted]	[redacted]	[redacted]	£17,300
STM	[redacted]	[redacted]	[redacted]	£16,160

Key: HS, health state; ICER, incremental cost-effectiveness ratio; LYG, life years gained; mono, monotherapy; PSM, partitioned survival model; QALY, quality-adjusted life year; R, rituximab; R^2, lenalidomide plus rituximab; STM, state-transition model.

7. Conclusion

An STM was provided at NICE's request to test the sensitivity of the modelling extrapolations to the modelling structure used. For the comparison of R^2 vs R-mono, the incremental OS are consistent across the partitioned survival model and state-transition model approaches and the associated change in the ICER is immaterial.

However, the suitability of the STM is still limited given the immaturity of the data used to inform the transitions and the irrelevance of R-mono as a comparator within UK clinical practice.

Therefore, the company advocate the use of the PSM for decision making.

8. References

1. Woods B, Sideris E, Palmer S, et al. NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care: a Critical review. 2017. Available at: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf. Accessed: 28 May 2019.
2. Celgene. Celgene Indolent Non-Hodgkin Lymphoma Advisory Board Meeting. (Final meeting report: NP-UK-REV-0069) March 2019 2019. (Updated: June 2019)
3. The R Foundation. The R Project for Statistical Computing. 2019. Available at: https://www.r-project.org/.
4. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluaions Alongside Clinical Trials - Extrapolation With Patient-Level Data. 2011. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed: 10 May 2019.
5. Williams C, Lewsey JD, Briggs AH and Mackay DF. Cost-effectiveness Analysis in R Using a Multi-state Modeling Survival Analysis Framework: A Tutorial. *Med Decis Making*. 2017; 37(4):340-52.
6. HMRN. 2019. Data on File.

Clinical Validation – Lenalidomide for treated follicular lymphoma (FL) [ID1374]

A clinical expert was interviewed who provided consent for an anonymised summary document to be used in support of the Celgene Limited submission to NICE: Lenalidomide for treated FL [ID1374].

Clinical validation was sought to justify:

1. Pooling of United Kingdom real-world data from the Haematological Malignancies Research Network (HMRN) for R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone) and R-CVP (rituximab, cyclophosphamide, vincristine, and prednisone)[1].
2. Curve selection for extrapolation[2] of overall survival (OS) and Progression-Free Survival (PFS) data for lenalidomide-rituximab and rituximab monotherapy from the AUGMENT study[3] and for R-CHOP/R-CVP from HMRN.

1. Pooling of HMRN data for R-CHOP and R-CVP

- In the absence of published comparative data for these two regimens in R/R FL, it is reasonable to accept the observed similarity in efficacy outcomes for R-CHOP and R-CVP from the HMRN data, and therefore clinically acceptable to pool the HMRN R-CVP and R-CHOP data.

2. Curve Selection for extrapolations

OS Extrapolations

- R-CHOP/R-CVP: The curves providing the lowest survival estimates at 20 years, Weibull & Exponential, are considered the most clinically plausible extrapolations.
- Lenalidomide-rituximab: Separation of the lenalidomide-rituximab and R-CHOP/R-CVP curves would be expected initially before the curves start to come together at around 5-10 years following treatment.
- Rituximab monotherapy: The Log-logistic curve provides the most clinically plausible estimate for 20-year survival (18%); the Generalised Gamma curve provides a clear over-estimate of 20-year survival (35%).

PFS extrapolations

- R-CHOP/R-CVP: The curves providing the lowest progression-free estimates at 20 years, Exponential & Generalised Gamma, possibly also Weibull, are considered the most clinically plausible.
- With respect to the potential for the PFS curves for lenalidomide-rituximab and R-CVP/R-CHOP to cross at some point following treatment, the clinical expert indicated that the curves were “more likely to merge than to cross”.

1. Celgene Data on File; HMRN Draft Report (v7.1 17th July 2019)
2. Celgene Data on File; R2 CE Model v2.1
3. Celgene Data on File; AUGMENT Clinical Study Report (CC-5013-NHL-007 CSR), 9th Nov 2018

NICE National Institute for Health and Care Excellence

Technical engagement response form

Lenalidomide for treated follicular lymphoma and marginal zone lymphoma [ID1374]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **5pm** on **Monday 16 December 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential

information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	Janssen-Cilag UK
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

Questions for engagement

Issue 1: The MAIC is highly uncertain	
1. Is there sufficient evidence to suggest that R-CHOP and R-CVP could be assumed equally efficacious in clinical practice?	No comment
2. Can the MAIC be considered appropriate given the extent of covariates that were not included for matching?	No comment
Issue 2: The model structure may be inappropriate	
3. Would re-analysis of the data, through use of a Markov model, be useful to reduce uncertainty?	We appreciate that in the AUGMENT phase 3 trial[1], median overall survival has not been reached in either arm in the final analysis (median follow-up :28.3 months), and that as such trial data can be deemed immature: • Using a state transition model (STM) such as a Markov model with immature survival data may introduce bias as post-progression survival will be predicted using only data from the patients with worst prognoses; i.e. patients progressing and dying early on in the study • A survival partition model (SPM) may introduce bias depending on the underlying hazard rate (rate of death moving forward). If the hazard rate is likely to increase over time, then a SPM may overestimate survival. If the hazard rate is likely to decrease over time, that is

[1] https://ascopubs.org/doi/full/10.1200/JCO.19.00010?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed

	the rate of deaths slows (a more likely scenario in situations where some patients have a better or more sustained response than others) then the SPM may underestimate survival • With immature data, the value of STM is very limited versus exploration of the likely underlying hazard rate which can be validated with clinical experts. Re-analysis of the data through use of a Markov model therefore does not appear to be the most appropriate approach to reduce uncertainty.
Issue 3: Kaplan-Meier extrapolations are highly uncertain	
4. Upon visual inspection, which parametric distributions appear to be most suitable for extrapolation?	No comment
Issue 4 – Utility scores appear inflated in comparison to population norms	
5. Would you expect health-related quality of life to be higher for the progression-free and post-progression health states compared with the general population (0.8)?	No comment
6. Do decrements of -0.026 and -0.056 for utility seem feasible for post-progression compared to progression-free survival?	No comment

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma

Critique of the company's response to Technical Engagement and responses to specific questions from NICE.

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Date completed 08/01/2020

Questions for ERG relating to company response to technical engagement (Received on 7 January 2020)

Utilities (p.22-24):

On the engagement T/C it was decided that HRQoL would be capped at general population norms. This was because EQ-5D data collected in the AUGMENT trial showed post-progressed and progression-free HRQoL values that were above those of the age-related general population.

The company have since added a further suggestion (p.22-23 response to engagement) to cap the progression-free health state to the general population norm, and uses the relative utility decrements from the AUGMENT trial to reduce the HRQoL for post-progressed (on/off treatment) health states.

The company have also used age-related general population utility values from Ara and Brazier (2010) rather than Kind et al (1999).

- Does the ERG agree with using relative decrements for HRQoL (rather than having all health states capped at general population norms), and for using Ara and Brazier general population norms (rather than Kind et al.)?

ERG: the ERG is fine with using the Ara and Brazier population norms and age decrements - these were already used in the original submission. As for the proposed way of implementing the decrements, the company have only changed this in their updated model of mid-December 2019 as part of a 'company revised base-case' which was not asked for to the ERG's knowledge. The way utilities were capped in the ERG base-case was as designed by the company in their original model. That said, the ERG does in itself consider it sensible to use differential utilities and HRQoL as observed in AUGMENT may be the best way to derive these. However, although the decrement for the progression off treatment is comparable between FL-only and FL+MZL ([redacted] vs. 2.9%), the decrement for post-progression on treatment is higher in FL-only ([redacted] as compared to 3.6% for FL+MZL). Given that FL-only is supposed to be a younger and fitter population this raises the question whether there are enough observations in AUGMENT for FL in this phase (post-progression on treatment) to properly inform these decrements. Decreasing the post-progression on treatment decrement would (slightly) increase the ICER.

Is splitting the decision problem by rituximab-refractory status appropriate (p.24-28):

The company has moved away from separating treatment choices for individuals with rituximab-refractory and non-rituximab refractory follicular lymphoma. The company consulted clinical experts think that early relapse (relapse within 24 months) to first-line R-chemo is a more clinically valid metric to separate treatment choices by.

AUGMENT and HMRN provide data for patients that have previously relapsed early (to first line R-chemo). This is defined as 'POD24' (relapse within 24 months) in AUGMENT. The company state that since POD24 individuals and non-POD24 individuals respond equally well to lenalidomide with rituximab in progression-free survival (based on graphs shown on p.27), that lenalidomide with rituximab is a relevant comparator for individuals with or without early relapse. Therefore data for all FL patients can be used, rather than POD24 alone.

The company argue that despite initial disagreement in how the model was split, the model that the company are using is still relevant, as it compares lenalidomide with rituximab for the full FL population (as lenalidomide with rituximab is argued to be equally effective in both early relapse and non-early relapse), to relevant comparators. The company further argue that rituximab refractory status is no longer relevant for this submission, but since the AUGMENT trial exludes r-refractory individuals, the analysis is for the non-r-refractory population.

- Have I interpreted this correctly?

If so: I have some concerns which you may share, as below. Do you have any comments on these concerns:

- Do the graphs shown on p.27 indicate that the POD24 and no POD24 data can be grouped, and we can consider all FL patients? Or should we be using the POD24 and no POD24 populations as separate subgroups?
  ERG Response: I don't see why the population should be split by POD24 status. The question is whether results are applicable to R-refractory patients, and I don't see how this can be evidence for that – see below (especially, as all data from AUGMENT are for non-R-Refractory only).
- From the graphs on p.27 we only have PFS data for POD24 – do we need to see overall survival for POD24 before assuming that lenalidomide with rituximab is equally effective for POD24 and no POD24 patients?
  ERG Response: OS data are immature, so will probably be inconclusive. In addition, I think the main problem is that data are from a non-R-refractory population only.
- Even if we are happy with using the 'all FL patients' rather than POD 24 or no POD 24 subgroups, are we equally happy to group early relapse and no early relapse data from HMRN together? It seems that there would be a worse PFS and OS for R-CHOP/R-CVP in the early relapse HMRN individuals than those that did not relapse early.
  ERG Response: The populations in AUGMENT and the HMRN are different. HMRN patients are all from the UK; therefore, initial therapy in most cases is R. That means that in AUGMENT POD24 patients have relapsed on something other than R and in HMRN most patients have relapsed after R. It's unclear what means for the results. It is definitely more valid to compare similar populations with regard to initial therapy. In addition, all our original critique regarding the MAIC still applies.
- Overall, do you have any other concerns with this change from the company?
  ERG Response: See below.

ERG Response: The company seems to argue that since results are similar for POD24 and non-POD24 patients, the results from AUGMENT are valid for all FL patients (not just non-R-refractory patients).

That raises two questions:

1. Is POD24 a substitute for R-refractory status; and

2. Are the results for POD24 and non-POD24 patients indeed similar.

I'm not convinced by either of those two:

1. POD24 is defined as relapse within two years of initial chemoimmunotherapy. But that does not need to be relapse after R. In fact, we have no idea what initial therapy was. Therefore, I don't think POD24 can be used as an indicator of R-refractory status.

2. Results are shown in Table 8 and Figure 11 of the company response to TE. Table 8 shows indeed similar HRs for PFS, but the curves in Figure 11 show considerable differences. So, I'm not convinced that results are indeed similar.

Finally, there is a major flaw with the company's line of reasoning. They present results from a non-R-refractory population (AUGMENT trial) to argue that results are similar for R-refractory patients and non-R-refractory patients; that is like Baron Munchausen pulling himself (and his horse) out of a swamp by his hair. Even if POD24 can be regarded as a substitute for R-refractory status and if we agree that results are similar for POD24 and non-POD24 patients. The data presented only apply to the population in AUGMENT (non-R-refractory); therefore, it is still unclear if the same applies to R-refractory patients.

Why does altering the treatment waning effect change the ICER in an illogical way (p.29-31):

The technical team notes that changing the treatment waning effect led to counterintuitive ICERs for the FL only population; increasing the treatment waning effect led to increasing ICERs. The ERG agreed that changing the treatment waning effect led to counterintuitive results, and noted that this relationship was not previously seen in the mixed FL and MZL population. The ERG noted that there may be an error in the new FL only model which is causing this.

P.30 of the company response to engagement shows the Weibull hazards for lenalidomide with rituximab and R-CHOP/R-CVP.

- Are the hazard functions what we would expect or do they indicate something unusual?

ERG: It is difficult to say what would have been expected. The rise of the hazard for the R^2 arm is indeed quite steep but the fact that at some point the hazards cross (and so the hazard of R^2 would lie above that of the comparator) is not unlikely per se. In fact, in the FL+MZL population the hazards also cross (see Figure 1), but at a much later point and therefore it has less impact and does not show when varying treatment effect duration from 5 to 7 or 10 years. But also, the log-normal function in the FL-only population shows the crossing hazards (see Figure 2), and will produce counterintuitive ICERs when increasing the treatment duration to around 10 years. The same goes for log-logistic and Gompertz (graphs not shown). An exponential function has constant hazards by definition (see Figure 3) and can never cause counter-intuitive results in varying treatment effect duration, as it is restricted in this sense.

The company have, in their addendum, expressed a preference for Weibull, log-normal, or exponential OS, with Weibull being the most logical and consistent choice. The ERG feels that the argument of counter-intuitive findings is not sufficient to discard both Weibull and log-normal and use only exponential.

Figure 1: Weibull hazards for FL+MZL population

Figure 2: Log-normal hazards for FL-only population

Figure 3: Exponential hazards for FL-only population

Does the ERG maintain that the counterintuitive ICER results are likely based on a model error?

ERG: No, the ERG did not discover any error in this respect.

- Is it possible that the counterintuitive ICER results are due to limited follow-up data (46.7 months), with small changes in the KM curves in the FL only population relative to the MZL + FL population causing the different relationships when extrapolated?

ERG: It is difficult to say what exactly is causing the hazards to behave as they do. Immaturity of the data cannot be ruled out as a cause. But as said, the company's now proposed exponential curve is almost the only curve that does not to some extent have these counter-intuitive ICER results, because it imposes constant hazards.

ERG Response to Celgene Response to TE.

1. Is there sufficient evidence to suggest that R-CHOP and R-CVP could be assumed equally efficacious in clinical practice?

ERG: Treatments are generally considered for different populations (R-CVP for older patients and R-CHOP for younger patients); therefore, the effectiveness of R-CHOP and R-CVP is difficult to compare.

2. Can the MAIC be considered appropriate given the extent of covariates that were not included for matching?

ERG: MAIC is now as good as it can get. However, the analysis is still based on an unanchored MAIC involving two single treatment arms from different studies. This analysis makes the assumption that all effect modifiers and prognostic factors are accounted for in the model, which in practice is difficult to achieve as, in this case, one or both studies do not measure a specific variable.

Therefore, the ERG still believes that the results of the MAIC should be treated with a high degree of caution due to the fact that potentially important covariates were not included in the matching models (because they were not measured in one of the trials), small sample sizes and assumptions about the equivalence of R-CHOP and R-CVP in the HMRN data and differences in PFS definitions and length of follow-up between the two data sources.

3. Would re-analysis of the data, through use of a Markov model, be useful to reduce uncertainty?

ERG: See ERG addendum dated 20 December 2019.

4. Upon visual inspection, which parametric distributions appear to be most suitable for extrapolation?

→ Celgene: R2 vs R-CHOP/CVP - In conclusion, the ERG and the NICE technical team have chosen the Weibull distribution as the most appropriate for OS extrapolations. Further UK clinical validation is supportive of the Weibull, and in addition, states the exponential as another plausible distribution for OS extrapolation. Therefore, the company advocate for the use of the Weibull or the exponential distributions for OS extrapolations in the R2 vs RCHOP/RCVP comparison.

→ R2 vs R-monotherapy - The ERG and NICE technical team have chosen the generalised gamma distributions for OS extrapolations for R2 and R-monotherapy. However, given expectation that this overestimates survival, for both arms, Celgene advocate for the use of the log-logistic distribution for R2 and R-monotherapy (supported by UK clinical validation).

ERG: The ERG have actually not chosen one OS curve. Given the uncertainty in the OS extrapolations and the expected impact this would have, the ERG have performed 6 base-cases, one for each OS curve. The rationale for this is mentioned in the ERG report following the company FL-only addendum in section 5.2.6.

5. Would you expect health-related quality of life to be higher for the progression-free and post-progression health states compared with the general population (0.8)?

→ Celgene agree that the health-related quality of life of FL patients should not be greater than the general population and agree to use the capped general population values in the base case. But Celgene uses different method to cap.

ERG: see above (utilities)

6. Do decrements of -0.026 and -0.056 for utility seem feasible for post-progression compared to progression-free survival?

→ Celgene acknowledge that there is still uncertainty with this decrement and have options within the model to use the post-progression utility values from Wild et al.

ERG: This may also need clinical expert opinion, as for HE it is difficult to judge what impact progression would have.

Additional Questions raised for the company during the engagement T/C:

- Is splitting the decision problem by rituximab-refractory status appropriate?

→ Celgene recognise NICE's preference to exclude interventions only available via the CDF from the decision problem. Additionally, based on the clinical opinion listed above, a valid alternative to the consideration of the rituximab-refractory population is to consider the population of patients who relapse early following initial treatment with a CD-20 antibody combined with chemotherapy. Based on clinical opinion, the most relevant comparator for these patients is immunochemotherapy (e.g. RCHOP and RCVP).

ERG: For NICE to decide (see also above).

- Should R-CVP and R-CHOP be pooled?

ERG: See above (point 1).

- Can the MAIC be improved by requesting data from HMRN

ERG: See above (point 2)

Questions posed for the company following the engagement T/C:

- Why does altering the treatment waning effect change the ICER in an illogical way?

ERG: See above.

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Lenalidomide with rituximab for previously treated follicular lymphoma and marginal zone lymphoma

Top-level ERG critique of the company's updated CE Model

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Date completed 19/12/2019

In this report, the ERG provides a top-level critique of the evidence provided by Celgene in an addendum on December 13th 2019. The addendum contains results of a state transition model (STM), complementary to the results of the partitioned survival model (PSM) as provided in the original submission. Given time constraints and after discussion with NICE, the ERG has limited this to be a top-level critique only. The full critique will be available after the committee meeting of January 22nd 2020 as agreed with NICE.

Summary and critique

The company provided STM results for a comparison with only R-mono, as R-CHOP and R-CVP data were based on the HMRN registry and therefore not measured as regularly – this posed methodological problems in constructing STM for R-CHOP and R-CVP, as stated by the company. R-mono does not suffer from these problems as much but the company considers this to be a less relevant comparator given its limited use in UK clinical practice.

- The ERG considers cross validation by comparing PSM versus STM results to be hampered by the absence of R-CHOP and R-CVP comparators in the STM model.

The company chose to model the STM with three health states: before next anti-lymphoma treatment (NALT), NALT, and death, as splitting the health states into on treatment and off treatment (as in the PSM) would cause difficulties in estimating probabilities because of low patient numbers transitioning.

- The ERG appreciates the rationale for this choice, but considers it to limit the comparability with the original submission and previously reported results.

Three curves were estimated (before NALT to NALT, before NALT to death, NALT to death) all were stratified, and choice for curves was based predominantly on AIC/BIC – with exception of NALT to death probability where 2nd best was used but this was well explained and based on survival as observed in HMRN 3rd line.

- Before NALT to death curves for R^2 and R-mono are considerably different (for instance, 10-yr estimated survival for R^2 was around 85% and for R-mono around 65%) while these were informed by a small number of events (5 in R^2 and 10 in R-mono), increasing the uncertainty. The ERG questions the clinical plausibility of the substantially different survival profiles while still in 2nd line.

Resulting costs, LYs, QALYs and ICERs were comparable to those found in the PSM – that is, when in the PSM the switch for PFS equal to TTNLT was turned to 'yes'. ICER resulting from the STM was £16,160 while for the PSM it was £17,300

- The ERG could not exactly reproduce the QALY and ICER results for the PSM as stated by the company in this addendum from the previously submitted version of the model, but the difference was minor. Costs and LYs were equal.
- Although the ICER for the STM is comparable to ICER in the PSM, this is only the case for the 'PFS equal to TTNLT' setting, which was not used in the analyses previously reported by the company or the ERG. The ICER for R^2 versus R-mono in the original company base-case (i.e. without this setting) was £20,274.

- The company refers to the 'PFS=TTNLT' setting in the PSM as a 3-health state approach. However, in the original PSM model, there were 5 health states (see also CS page 128):

  • PF (on-treatment) = ToT data

  • PF (off-treatment) = PFS – ToT

  • PP (off-treatment) = TTNLT – PFS

  • PP (on-treatment) = OS – TTNLT

  • Death = 1 – OS

  By setting TTNLT equal to PFS, the 'PP off-treatment' would be excluded, but the 'PF off-treatment' is still in. So, there would still be a difference between model structures of the PSM and STM in the sense that the PSM includes on and off treatment for PF (and therefore has 4 health states), while the STM does not. Any bias caused by this would only be situated in the costs of the PF states and will therefore probably be limited, but the direction of the potential bias is unclear, given that it is difficult to derive from the addendum how costs for on and off treatment were allocated to the PF state in the STM.

Conclusions of the company were that the STM is of limited suitability, given the immaturity of the data used to inform it, and irrelevance of R-mono for UK clinical practice. Overall survival and ICER were nevertheless comparable to what was found from the PSM. The company advocates the use of PSM results for decision making

- The ERG agrees that the STM is of limited suitability, but mostly because of the absence of STM results for R-CHOP and R-CVP, which will only allow cross-validation for the R-mono comparison. The company concludes that results for R-mono are comparable between PSM and STM, but this is only the case when the PSM is put to a setting (TTNLT=PFS) not used before in base-case or scenarios. So, in essence, although the company did provide an STM, it does not apply to any relevant analysis. Therefore, the STM cannot be used to cross-validate or confirm the PSM structure and results.