Axicabtagene ciloleucel for treating relapsed or refractory low-grade non-Hodgkin lymphoma [ID1685]

Produced by Aberdeen HTA Group

Authors Thenmalar Vadivaloo[1]
Graham Scotland[2]
Corinne Booth[3]
Clare Robertson[1]
Moira Cruickshank[1]
Lorna Aucott[1]
Charlotte Kennedy[2]
Paul Manson[1]
Gavin Preston[4]
Miriam Brazzelli[1]

1 Health Services Research Unit, University of Aberdeen, UK
2 Health Economics Research Unit, University of Aberdeen, UK
3 Health Economist, Independent Consultant
4 NHS Grampian, Aberdeen, UK

Correspondence to Miriam Brazzelli
Health Services Research Unit, University of Aberdeen
3rd Floor, Health Sciences Building, Foresterhill
Aberdeen, AB25 2ZD
m.brazzelli@abdn.ac.uk

Date completed 15/03/2022

Contains AIC/CIC

Source of funding: This report was commissioned by the NIHR Systematic Reviews Programme as project number **135434**.

Declared competing interests of the authors

No competing interests to disclose.

Acknowledgements

Copyright is retained by Kite for Tables 3, 7, 13,15-20, Figures 1,2,3 and 4 and text referenced on page 11, 14, 23.

Rider on responsibility for report

The view 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*: Vadiveloo T, Scotland G, Booth C, Robertson C, Cruickshank M, Aucott L, Kennedy C, Manson P, Preston G, Scotland G, Brazzelli M. Axicabtagene ciloleucel for treating relapsed or refractory low-grade non-Hodgkin lymphoma [ID1685]. Aberdeen HTA Group, 2022.*

Contribution of authors

Moira Cruickshank and Clare Robertson summarised and critiqued the clinical effectiveness evidence; Thenmalar Vadiveloo and Lorna Aucott checked and critiqued the statistical analyses presented in the company submission; Graham Scotland was the health economics lead for the appraisal and with assistance from Corinne Booth and Charlotte Kennedy reviewed and critiqued the cost-effectiveness evidence and the economic model; Paul Manson checked and critiqued the company's search strategies; Gavin Preston provided clinical guidance and comments on the draft report. Miriam Brazzelli was the clinical effectiveness lead for the appraisal and coordinated all its aspects. All authors contributed to the writing of this report and approved its final version.

Table of contents

List of tables

List of figures

List of abbreviations

4L+	Fourth-line plus (three or more lines of prior therapy)
AE	Adverse event
BOR	Best overall response
CAR T-cell	Chimeric antigen receptor T-cell
CR	Complete response
CRD	Centre for Reviews and Dissemination
CRS	Cytokine release syndrome
CS	Company submission
CSR	Clinical study report
DOR	Duration of response
ERG	Evidence review group
FAS	Full analysis set
FL	Follicular lymphoma
HRQoL	health-related quality of life
IAS	Inferential analysis set
iNHL	Indolent non-Hodgkin lymphoma
mITT	Modified intent-to-treat
MZL	marginal zone lymphoma
NHL	Non-Hodgkin lymphoma
NICE	National Institute for Health and Care Excellence
ORR	Objective response rate
OS	Overall survival
PFS	Progression free survival
POD24	Progression of disease within two years of receiving front-line chemoimmunotherapy
PR	Partial response
r/r	Relapsed or refractory
RCT	Randomised controlled trial
SAS	Safety analysis set
SLR	Systematic literature review
TEAE	Treatment-emergent adverse event

1. Executive summary

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.

All issues identified represent the ERG's view, not the opinion of NICE.

1.1 Overview of the submitted evidence and ERG's key issues

The focus of the submission received from Kite is axicabtagene ciloleucel (referred to throughout as axi-cel) for treating follicular lymphoma (FL), which is the most common subtype of indolent non-Hodgkin lymphoma, and specifically relapsed or refractory FL.

The clinical evidence submitted by the company consists of an on-going, single-arm, multicentre, open-label phase II trial: ZUMA-5 in which most FL patients had received at least three prior lines of therapy. The overall response rate (ORR; defined as the incidence of complete response [CR] or partial response [PR]) was [redacted] for the inferential analysis set (IAS). CR was achieved by [redacted] of participants. The median duration of response (DOR) was not reached in all responders: [redacted]. The median follow-up for DOR was [redacted]. [redacted] responders had an ongoing response at censoring. At the time of analysis, [redacted] of participants were alive and progression-free. The median PFS [redacted] The median follow-up time for PFS was [redacted]. The median OS was not reached [redacted] The median follow-up time for OS was [redacted] patients had died at the time of analysis. The clinical outcomes used in the economic model are progression-free survival (PFS), overall survival (OS) and adverse event

incidence. The company's literature review identified several studies providing evidence in relevant contexts but did not use any of this evidence in the submission. It is unclear to the ERG if the company's strategy was appropriate as reasons for not including each individual study were not reported by the company, despite being requested at clarification. Instead, the company presented comparative evidence from an external cohort study, SCHOLAR-5. Although there were differences in the distribution of ECOG score between ZUMA 5 and SCHOLAR-5, based on the opinion of their clinical expert, the ERG accepts SCHOLAR-5 as the comparator given the lack of randomised evidence.

The company present a de Novo economic model to determine the cost-effectiveness of axi-cel versus therapies currently available in the NHS in England for [redacted], referred to as the r/r FL 4L+ population throughout. The model takes the form of a partitioned survival model, with efficacy inputs for axi-cel derived from parametric survival analysis of OS and PFS data for the relevant subgroup of ZUMA-5. Efficacy inputs for current 4L+ care are derived from parametric survival analysis of propensity score weighted PFS and OS data from the SCHOLAR-5 study. The company assume that a proportion of patients treated with axi-cel can be considered long-term survivors from a future time point, and thereafter experience zero risk of progression and overall survival in line with the SMR adjusted general population mortality. Non-long-term survivors continue to follow the hazard of progression and death based on the curves fitted to the ZUMA-5 data. The company base case assumes 25% of axi-cel treated patients are long term survivors and applies these extrapolation assumptions from 5 years. Costs and utility values are derived from various sources.

Table 1 presents a summary of the key issues identified by the ERG.

Table 1 Summary of key issues identified by the ERG

Issues	Summary of issue	Report sections
Issue 1	Differences between the ZUMA-5 and SCHOLAR-5 cohorts in term of prior treatment received by SCHOLAR-5 patients	Section 3.3 and 3.6
Issue 2	The proportion of patients who can be considered long term survivors following treatment with axi-cel	Section 4.2.6
Issue 3	The PFS and OS extrapolation assumptions for axi-cel non-long-term survivors	Section 4.2.6
Issue 4	Health state utility values applied in the model	Section 4.2.7
Issue 5	The capping of time on treatment for comparator therapies, and modelling subsequent treatment costs	Section 4.2.8

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are the modelling of OS for non-long-term survivors, the modelling time on treatment for current 4L+ therapies and subsequent treatment costs, and the source of utility values applied.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Delaying/preventing progression of disease and increasing overall survival compared to current 4L+ care for patients with r/r FL.

Overall, the technology is modelled to affect costs by:

- Having higher acquisition costs compared to other available treatments
- Delaying or preventing progression of disease which incurs further subsequent treatment costs
- A higher modelled rate of adverse events compared to current care

- Extending expected survival time in the pre- and post-progression health states, which increases health state monitoring costs.

The modelling assumptions that have the greatest effect on the ICER are:

- The size of the overall survival benefit, which is determined by:
  - the parametric curve selection for OS in the technology and comparator arm of the model
  - the assumed proportion of patients that can be considered long-term survivors following treatment with axi-cel.
  - The OS extrapolation assumptions applied to axi-cel long-term survivors and non-long-term survivors
- The capping of time on treatment for current comparator therapies on overall survival rather than progression free survival. This assumption also affects the subsequent treatment costs applied in the model.

1.3 The decision problem: summary of the ERG's key issues

In general, the company decision problem is in line with the NICE final scope and no major issues were identified by the ERG. The CS addresses a more specific population than that specified in the NICE final scope and focuses on follicular lymphoma (FL), a subtype of indolent non-Hodgkin lymphoma, and specifically on FL patients who have received three or more prior lines of therapy (4L+ patients). The ERG in consultation with their clinical expert considers the company's description of the current treatment pathway and treatment options available for people with relapsed or refractory FL (r/r FL) accurate and agrees with the company's positioning of axi-cel in the treatment pathway

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

- The main source of clinical effectiveness evidence for axi-cel consists of the ongoing ZUMA-5., single-arm trial. The sample sizes of the analysis cohorts are generally small.

- Data from ZUMA-5 are immature with [redacted] within the current 18-month follow-up analysis.

- Some patients in SCHOLAR-5 received treatments that are not aligned with clinical practice in England, including idelalisib, which is accepted for use within NHS Scotland for the treatment of adults with FL refractory to 2 prior lines of therapy.

Issue 1 Comparability of ZUMA-5 with SCHOLAR-5 data

Report section	Section 3.3 and 3.6
Description of issue and why the ERG has identified it as important	Differences between the ZUMA-5 and SCHOLAR-5 cohorts in terms of prior treatment received by SCHOLAR-5 patients, and generalisability of SCHOLAR-5 to the NHS in England.
What alternative approach has the ERG suggested?	The ERG is not able to suggest an alternative approach, but the lack of randomised evidence leads to uncertainty in the magnitude of progression-free and overall survival benefit that can be expected with axi-cel versus currently available 4L+ treatments. There is also some uncertainty about how applicable the SCHOLAR-5 data are to the NHS in England, as a significant proportion of patients received treatments not routinely available or used in the NHS. However, on balance, the ERG believes this latter issue may bias against axi-cel.
What is the expected effect on the cost-effectiveness estimates?	Uncertainty relating to the magnitude of PFS and OS benefits, driven by the lack of randomised evidence, translates into uncertainty in the economic case.
What additional evidence or analyses might help to resolve this key issue?	The ERG judges this to be unresolvable uncertainty given the available evidence to inform comparative effectiveness.

1.5 The cost-effectiveness evidence: summary of the ERG's key issues

The ERG identifies the following key issues and uncertainties in the company's economic case:

Issue 2 The proportion of patients who can be considered long term survivors following treatment with axi-cel

Report section	Section 4.2.6
Description of issue and why the ERG has identified it as important	Whilst plausible based on previous experience with CAR T-Cell therapies in haematological cancer, the company's long-term survivor assumptions remain uncertain given the immaturity of the PFS and OS data from ZUMA-5. Whilst the ERG accept that it is plausible to expect a proportion of axi-cel treated patients to achieve long-term survivor status, there are no data available to estimate the proportion to which this assumption should apply. There is further uncertainty around the mortality hazard that long-term survivors might be able to achieve relative to the age and sex-matched general population. A standardised mortality ratio 1.09 is applied in the company base case.
What alternative approach has the ERG suggested?	The company acknowledge the current uncertainties and have conducted a sensitivity analysis to explore the impact of uncertainty around this issue. The ERG accept the company's base case long-term survivor proportion and timing of implementation (5 years) in its own base case but believe that scenario analyses around these inputs should be considered carefully by the committee.
What is the expected effect on the cost-effectiveness estimates?	Reducing the long-term survivor proportion has an upward impact on the ICER substantially, as well as applying it from a later timepoint. Increasing the proportion to which it applies, or applying it from an earlier time point, reduces the ICER.
What additional evidence or analyses might help to resolve this key issue?	The company have indicated that they will update their PFS and OS model inputs using an updated data cut from ZUMA-5 during technical engagement. This may help to better inform the shape of the time to event distributions. However, as the additional follow-up time will be limited, it is likely that the long-term survivor proportion will remain a key area of uncertainty. The company have provided the functionality in their model to address this.

Issue 3 The PFS and OS extrapolation assumptions for axi-cel non-long-term survivors

Report section	Section 4.2.6
Description of issue and why the ERG has identified it as important	Related to issue 2 above, the company fit parametric curves to the PFS and OS data of the overall subgroup of ZUMA-5 that matches the proposed positioning. However, they assume 25% of these patients achieve a reduced hazard of mortality in line with SMR adjusted general population mortality. From 5 years they use the fitted PFS and OS curves to model the hazard of progression and death only for non-long-term survivors. Since the curves were fitted for the whole patient population, which we assume includes patients achieving long term survivorship, the ERG believes this approach may underestimate the hazard of progression and death for non-long-term survivors. Adding to the uncertainty, the proportion of the surviving model cohort that are considered long-term survivors is fixed over time in the model. In reality, it should be increasing as non-long-term survivors face a higher risk of death. These issues lead to uncertainty with respect to the extrapolated survival gains in the progression-free and progressed model health states.
What alternative approach has the ERG suggested?	The ERG requested scenarios to explore these uncertainties at the clarification stage, which the company provided by applying SMR adjustments (of 1.09 and 1.2) to inflate the hazards of death and progression in non-long-term survivors from 5 years onwards. The ERG has extended the range of SMR adjustments applied in chapter 6 of this report. The company also provided an adjustment to allow the proportional split of the surviving cohort, between long-term and non-long-term survivors, to update over time.
What is the expected effect on the cost-effectiveness estimates?	Inflating the risk of progression and death in non-long-term survivors results in increases in the ICER. Allowing the proportion of survivors who are long-term/non-long-term survivors to update over time, in line with the separate hazards applied, produces reductions in the ICER.
What additional evidence or analyses might help to resolve this key issue?	The ERG does not believe it will be possible to resolve this issue through additional evidence or analysis. However, it remains an area of uncertainty that should be considered.

Issue 4 Health state utility values

Report section	4.2.7 Health-related quality of life
Description of issue and why the ERG has identified it as important	There is a lack of robust utility data available in the relevant patient population who would be eligible to receive axi-cel in practice. A literature search identified some potentially relevant studies, but instead, the company used a similar approach to that accepted in TA627 based on utility data collected in the AUGMENT study with values capped at population norms. The ERG is concerned that as the majority of patients in the AUGMENT study are at an earlier stage in the disease pathway, they would be expected to have a higher quality of life than patients receiving treatment at fourth line onward.
What alternative approach has the ERG suggested?	Alternative utility values were identified in a UK study reported by Wild et al., where EQ-5D data were collected in r/r FL patients. While there are also limitations with this study, the utility values are lower than those in AUGMENT and may better reflect the quality of life of patients at this stage of the treatment pathway. These values have also been used in other relevant NICE appraisals of FL treatments (TA604)
What is the expected effect on the cost-effectiveness estimates?	Wild et al. utility values were used in the sensitivity analysis, and this had a small upward impact on the ICER. Using these values increases face validity but does not resolve the uncertainty associated with a lack of robust quality of life data in this patient group.
What additional evidence or analyses might help to resolve this key issue?	The ERG agrees with the company that this remains a key source of uncertainty in the analysis due to the absence of robust data available. A range of alternative utility values, from other sources, were used in sensitivity analysis, all with minimal impact on the results. Further clinical validation of the Wild et al utility values would be useful.

Issue 5 The capping of time on treatment for comparator therapies, and modelling subsequent treatment costs

Report section	Section 4.2.8
Description of issue and why the ERG has identified it as important	For therapies included in the basket of current 4L+ care, the company utilise median numbers of treatment cycles reported in relevant summaries of product characteristics and then fit exponential distributions to estimate time on comparator treatments. However, the time on treatment curves are not necessarily consistent with the derived PFS and OS curves for the comparator arm in this indication (r/r/ FL at 4L+), and in their base case the company cap time on treatment (ToT) so it can't exceed overall survival. This assumes that treatment can continue beyond progression. Furthermore, the company then recycle the mean 4L+ treatment acquisition and administration costs derived for the comparator arm and apply this as a one-off cost of subsequent treatment to the estimated proportion of the cohort that progresses in each cycle of the model. This method would appear to overestimate comparator therapy costs which, based on clinical advice to the ERG, would be stopped upon progression. It may also overestimate subsequent treatment costs. A further uncertainty relates to the assumption that subsequent treatment costs, per progressed patient, are assumed equal between the treatment arms. There may be potential for subsequent treatment costs to be higher in the progressed disease state for those treated with axi-cel, as these patients may have more treatment options left available and may respond for longer.
What alternative approach has the ERG suggested?	In their original submission, the company also provided a scenario analysis whereby they capped time on treatment for comparator therapies to PFS rather than OS. The ERG is of the opinion that the latter assumption is more appropriate based on its clinical advice. It is also more consistent with the assumption that all patients who progress receive subsequent treatment costs in line with the modelled 4L+ comparator costs. The ERG has also assessed the impact of reducing subsequent treatment costs in the current 4L+ care arm relative to those applied in the axi-cel arm.
What is the expected effect on the cost-effectiveness estimates?	The change results in modest increases in the ICER.
What additional evidence or analyses might help to resolve this key issue?	This remains an area of uncertainty for which it is difficult to identify alternative data. The company note that ToT data was not reported for SCHOLAR-5. If such data could be obtained, it could help to resolve the above uncertainty. Alternatively, further clinical opinion could be sought on the suitability of using PFS rather than OS to cap time on treatment for current 4L+ therapies used in NHS England.

1.6 Other key issues: summary of the ERG's view

The company believe the axi-cel may be suitable for consideration on the cancer drug fund. They also argue that it will be used as an end-of-life medicine in this indication. However, both median overall survival and modelled life expectancy in the comparator arm are [redacted] (see chapter 7).

1.7 Summary of ERG's preferred assumptions and resulting ICER

Given the uncertainties outline above, and other issues raised in the report, the ERG prefers to:

1) Apply the company's scenario switch which allows the proportional split of the surviving cohort, between long-term survivors and non-long-term survivors, to be updated on a cycle-by-cycle basis from the time that the long-term survivor assumptions are applied (5 years).
2) Inflate the hazard of progression and death by 1.2 in non-long-term survivors from the time the long-term survivor assumption is applied (5 years).
3) Cap overall survival of non-long-term survivors at SMR adjusted general population mortality, to ensure the risk of death in non-long-term survivors is never lower than that in long-term survivors.
4) Cap the current 4L+ time on treatment to the selected PFS curve for current 4L+ care, rather than the selected OS curve.
5) Apply alternative Wild et al./Pettengell et al. utility values for progression-free and progressive disease states that are available from the literature.
6) Retain the preferred progression-free health state utility for long-term survivors from 5 years, rather than assuming general population utility.

Further scenario analysis around the ERG base case explores the impact of: alternative PFS and OS curve selections; alternative adjustments to the risks of progression or death in non-long-term survivors; relative reductions in the costs of subsequent therapy following progression on current 4L+ care; and changes to the long-term survivor proportion (see section 6.3).

Table 2 Summary of ERG's preferred assumptions and ICER

Preferred assumption	Incremental cost	Incremental QALY	ICER £/QALY	Change from company base case
Company base-case	[redacted]	[redacted]	£48,272	NA
1. Time dependent updating of long-term survivor proportion from 5 years	[redacted]	[redacted]	£46,105	-£2,168
2. Increase progression and mortality risks by 20% after 5 years non-long-term survivors	[redacted]	[redacted]	£52,326	£4,054
3. Cap overall survival of non-long-term survivors at SMR adjusted general population mortality	[redacted]	[redacted]	£48,354	£82
4. Capping the current 4L+ time on treatment to the selected PFS curve for current 4L+ care	[redacted]	[redacted]	£54,163	£5,891
5. Apply Wild et al/Pettengell et al. utility values for progression free and progressive disease states.	[redacted]	[redacted]	£49,296	£1,024
6. Retain PF health state utility from Wilde et al. for long-term survivors (only relevant with 5 above)	[redacted]	[redacted]	£49,993	£1,721
Combined changes (ERG base case)	[redacted]	[redacted]	£56,332	£8,060

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

The relevant health condition for the submission received from Kite is relapsed or refractory low-grade non-Hodgkin lymphoma in adults. The company's description of this health condition in terms of prevalence, symptoms and complications appears generally accurate and in line with the decision problem. The relevant intervention for this submission is axicabtagene ciloleucel (axi-cel).

2.2 Background

The company submission (CS) describes non-Hodgkin lymphoma (NHL) as diverse group of cancers that originate in the lymphatic system. The CS focuses on follicular lymphoma (FL) as the most common type of indolent (slow-growing) NHL (iNHL).[1] FL mainly affects people aged over 60 years and, while, it is associated with longer survival times, it is less likely to be cured than faster-growing lymphomas and is associated with reduced life expectancy and impaired health-related quality of life (HRQoL) compared to the general population.[2]

FL has an annual incidence of approximately 3.3 per 100,000 people, and it is estimated that around 2,200 people are diagnosed with FL each year in the UK.[3] The 10-year prevalence is 24.7 per 100,000 people, and it is estimated that 16,220 people in the UK will have been diagnosed with FL during the last 10 years.[3] The most common physical symptom of FL is a painless swelling in the neck, armpit, or groin, caused by enlarged lymph nodes.[4] FL is also associated with 'B-symptoms' such as night sweats, erratic fever, weight loss, and unexplained itching.[4] Patients with FL presenting with multiple sites of lymphadenopathy can endure restricted movement, disfigurement, pain, and bone marrow disease that can result in anaemia, leukopenia, and thrombocytopenia.[5, 6] FL is also associated with poorer mental health, with patients experiencing depression and stress, as well as the emotional upset of living with a chronic disease that is incurable and will progress.[7-9] HRQoL is further affected by treatment toxicity effects, and HRQoL is likely to deteriorate with each treatment relapse. Patients with relapsed FL are more likely to experience lower physical, emotional, functional, and social wellbeing HRQoL scores and higher levels of anxiety, depression and activity impairment levels compared with disease-free patients.[5, 10] The burden of illness in patients

with three or more lines of systemic therapy is, therefore, expected to be particularly high. FL is also associated with a high carer burden. In a Canadian cross-sectional cohort of patients with iNHL, including FL, most of the care (74%) was unpaid assistance from a partner or spouse, relative or friend.[11] Carers in the study provided a mean of 9.8 (SD 13.4) days of care in the 30 days prior to data collection and missed a mean of 11.3 (SD 16.2) days of work because of the care they provided.

Treatment decisions for FL are based on several factors, including the stage and grade of the disease, and risk categorisation based on demographic and basic disease characteristics. The company provides a summary of the classifications systems for FL in Table 3, document B of the CS and this is reproduced by the ERG as Table 3.

Table 3 Classification systems for follicular lymphoma

<table>
<tr><th>WHO/REAL</th><th>Cotswolds modified Ann Arbor</th><th>FLIPI score</th></tr>
<tr><td rowspan="2">Grade 1: 0–5 centroblasts
Grade 2: 6–15 centroblasts
Grade 3: >15 centroblasts

Grade 3B: absence of centrocytes</td><td rowspan="2">Stage I: single lymph node group or organ
Stage II: multiple lymph node groups/organ on same side of diaphragm
Stage III: multiple lymph node groups/organ on both sides of diaphragm
Stage IV: bone marrow or distant organ involvement.
Stage X: bulky disease with nodal mass >10 cm
Stage E: extra-nodal extension or single isolated site of extra-nodal disease
Stage A/B: absence or presence of symptoms – B-symptoms include weight loss >10%, fever, drenching night sweats</td><td>Factors (1 point for each variable present):
• Age >60 years
• Ann Arbor Stage III–IV
• Haemoglobin level <12 g/dl
• LDH level >ULN
• ≥4 nodal sites of disease</td></tr>
<tr><td>Risk category (factors):
• Low (0–1)
• Intermediate (2)
• High (3–5)</td></tr>
<tr><td colspan="3">Key: FLIPI, Follicular Lymphoma International Prognostic Index; LDH, lactate dehydrogenase; REAL, Revised European-American Lymphoma; ULN, upper limit of normal; WHO, World Health Organization.
Source: Hernandez-Ilizaliturri 2020.[12]</td></tr>
</table>

The aim of treatment is usually to keep the disease in remission for as long as possible.[13] Remission may last for several years, but approximately 10-20% of FL patients will experience multiple relapses. The time spent in remission usually shortens with each successive relapse as the disease becomes more resistant to treatment (known as treatment refractoriness), thus reducing the patient's overall lifespan. High-risk sub-populations include patients who are chemoimmunotherapy resistant and fail to achieve a response within six months of completing initial chemoimmunotherapy and patients who have double-refractory

disease (patients who are refractory to the first two lines of therapy, including both an alkylating agent and an anti-CD20 monoclonal antibody). People who experience progression of disease within two years of receiving front-line chemoimmunotherapy (defined as 'POD24') have a particularly poor prognosis, with only a 50% overall survival (OS) estimate at five years, compared with 90% OS estimate at five years for people without POD24.[14-16] Around 10% of people diagnosed with FL in the UK will receive four or more lines of therapy. This is approximately 220 patients, around 198 of whom will receive their treatment in England or Wales.[10] The survival prognosis of patients with relapsed or refractory FL who have had ≥4 lines of therapy is generally poor.

While there are several guidelines for the treatment of symptomatic advanced-stage FL, there is no consensus on treatment or standard of care for patients beyond the third line of treatment.[17-22] These patients typically follow an aggressive, chemotherapy-resistant disease course, with poor prognosis. By the time patients have received three or more lines of prior therapy (4L+), patients will usually have received multiple rituximab-based regimens and are, therefore, expected to have suboptimal response to further rituximab-based treatment. In the absence of an established standard of care, current 4L+ therapy consists of recycling earlier-line treatment options or resorting to generic haemato-oncology or experimental/compassionate use treatments. Treatment decisions are made on a case-by-case basis, considering factors such as patient fitness, treatment goals, response, and durability of response to prior therapy.

The proposed place of axi-cel in the treatment pathway is presented in Document B, Figure 3 of the CS and is reproduced below as Figure 1. The ERG notes that the NICE Pathways service has been withdrawn since the company accessed the treatment pathway in August 2021. The company clarified that eligibility for axi-cel is not expected to differ depending on the stage of disease, and will not differ, irrespective of the route the patient has taken to reach 4L/4L+ treatment, where the treatment goal is to achieve sustained clinical remission. The ERG agrees that the company's proposed pathway is representative of current clinical practice and the anticipated positioning of axi-cel is within its licensed indication.

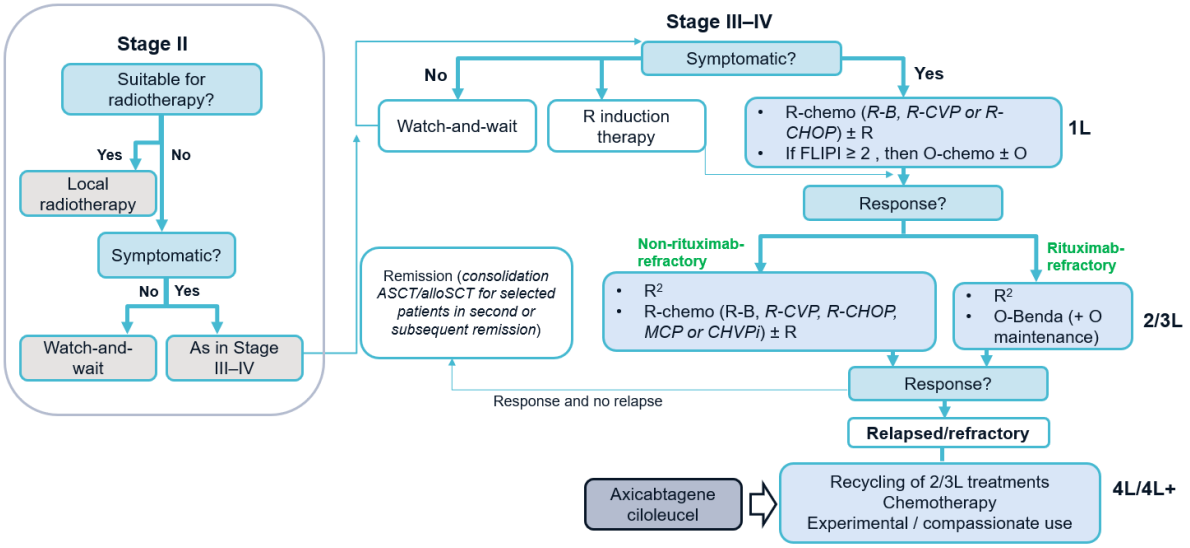

Figure 1 Clinical care pathway for patients with follicular lymphoma and proposed axi-cel positioning

Key: 1L, first-line; 2L, second-line; 4L, fourth-line; alloSCT, allogeneic stem cell transplantation; ASCT, autologous stem cell transplant; Benda, bendamustine; CHOP, cyclophosphamide, doxorubicin, vincristine and prednisolone; CHVPi, cyclophosphamide, doxorubicin, etoposide, prednisolone and interferon-α; CVP, cyclophosphamide, vincristine and prednisolone; FLIPI, Follicular Lymphoma International Prognostic Index; MCP, mitoxantrone, chlorambucil and prednisolone; NICE, National Institute for Health and Care Excellence; O, obinutuzumab; R, rituximab; R-B, rituximab with bendamustine; R^2, lenalidomide with rituximab.
Source: NICE Pathways – Treating follicular lymphoma[23]

2.3 Critique of company's definition of decision problem

A summary of the company's decision problem in relation to the NICE final scope is presented in Table 4 below. A critique of adherence of the company's economic modelling to the NICE reference case is presented in Chapter 4.

Table 4 Summary of the company's decision problem

	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 relapsed or refractory non-Hodgkin lymphoma	[redacted]	The anticipated marketing authorisation for axicabtagene ciloleucel is for the treatment of '[redacted] As such, this submission is focused on FL, a subtype of indolent non-Hodgkin lymphoma, and specifically on FL patients who have received three or more prior lines of therapy (4L+ patients)	The ERG agrees that the population addressed in the CS is appropriate for this appraisal
Intervention	Axicabtagene ciloleucel	Axicabtagene ciloleucel	Not applicable	The intervention described in the CS matches that described in the NICE final scope. Axicabtagene ciloleucel has a marketing authorisation for treating relapsed or refractory diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma, after 2 or more lines of systemic therapy. A FL variation was submitted to the EMA on 23 July 2021. CHMP opinion is expected in April 2022. The application for GB filing will be submitted in April 2022 for a marketing authorisation extension of axi-cel (Yescarta) to [redacted] The anticipated date of marketing authorisation for this indication is [redacted]

Comparator(s)	• Rituximab monotherapy • Rituximab in combination with chemotherapy • Obinutuzumab with bendamustine • Lenalidomide with rituximab • Clinical management	• Rituximab in combination with chemotherapy • Obinutuzumab with bendamustine • Lenalidomide with rituximab • Clinical management without axicabtagene ciloleucel including chemotherapy (such as cyclophosphamide, fludarabine, bendamustine or chlorambucil)	Rituximab monotherapy is only recommended as an option for the treatment of r/r FL when all alternative treatments have been exhausted (that is, if there is resistance to or intolerance of chemotherapy). If it was being considered for use in patients with r/r FL after three or more lines of systemic therapy, it would be reserved for patients not fit enough to receive intensive active treatment as is the case for best supportive care, thereby constituting a cohort of patients widely considered not suitable or appropriate for consideration of CAR T-cell therapy. Indeed, clinical experts note that by the time patients reach the 4L+ treatment setting, they will have received rituximab monotherapy multiple times and, thereby, additional rituximab monotherapy would most likely be ineffective in this setting.[24] Neither rituximab monotherapy nor best supportive care are therefore relevant comparators for patients being considered for axicabtagene ciloleucel Of the other comparators listed, we would expect obinutuzumab with bendamustine and lenalidomide with rituximab to typically be used earlier in the treatment pathway than the 4L+ treatment setting. In addition, we would expect that chemotherapy (clinical management without axicabtagene ciloleucel) would be used after the 4L+ setting, following approval of axicabtagene ciloleucel.	The ERG clinical expert agrees that the company's choice of comparators is appropriate for this appraisal.

	ent without axicabtagene ciloleucel including chemotherapy (such as cyclophosphamide, fludarabine, bendamustine or chlorambucil) • Best supportive care		However, we have considered these as part of a blended comparator representing current care in the decision problem addressed.	
Outcomes	• Overall sur	• Overall survival • Progression-free survival • Response rates	Health-related quality of life data were not collected in ZUMA-5 and are therefore informed by the existing literature base	The outcomes reported in the CS match the NICE final scope. The ERG clinical expert considers the outcomes to be appropriate

	viv al • Pro gre ssi on-fre e sur viv al • Res pon se rate s • Ad ver se eff ect s of trea tme nt • He alth - rela ted qua lity of life	• Adverse effects of treatment • Health-related quality of life		for addressing the topic of this appraisal
Eco nom ic anal ysis	The referen ce case stipula tes that the cost-effecti veness of treatm ents should be expres sed in	Cost-effectiveness is expressed in terms of incremental cost per QALY. The time horizon is set at 40 years; sufficient to capture the plausible maximum life expectancy for the population modelled (who have a mean age of [redacted] years at model entry).	Not applicable	

	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 relate to NHS and PSS resources and are valued using the prices relevant to the NHS and PSS. The cost year of the analysis is 2019/20, though the latest available drug prices were used whenever possible using MIMS UK and eMIT databases.		

	Costs will be considered from an NHS and Personal Social Services perspective.			
Subgroups	No subgroups were specified in the NICE final scope			The company presents subgroup analyses for objective and complete response rates for baseline and treatment characteristics in Appendix E of the CS; however, the analyses are for FL patients with ≥2 lines of prior therapy and are, therefore, not the relevant patient population for this appraisal.
Special considerations including issues related to equity or equality	The availability and cost of biosimilar and generic products should be taken into account. Guidance			The ERG agrees with the company that there are no foreseen equality issues with axi-cel. The CS states that there are existing inequalities in current non-immunochemotherapy treatment options available in England compared with Wales and Scotland where idelalisib (a licensed Pi3Kδ inhibitor) is available through routine baseline commissioning to patients who have refractory FL after two prior lines of treatment.

	will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation			

	granted by the regulator			

Key: 4L+, fourth-line plus (three or more lines of prior therapy); CAR T-cell, chimeric antigen receptor T-cell; FL, follicular lymphoma; NICE, National Institute for Health and Care Excellence; r/r, relapsed or refractory.

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

Full details of the methods used to identify and select the clinical evidence relevant to this appraisal are reported in Appendix D of the CS. The ERG'S appraisal of the company's systematic review methods is summarised in Table 5.

Table 5 ERG's appraisal of the systematic review methods presented in the CS

Review process ERG	ERG response	Comments
Were appropriate searches (e.g., search terms, search dates) performed to identify all relevant clinical and safety studies?	Yes	The CS provides full details of the searches used to identify the studies for the clinical effectiveness review. The search strategies include relevant controlled vocabulary and text terms with appropriate use of Boolean operators and are fully reproducible. Details provided in Appendix D of the CS.
Were appropriate bibliographic databases/sources searched?	Yes	Sources included Embase, Medline, and CENTRAL for primary research. Relevant conference proceedings and trial registers were also searched. Full details are provided in Appendix D of the CS.
Were eligibility criteria consistent with the decision problem outlined in the NICE final scope?	Yes	Searches were not restricted by any eligibility criteria so all results were discovered and only those relevant to the scope were selected.
Was study selection conducted by two or more reviewers independently?	Yes	Appendix D, section D.1.3.1: *"At primary screening, all abstracts were assessed against pre-defined eligibility criteria (**Error! Reference source not found.**) by two reviewers with any uncertainty resolved with a third independent reviewer. At secondary screening (full-text review) publications were independently assessed by two reviewers and discrepancies resolved by consulting a third*

		reviewer and on reaching consensus." Section D.1.3.2: *"The study selection methodology of the SLR update was aligned with the original SLR; screening (both primary and secondary) was performed by two independent reviewers, with discrepancies resolved with a third independent reviewer."*
Was data extraction conducted by two or more reviewers independently?	No	Appendix D, section D.1.3.1: *"Data extraction was performed by one researcher and validated by another independent researcher. Any disagreements were resolved by consulting with the third reviewer."* Section D.1.3.2: *"For data extraction, the template of the original SLR was used, with data extraction conducted by one reviewer, and quality checked against the original source by a second reviewer."*
Were appropriate criteria used to assess the risk of bias of identified studies?	Yes	Appendix D, section D.3.1: *"The quality of each RCT identified in the SLR was assessed using the Cochrane Risk of Bias Tool. Each RCT was rated as low risk, unclear risk or high risk of bias. The Downs and Black checklist was used to assess the bias in non-randomised studies."* Section D.3.2: *"In the SLR update, the quality assessment of the included non-RCTs was performed using the Downs and Black checklist including an assessment of the ZUMA-5 study".* The ERG considers the company's assessments to be appropriate
Was the risk of bias assessment conducted by two or more reviewers independently?	No	At clarification: *"For the SLR, each study that met the criteria for inclusion was critically appraised by a single reviewer and reviewed by a second reviewer using the Cochrane Collaboration's tool for*

		assessing the risk of bias, in line with NICE requirements. *Similarly, in the SLR update, quality assessment of the included studies was performed as part of the data extraction process, i.e., each checklist item was extracted from the included full-text articles by one reviewer, and quality checked against the original source by a second reviewer."* The ERG considers the company's strategy to be satisfactory
Was identified evidence synthesised using appropriate methods?	Yes	The company did not conduct a meta-analysis or a NMA but they compared the outcomes of ZUMA-5 with those of SCHOLAR-5 which is an external cohort study. To account for imbalances between the populations in the two studies they used propensity scoring methods, specifically standardised mortality ratio weighting. Although it was not transparent how this was performed, the ERG felt the weighting has improved comparability between the ZUMA-5 and SCHOLAR-5.

The CS reports that 16 studies reporting data for the 4L+r/r FL setting were identified by the original SLR and further studies were also identified in the SLR update. Section D.5 of the CS Appendices states: "*While some data was identified from the SLRs for the 4L+ r/r FL reporting on current treatment options for this setting, the strength of evidence was insufficient to enable robust treatment comparisons of this data with the ZUMA-5 study. There were several reasons for this including:*

- *The low availability of evidence specific for 4L+ r/r FL*
- *The scarcity of RCTs and other types of controlled study designs, which increases the risk of bias in effect estimates and challenges an assessment of comparative effectiveness*

- *The small sample sizes, which increase uncertainty around estimates of the treatment effect*
- *The considerable heterogeneity in patient characteristics and clinical endpoints, making reliable inter-study comparisons difficult*

As such, no literature-based treatment comparisons were conducted for ZUMA-5. Instead, the international, multicentre, external control cohort study, SCHOLAR-5 was used to provide a synthetic control arm for ZUMA-5 and comparative analyses were conducted for r/r FL patients meeting ZUMA-5 eligibility criteria."

At clarification, the ERG requested reasons for each individual study being unsuitable for a comparison with ZUMA-5. The company responded: "*ZUMA-5 was a single-arm study because there is no standard of care (SoC) for this population. As a single-arm study, direct comparison to a comparison arm was not possible. Patients [with r/r FL] typically receive salvage therapy or potentially allogeneic SCT, but the exact nature and outcome varies greatly depending on patient characteristics including age, disease stage, tumour burden, and the number of prior lines of therapy. This may lead to potential bias when carrying out indirect comparisons of results from published studies. Therefore, to further determine the clinical benefit associated with CAR-T therapy, an accurate detailed description of available treatment options in the relevant patient population and associated outcomes was required. In the absence of comparable data, Kite Pharma constructed an external cohort of real-world FL (grades 1-3A) patients who would be eligible for ZUMA-5. This real-world cohort was used as an external control for the ZUMA-5 clinical trial.*"

In addition, the CS states that its SLR identified three studies in the grey literature that reported potentially relevant comparative efficacy data: Batlevi 2020, Link 2019, Fuji 2020.[25-27] The company's justification for not using these studies was that none reported baseline characteristics for the relevant population, and none were conducted in Europe (two in USA: Batlevi 2020, Link 2019; one in Japan: Fuji, 2020). The ERG agrees with these assertions. However, overall, it is unclear to the ERG whether it was appropriate for the company to choose to not use any of the identified evidence in this appraisal. The ERG conducted a quality assessment of the methods used by the company for the systematic review of clinical evidence using the Centre for Review and Dissemination (CRD) criteria.[28] The results are presented in Table 6.

Table 6 Quality assessment of the company's systematic review of clinical effectiveness evidence

CRD quality item	Yes/No/Unclear
1. Are any inclusion/exclusion criteria reported relating to the primary studies, which address the review question?	Yes
2. Is there evidence of a substantial effort to search for all of the relevant research?	Yes
3. Is the validity of included studies adequately assessed?	Yes
4. Are sufficient details of the individual studies presented?	Yes
5. Are the primary studies summarised appropriately?	Yes

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

3.2.1 Included studies

Details of the key clinical effectiveness evidence are provided in Document B, Section B.2 of the CS. The company presents clinical effectiveness evidence from one on-going, single-arm, multicentre, open-label phase II trial, ZUMA-5. Details of the trial are summarised in Document B, Table 5 of the CS and reproduced in Table 7 below. The methods of ZUMA-5 are reported in Document B, Section 2.3 of the CS and the participant flow is reported in Document B, Section 2.4.1 of the CS. The objective of ZUMA-5 was to evaluate the efficacy of axi-cel, as measured by overall response rate (ORR), in people with relapsed or refractory follicular lymphoma (r/r FL) or marginal zone lymphoma (MZL). The CS states that the focus of the submission is on participants with r/r FL who had already received three or more lines of prior therapy, albeit reporting also baseline and outcome data for participants who had received two or more lines of prior therapy. ZUMA-5 was conducted at 15 sites in the USA and two in France.

The key eligibility criteria for ZUMA-5 are reported in Document B, Section B.2.3, Table 6 of the CS. The study schema for ZUMA-5 is presented in Document B, Section B.2.3, Figure 4 of the CS and is reproduced as Figure 2 below.

Table 7 Summary of clinical effectiveness evidence [reproduced from Table 5, Document B of the CS]

Study (NCT)	ZUMA-5 (NCT03105336)				
Study design	ZUMA-5 is an ongoing Phase II, multicentre, open-label study evaluating the efficacy and safety of axi-cel in r/r iNHL.				
Population	Adult subjects with r/r B-cell iNHL of FL or MZL histological subtypes who have received 2 or more prior lines of therapy. The FL cohort of patients who have received three or more lines of prior therapy is the focus of this submission.				
Intervention(s)	Axi-cel				
Comparator(s)	Not applicable				
Indicate if trial supports application for marketing authorisation	Yes	✓	**Indicate if trial used in the economic model**	Yes	✓
	No			No	
Rationale for use/non-use in the model	ZUMA-5 presents the pivotal, regulatory, clinical evidence in support of axi-cel in r/r FL				
Reported outcomes specified in the decision problem	• Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life				
All other reported outcomes	• Incidence of anti-CD19 CAR antibodies • Levels of anti-CD19 CAR T-cells in blood • Levels of cytokines in serum				
Key: FL, follicular lymphoma; iNHL, indolent non-Hodgkin lymphoma; MZL, marginal zone lymphoma; r/r, relapsed/refractory. **Notes:** bolded outcomes are those used in the economic modelling. No outcomes were bolded in Table 5 of the CS. Table 24 of the CS states that clinical parameters of the model were PFS, OS and AE incidence.					

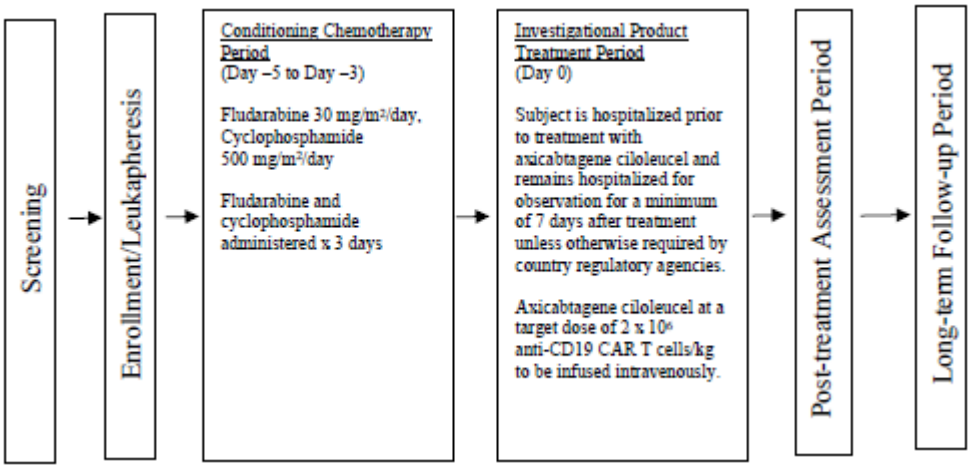

Figure 2 Study scheme for ZUMA-5

Key: CAR, chimeric antigen receptor. **Source:** ZUMA-5 Clinical Study Protocol.

The company assessed risk of bias of ZUMA-5 using the Downs and Black checklist. In general, the ERG agrees with the company's assessment of the study and that the overall risk of bias is low, in the context of a single-arm study, albeit with the bias inherent in non-randomised studies. In addition, ZUMA-5 was funded by Kite, a Gilead company, which declared a role in study design, data collection, data analysis and data interpretation. Details of the baseline characteristics of the full analysis set (FAS), safety analysis set (SAS; also referred to in the CS as the modified ITT [mITT] population for efficacy analyses) and inferential analysis set (IAS) of participants with two or more and three or more lines of prior therapy are presented in Document B, Section B.2.3.1, Table 7 of the CS. The company provided an amended version of the table at clarification, an adapted version of which, is presented as Table 8 below, reporting those participants with three or more lines of prior therapy.

Table 8 Baseline characteristics of participants in ZUMA-5 with ≥3 lines of prior therapy [adapted from Table 7 of company's clarification response]

Characteristics	FL patients with three or more lines of prior therapy		
	FAS (n = 80)	SAS (n = 78)	IAS (n = 60)
Median age, years (min-max range)	[redacted]	[redacted]	[redacted]
Aged ≥65 years, n (%)	[redacted]	[redacted]	[redacted]
Aged <65 years, n (%)	[redacted]	[redacted]	[redacted]
Male, n (%)	[redacted]	[redacted]	[redacted]
Female, n (%)	[redacted]	[redacted]	[redacted]
ECOG performance status, n (%)			
0	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]
FL histological category at trial entry, n (%)			
Grade 1	[redacted]	[redacted]	[redacted]
Grade 2	[redacted]	[redacted]	[redacted]
Grade 3a	[redacted]	[redacted]	[redacted]
FLIPI total score, n (%)			
Low risk (0–1)	[redacted]	[redacted]	[redacted]
Intermediate risk (2)	[redacted]	[redacted]	[redacted]
High risk (3–5)	[redacted]	[redacted]	[redacted]

Characteristics	FL patients with three or more lines of prior therapy		
	FAS (n = 80)	SAS (n = 78)	IAS (n = 60)
Relapsed/refractory disease[a], n (%)			
Relapsed	[redacted]	[redacted]	[redacted]
Refractory	[redacted]	[redacted]	[redacted]
Double-refractory subgroup[a], n (%)	[redacted]	[redacted]	[redacted]
Median no. of prior therapies (range)	[redacted]	[redacted]	[redacted]
Number of prior lines of therapy, n (%)			
1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]
≥5	[redacted]	[redacted]	[redacted]
Prior auto-SCT, n (%)	[redacted]	[redacted]	[redacted]
Prior PI3K inhibitor, n (%)	[redacted]	[redacted]	[redacted]
Prior anti-CD20 single agent, n (%)	[redacted]	[redacted]	[redacted]
Prior alkylating single agent, n (%)	[redacted]	[redacted]	[redacted]
Prior anti-CD20 + alkylating agent, n (%)	[redacted]	[redacted]	[redacted]
Time to relapse from first therapy[b], n (%)	[redacted]	[redacted]	[redacted]
≥24 months	[redacted]	[redacted]	[redacted]
<24 months	[redacted]	[redacted]	[redacted]
Prior lenalidomide, n (%)	[redacted]	[redacted]	[redacted]
Bone marrow assessment at baseline, n (%)[c]			
Lymphoma present	[redacted]	[redacted]	[redacted]
Lymphoma present but not FL	[redacted]	[redacted]	[redacted]
Lymphoma not present	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]

Key: auto-SCT, autologous stem cell transplantation; ECOG, Eastern Cooperative Oncology Group; FAS, full analysis set; FL, follicular lymphoma; IAS, inferential analysis set; PI3K, phosphoinositide 3-kinase
Notes: [a] Patients with FL who progressed within 6 months of completion of the most recent prior treatment are defined as refractory. Patients with FL who progressed >6 months of completion of the most recent prior treatment are defined as relapsed. Patients with FL who progressed within 6 months of completion each of the first 2 lines of prior treatment are defined as double-refractory. [b]Time to relapse is defined as the time from initiation of the first line anti-CD20-chemotherapy combination therapy to progression. Number of subjects with time to relapse is based on those who had progressed with date of progression. Percentages are based on the number of subjects who ever received anti-CD20-chemotherapy combination therapy. [c] bone marrow assessment at baseline for lymphoma presence is based on investigator reported Lugano bone marrow assessment/bone marrow assessment using aspirate or core biopsy at screening. If these are not available, lymphoma presence is based on diagnosis history of bone marrow involvement.

Source: ZUMA-5 CSR 18-Month Addendum.[29] *Total is not 100% due to rounding

The median age of participants was [redacted] years ([redacted] years in the IAS) and [redacted] of participants were males. [redacted] participants had ECOG scores of 0 than 1 but the difference was not substantial. [redacted] of the participants had Grade 2 FL, with the remaining participants being split quite evenly between Grades 1 and 3a. [redacted] of the participants were high risk, according to the FLIPI total score, and the [redacted] had refractory disease rather than relapsed. The median number of prior therapies was [redacted] and around [redacted] of participants had [redacted] prior therapies. Time to relapse from first anti-CD20-chemotherapy was [redacted] months in [redacted] of participants. The ERG's clinical expert notes that progression of disease within 24 months of initiating treatment is the strongest predictor of aggressive disease. In general, the ERG's clinical expert is of the opinion that the baseline characteristics of the participants in ZUMA-5 are representative of patients with r/r iNHL seen in clinical practice in the UK.

3.2.2 Primary and secondary efficacy endpoints

The outcome measures listed in the NICE final scope for this appraisal were: OS, progression-free survival (PFS), response rates, adverse effects and HRQoL.

Primary endpoint: ZUMA-5

The primary endpoint of ZUMA-5 was the ORR in patients with r/r FL with two or more lines of therapy who had the opportunity to be followed for at least 18 months from first disease assessment date following axi-cel treatment, defined as the incidence of participants achieving complete response (CR) or partial response (PR), as determined by independent central review per Lugano classification. Thus, the primary endpoint of ZUMA-5 is not relevant to this appraisal, the focus of which is patients with r/r FL with three or more lines of prior therapy.

Secondary endpoints: ZUMA-5

The secondary endpoints reported in the CS and relevant to this appraisal (i.e. for participants with three or more lines of prior therapy) are the following, reported in terms of the IAS (i.e. patients treated with any dose of axi-cel who had the opportunity to be followed up for at

least 18 months from first disease assessment date; [redacted] and post-hoc analyses of the mITT population (i.e. patients treated with any dose of axi-cel; [redacted]):

- **ORR** (defined as incidence of CR or PR by independent central review per Lugano classification in patients who had the opportunity to be followed for at least 18 months from first disease assessment date): ORR was [redacted] of the IAS ([redacted]; exact test for ORR ≤40%: [redacted]). ORR was [redacted] of the mITT population
- **CR** (defined as incidence of CR by independent central review per Lugano classification): [redacted] participants achieved a CR ([redacted]; exact test for CR[redacted]). The CR rate in the mITT population was [redacted]
- **Duration of response** (DOR; defined only for participants who achieved an OR and is the time from first objective response to disease progression or death, by central and investigator assessment): The median DOR was not reached in all responders: [redacted]. Median follow-up for DOR was [redacted]. [redacted] responders had an ongoing response at censoring. In the [redacted] participants with a CR, median DOR was not reached and [redacted] had an ongoing response at data cut-off. In the mITT population, median DOR in the [redacted] responders was [redacted]; the median follow-up time was [redacted]
- **Best objective response** (BOR; defined as incidence of CR, PR, stable disease (SD), progressive disease (PD) or non-evaluable (NE) as best response by the Lugano classification (by central read or investigator read)): in the IAS, CR was achieved by [redacted] PR achieved by [redacted]; stable disease achieved by [redacted]. The remaining [redacted] participants were classified as either "undefined/no disease" or "not done". The CS presents Kaplan-Meier plots for DOR and DOR by best response in Appendix L (IAS) and Document B, Figures 6 and 7 (mITT).

The CS presents summaries of response and duration of response data for the IAS and mITT in Document B, Tables 11 and 12 of the CS, adapted as Table 9 below.

Table 9 Summary of response using central assessment per Lugano classification; FL patients with three or more lines of prior therapy, IAS [adapted from Table 11 and Table 12, Document B, of the CS]

	IAS (N = [redacted]	mITT (N=[redacted])
Objective response rate (CR + PR), n (%) [95% CI]	[redacted]	[redacted]
p-value vs historical control rate	[redacted]	
Best objective response		
Complete response rate, n (%) [95% CI]	[redacted]	[redacted]
Partial response, n (%) [95% CI]	[redacted]	[redacted]
Stable disease, n (%) [95% CI]	[redacted]	[redacted]
Progressive disease, n (%) [95% CI]	[redacted]	[redacted]
Duration of response		
Median duration of response in all responders, months (range)	[redacted]	[redacted]
Median duration of response in CRs, months (range)	[redacted]	

Key: CI, confidence interval; CR, complete response; CSR, clinical study report; IAS, inferential analysis set; NE, not evaluable; PR, partial response.
Source: ZUMA-5 CSR 18-Month Addendum.

The CS reported further outcomes in terms of the IAS and mITT:

- **PFS** (defined as the time from date of axi-cel infusion to date of disease progression per Lugano assessment or death due to any cause): in the IAS, [redacted] had progressed and [redacted] had died at the time of analysis; thus, [redacted] were alive and progression-free. Median PFS [redacted] Median follow-up time for PFS was [redacted]. Estimated PFS rates at months 12 and 18 were [redacted], respectively. In the mITT population, median PFS was [redacted] ([redacted]), with a median follow-up of [redacted] months. A total of [redacted] participants had progressed

or died at the time of analysis. Estimated PFS rates at months 12 and 18 were [redacted] respectively. The CS presents Kaplan-Meier plots for PFS in the IAS (Appendix L) and the mITT (Document B, Figure 8).

- **OS** (defined as time from axi-cel infusion to date of death due to any cause): [redacted] patients had died at the time of analysis and [redacted] were alive. Median OS was not reached [redacted] Median follow-up time for OS was [redacted] Estimated OS rates at months 12 and 18 were [redacted] and [redacted], respectively. In the mITT population, median OS was not reached [redacted], with a median follow-up of [redacted] months. Estimated OS rates at months 12 and 18 were [redacted] and [redacted], respectively. The CS presents Kaplan-Meier plots for OS in the IAS (Appendix L) and mITT (Document B, Figure 10).

A summary of PFS and OS outcomes is presented in Table 10 below.

Table 10 Summary of PFS and OS outcomes for IAS and mITT populations

	IAS (N = [redacted])	mITT (N=[redacted])
Progression free survival		
Median (95%CI) PFS	[redacted]	[redacted]
Median follow-up, months	[redacted]	[redacted]
Progression/death, n (%)	[redacted]	[redacted]
Estimated PFS rate at month 12, % (95%CI)	[redacted]	[redacted]
Estimated PFS at month 18, % (95%CI)	[redacted]	[redacted]
Overall survival		
Median (95%CI) OS	[redacted]	[redacted]
Median follow-up, months	[redacted]	[redacted]
Estimated OS rate at month 12, % (95%CI)	[redacted]	[redacted]
Estimated OS at month 18, % (95%CI)	[redacted]	[redacted]
Key: CI, confidence interval; PFS: progression free survival; OS: overall survival; NE: not evaluable		

3.2.3 Adverse reactions

The company presents an overview of safety outcomes from the 18-month analysis of the ZUMA-5 FL patients in section B.2.10 of the CS. The safety analysis set (SAS) was used for all safety analyses for the study, and comprised all patients treated with any dose of axi-cel. No adverse event (AE) data for SCHOLAR-5 are reported in the CS. SCHOLAR-5 is described by the ERG in section 3.3. It is reported in the SCHOLAR-5 CSR that: "*Given the retrospective, observational design of the study, any reporting of adverse drug events had occurred prior to data collection and no additional reporting of AEs took place during this study*."

Published AE data are available for the SCHOLAR-5 Cohort C participants, the prospective cohort created from an open-label Phase II study, DELTA;[30] however, these data include patients who had received ≥2 or more lines of therapy, who are not part of the scope of this appraisal and were treated with idelalisib, which is currently unavailable to 4L+ FL patients in England. The ERG, therefore, feels that it is inappropriate to consider the AE data for the DELTA study in this appraisal.

The company states in Appendix F of the CS that no further studies reporting additional adverse events were identified. The company's economic model compares the AE frequencies from ZUMA-5 with AE frequencies for comparators as reported in the trials that informed the modelling for NICE appraisal TA627 (lenadlidomide with rituximab for previously treated FL).[21] A critique of the company's economic modelling of AE data is presented in chapter 4.

The company presents a summary of common adverse events in Table 17 of the CS, and a summary of serious adverse events (SAEs) that occurred in ≥2% of patients in Appendix N of the CS. Of the patients with ≥3 lines of prior therapy (n=78), the most common any grade adverse events (AEs) of patients with ≥3 lines of therapy were pyrexia ([redacted] patients [redacted] hypotension ([redacted] patients [redacted]), and headache ([redacted] patients [redacted%]). The most common Grade ≥3 AEs were neutropenia ([redacted] patients [redacted]), anaemia ([redacted] patients [redacted], and pyrexia ([redacted] patients [redacted%]). The most common SAEs experienced by patients with ≥3 lines of therapy were pyrexia ([redacted] patients [redacted]), pneumonia ([redacted] patients [redacted]), confusional state ([redacted]patients [redacted]), and encephalopathy ([redacted] patients [redacted%]).[31] The most common Grade ≥3 SAEs were

encephalopathy ([redacted] patients [redacted]), pneumonia ([redacted] patients [redacted]), and confusional state ([redacted]patients [redacted]).

The company presents details of treatment-emergent adverse events (TEAE) in Table 16 of the CS. A summary of TEAE and treatment-related AEs is presented in Table 11. Of the patients with ≥3 lines of prior therapy, [redacted] patients ([redacted]) experienced at least one serious TEAE, and [redacted] patients ([redacted]) experienced a Grade ≥3 serious TEAEs; [redacted] patients ([redacted]) experienced a serious treatment-related TEAE, and[redacted] patients ([redacted]) experienced a Grade ≥3 serious treatment-related TEAE. At the 18-month analysis data cut-off date, [redacted] Common treatment-related adverse events occurring in ≥20% of patients are presented in Table 18 of the CS. The most common any grade treatment-related AEs of patients with ≥3 lines of therapy were pyrexia ([redacted] patients [redacted]), hypotension [redacted]patients [redacted]) and headache ([redacted] patients [redacted]). The most common Grade ≥3 treatment-related AEs were neutropenia ([redacted] patients [redacted]) pyrexia ([redacted] patients [redacted]), hypoxia ([redacted] patients [redacted]), and encephalopathy ([redacted] patients [redacted]).

Table 11 Safety summary for FL patients in ZUMA-5 with three or more lines of therapy

Type of adverse event	FL patients with three or more lines of prior therapy SAS (n = 78)	
	Number (%) of patients experiencing AE any grade	Number (%) of patients experiencing AE Grade ≥3
TEAE	[redacted]	[redacted]
Serious TEAE	[redacted]	[redacted]
Treatment-related TEAE	[redacted]	[redacted]
Serious treatment-related TEAE	[redacted]	[redacted]
Death due to treatment-related TEAE	[redacted]	[redacted]
Common treatment-related adverse events occurring in ≥ 20% of FL patients in ZUMA-5		
Pyrexia	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Tremor	[redacted]	[redacted]
Chills	[redacted]	[redacted]
Sinus tachycardia	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Confusional state	[redacted]	[redacted]
Hypoxia	[redacted]	[redacted]
Encephalopathy	[redacted]	[redacted]

Key: AEs, adverse events FL, follicular lymphoma; SAS, safety analysis set; TEAE, treatment-emergent adverse event.
Source: ZUMA-5 CSR 18-Month Addendum[29]

Adverse events of special interest

The company presents adverse events of special interest in section B.2.10.4 of the CS; including cytokine release syndrome (CRS), neurological events, and cytopenia and the ERG presents a summary of these data in Table 12.

- **Cytokine release syndrome** Of the FL patients with ≥3 lines of prior therapy, [redacted] ([redacted]) experienced a CRS event, of which [redacted] had Grade ≥3 CRS, and [redacted] had Grade 5 CRS, who

[redacted]

[redacted]. The most common symptoms of CRS Grade ≥3 were hypoxia ([redacted] patients [redacted]), pyrexia ([redacted] patients [redacted]) and hypotension ([redacted] patients [redacted]). The median time to onset of CRS was [redacted] days (range: [redacted]) following axi-cel infusion. At the 18-month analysis data cut-off date, CRS had resolved in [redacted]. For the [redacted] patients with FL whose CRS had resolved, the median duration of CRS was [redacted]days (range: [redacted]).

- **Neurological events** [redacted] ([redacted]) patients with ≥3 lines of prior therapy had at least one neurological event of any grade, and [redacted] ([redacted]) had Grade ≥3 neurological events. [redacted] had a Grade 5 neurological event. The most common Grade ≥3 or higher neurological events were encephalopathy ([redacted] patients [redacted]), and confusional state ([redacted] patients [redacted]). The median time to onset of neurological event was [redacted]days (range: [redacted]); [redacted] had neurological events with an onset >80 days after the axi-cel infusion. The company state that the clinical experts they consulted indicated that the observed delayed/late-onset, low-grade neurological events were not likely to have any considerable impact.[10] [redacted]had unresolved neurological events at the 18-month analysis data cut-off. Of these patients, [redacted]. For the [redacted] patients with FL whose neurological event had resolved, the median duration of the event was [redacted]_days (range: [redacted]).
- **Cytopenia** Of the patients with ≥3 lines of therapy, [redacted]) experienced a cytopenia of any grade, and [redacted] experienced a Grade ≥3 cytopenia. Of the patients with ≥3 lines of therapy, [redacted] experienced Grade ≥3 neutropenia; [redacted] experienced Grade ≥3 thrombocytopenia; and [redacted] experienced Grade ≥3 anaemia. For FL patients whose events had resolved, the mean (standard deviation) and median (range) times to onset of cytopenias were [redacted] ([redacted]) and [redacted] ([redacted]) days after axi-cel infusion. The median duration of cytopenias were [redacted] (range: [redacted]) days.

Table 12 Summary of adverse events of special interest for FL patients in ZUMA-5 with three or more lines of therapy

Type of adverse event of special interest	**FL patients with three or more lines of prior therapy SAS (n = 78)**		
	Number (%) of patients experiencing AE any grade	**Number (%) of patients experiencing AE Grade ≥3**	**Number (%) of patients experiencing AE Grade 5**
Any CRS event[a]	[redacted]	[redacted]	[redacted]
Symptoms of CRS[b]			
Pyrexia	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]
Chills	[redacted]	[redacted]	[redacted]
Hypoxia	[redacted]	[redacted]	[redacted]
Sinus tachycardia	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]
Tachycardia	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]
Malaise	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]
Myalgia	[redacted]	[redacted]	[redacted]
Any neurological event	[redacted]	[redacted]	[redacted]
Type of neurological event, n (%)			
Tremor	[redacted]	[redacted]	[redacted]
Confusional state	[redacted]	[redacted]	[redacted]
Encephalopathy	[redacted]	[redacted]	[redacted]
Aphasia	[redacted]	[redacted]	[redacted]
Somnolence	[redacted]	[redacted]	[redacted]

Agitation	[redacted]	[redacted]	[redacted]
Disturbance in attention	[redacted]	[redacted]	[redacted]
Dysarthria	[redacted]	[redacted]	[redacted]
Paraesthesia	[redacted]	[redacted]	[redacted]
Delirium	[redacted]	[redacted]	[redacted]
Hallucination	[redacted]	[redacted]	[redacted]
Patients with any cytopenia	[redacted]	[redacted]	NR
Patients with neutropenia, n (%)	[redacted]	[redacted]	NR
Neutropenia	[redacted]	[redacted]	NR
Neutrophil count decreased	[redacted]	[redacted]	NR
Febrile neutropenia	[redacted]	[redacted]	NR
Patients with thrombocytopenia, n (%)	[redacted]	[redacted]	NR
Thrombocytopenia	[redacted]	[redacted]	NR
Platelet count decreased	[redacted]	[redacted]	NR
Patients with anaemia, n (%)	[redacted]	[redacted]	NR

Key: AE, adverse event; FL, follicular lymphoma; SAS, safety analysis set

Notes: a CRS events are graded according to a modification of the criteria of Lee et al. (2014). Percentages are calculated using the total number of patients in the analysis set as the denominator. b individual CRS symptoms are coded using MedDRA Version 23.0 and graded per CTCAE Version 4.03. Percentages are calculated using the number of patients with any TE CRS of any grade. c. multiple incidences of the same adverse event in one patient are counted once at the worst grade for this patient. Events (neutropenia, thrombocytopenia or anaemia) with onset on or after axicabtagene ciloleucel infusion date are summarised. Thrombocytopenia is identified using the standard Medical Dictionary for Regulatory Activities (MedDRA) query (SMQ) for haematopoietic thrombocytopenia (narrow search). Neutropenia is identified using MedDRA search terms defined by Kite. Anaemia (including aplastic anaemia) is identified using the SMQ haematopoietic erythropenia (broad search).
Source: ZUMA-5 CSR 18-Month Addendum[29]

Infections

The company presents infection data in section B.2.10.4.3 in the CS. The infection data are not presented by the number of lines of prior therapy received by the FL patients. Infections were experienced by [redacted] patients ([redacted]), of whom [redacted] ([redacted]) had worst Grade 3 infections, and [redacted] ([redacted]) had worst Grade 4 infection. [redacted] Worst Grade 3

events included pneumonia ([redacted] patients [redacted%]) and urinary tract infection ([redacted]patients, [redacted]). The single worst Grade 4 event was sepsis ([redacted] patient, [redacted]). The company states that [redacted] The company reports that [redacted]; however, these data are not reported in the CS or the ZUMA-5 CSR 18-month addendum.

It is the ERG clinical expert's opinion that the AEs reported in the CS are in keeping with the AEs related to the use of axi-cel in diffuse large B cell lymphoma where it is already approved. CAR-T is a single treatment, and most AEs occur within 30 days of treatment, with a far lower risk of AEs beyond that time. This differs from SOC where the risk of AEs remains similar for the duration of treatment, which is often 6 months depending on the regimen used. Like CAR-T, there is still a risk of AEs after treatment, but this is much smaller and gradually declines with time post-treatment.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company did not conduct any formal indirect or mixed treatment comparison but instead presented a comparison with SCHOLAR-5, described in the CS as an international, multicentre, external cohort control study for the purpose of providing comparative evidence for axi-cel in patients fulfilling the eligibility criteria of ZUMA-5. The SCHOLAR-5 CSR (Table 3; ref 21, Doc B) presents a comparison of the inclusion and exclusion criteria of SCHOLAR-5 and ZUMA-5.[32] In general, the criteria are aligned appropriately, and the ERG clinical expert has no concerns. SCHOLAR-5 consisted of three cohorts, described in full in Section B.2.9.1.1 of the CS. In brief, Cohort A and Cohort B were retrospective cohorts created from medical records of a total of seven sites in the UK, France, Spain, Portugal, and the USA. Cohort C consisted of participants of a single-group, open-label, Phase II study, DELTA (Gopal 2014), conducted at 41 sites in the USA and Europe of which the main inclusion criteria were a confirmed diagnosis of B-cell iNHL, including (among others) histological types FL grade 1, 2 or 3a.[30] Inclusion criteria also specified prior treatment with ≥2 prior chemotherapy-based or immunotherapy-based regimens for iNHL, prior treatment with rituximab and an alkylating agent for iNHL and refractoriness to both rituximab and an alkylating agent. The CS states that cohorts were restricted to FL patients with at least three

prior lines of treatment before construction of the analysis set. The ERG noted that SCHOLAR-5 included patients outside of the UK and some of the treatments received by these patients are not in line with clinical practice in England.

Propensity scoring methods - specifically standardised mortality ratio (SMR) weighting - were applied to account for imbalances of confounders between ZUMA-5 and SCHOLAR-5 populations. The ERG felt it was not transparent on how the SMR weighting was applied to the propensity scoring. However, the weighting has improved comparability between the ZUMA-5 and SCHOLAR-5. Baseline characteristics of SCHOLAR-5 and ZUMA-5 patients pre- and post-weighting are presented in Table 14 of the CS and reproduced as Table 13 below. The ERG notes that the abbreviation EES in the table is not defined by the company. However, the abbreviation ESS is defined as estimated sample size, and the ERG believes that incidences of EES should read ESS.

Table 13 Baseline characteristics of patients pre-and post-weighting; FL patients with three or more lines of prior therapy, SCHOLAR-5 ESS, ZUMA-5 mITT [reproduced from Table 14, Document B of the CS]

Characteristics	Pre-weighting			Post-weighting		
	SC-5 (n = 82)	Z-5 (n = 78)	p-value [SMD]	SC-5 (EES = 77)	Z-5 (n = 78)	p-value [SMD]
POD24, n (%) Yes No Missing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior lines of therapy Mean (SD) Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Relapsed/refractory to prior line of therapy Relapsed Refractory	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Characteristics	Pre-weighting			Post-weighting		
	SC-5 (n = 82)	Z-5 (n = 78)	p-value [SMD]	SC-5 (EES = 77)	Z-5 (n = 78)	p-value [SMD]
Missing	[redacted]	[redacted]		[redacted]	[redacted]	
Prior SCT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]		[redacted]	[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]		[redacted]	[redacted]	
Tumour bulk ≥ 7 cm	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]		[redacted]	[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]		[redacted]	[redacted]	
Time since last treatment, months	[redacted]		[redacted]	[redacted]		[redacted]
Mean (SD)	[redacted]	[redacted]		[redacted]	[redacted]	
Median (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]		[redacted]	[redacted]	
CR or PR to prior line of therapy	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]
Yes	[redacted]	[redacted]		[redacted]	[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]		[redacted]	[redacted]	
Age ≥ 65 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]		[redacted]	[redacted]	
No	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Missing	[redacted]	[redacted]		[redacted]	[redacted]	
Prior anti-CD20 + alkylator combination	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Yes	[redacted]	[redacted]		[redacted]	[redacted]	

Characteristics	Pre-weighting			Post-weighting		
	SC-5 (n = 82)	Z-5 (n = 78)	p-value [SMD]	SC-5 (EES = 77)	Z-5 (n = 78)	p-value [SMD]
No Missing	[redacted]	[redacted]		[redacted]	[redacted]	

Key: CR, complete response; ESS, estimated sample size; mITT, modified intent-to-treat; POD24, progressed disease within 24 months after initiation of first-line anti-CD20 chemo combination therapy; PR, partial response; SC-5, SCHOLAR-5; SCT, stem cell transplant; SD, standard deviation; SMD, standardised mean difference; Z-5, ZUMA-5.
Note: Percentages may not add up to 100% due to rounding.
Source: SCHOLAR-5 Technical Report.[32]

3.4 Critique of the indirect comparison and/or multiple treatment comparison

Indirect or multiple treatment comparisons were not conducted by the company for this appraisal.

3.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG requested the time to event data for progression-free survival and overall survival, but the company explained that they do not have permission to share their patient-level data (i.e., the time to event raw data underpinning the Kaplan Meier curves).

3.6 Conclusions of the clinical effectiveness section

The company decision problem is appropriate for addressing the NICE final scope for this appraisal. The company did not conduct any formal indirect or mixed treatment comparison. The key clinical effectiveness evidence for axi-cel for treating relapsed or refractory follicular lymphoma was based on a comparison with SCHOLAR-5 cohorts which were created from three data sources. Two of the data sources were retrospective cohort (real-world analysis set) which contained 58 patients and the third data source was a prospective cohort created from an open-label Phase II study, DELTA which contained 24 patients. The ERG noted that there were differences in the distribution of ECOG performance score (0 and 1) between ZUMA 5 and SCHOLAR 5. Another possible source of bias is that some patients in SCHOLAR 5, received treatments not approved for routine use by NHS England (e.g., idelalisib as part of the DELTA study). It is, therefore, plausible that the results from SCHOLAR-5 may overestimate OS for the current 4L+ treatments used in NHS England, which potentially acts against axi-cel; however, we do not have data to verify this. It would have been preferable to have comparator cohorts more in line with current NHS practice in England.

With regards to propensity scoring methods, specifically SMR weighting, the ERG felt the weighting has improved comparability between the ZUMA-5 and SCHOLAR-5; however, it is not transparent how this was performed.

After reviewing the analysis of the outcomes presented in the CS, the ERG agrees with the company that there is a beneficial effect on OS, PFS and RR rate from axi-cel. The Kaplan Meier plots show a reduction in the risk of disease progression and death, however, the ERG noted that the median PFS and OS were not reached for ZUMA-5. Although the confidence intervals around the effect sizes were wide, the large effect sizes on the ORR and CR show the difference between the two cohorts.

The ERG has inspected the adverse events being reported in ZUMA-5 in section B.2.10 of the CS. The ERG is not concerned with the proportions of serious adverse events or rates of adverse events. No adverse event (AE) data for SCHOLAR-5 are reported in the CS.

4 COST EFFECTIVENESS

4.1 ERG comment on company's review of cost-effectiveness evidence

The company conducted a systematic literature review of economic evaluations. Details were provided in appendix G of their submission. Comprehensive searches were originally undertaken to May 2020, and then later updated to May 2021. The review aimed to include all economic evaluations, and resource use and costing studies, of any interventions in adults with relapsed or refractory indolent non-Hodgkin lymphoma - grade 1-3a follicular lymphoma, or nodal or extra nodal marginal zone lymphoma.

The review identified a total of 33 studies, of which 19 were full economic evaluations. Details of the included study designs, modelling approaches, modelling inputs and findings were all tabulated from comparison in appendix G of the company submission. In their main submission document, the company have focused on three economic modelling studies that have informed previous NICE appraisals in r/r/ FL: TA604 (idelalisib), TA627 (lenalidomide with rituximab) and TA629 (obinutuzumab with bendamustine [TA472 CDF review]).[21, 22, 33, 34] The company notes that insights were drawn from these appraisals throughout their own submission. They further note that in addition to those studies identified in their review, they drew insights from three previous NICE appraisals of CAR T-cell therapies in advanced previously treated lymphoma indications, and a published mock appraisal of regenerative and cell therapy products.[35-38]

The ERG is satisfied that the company have undertaken a thorough review of the published economic evidence of relevance to this appraisal. Rather than using the existing economic evidence base to draw conclusions about the cost-effectiveness of axi-cel for r/r FL, the focus of their review was on gaining insights on methodological approaches, inputs and assumptions of relevance to the current appraisal.

4.2 Summary and critique of the company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 14 NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Aligns with reference case
Perspective on costs	NHS and PSS	Aligns with reference case
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Aligns with reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Aligns with reference case
Synthesis of evidence on health effects	Based on systematic review	Aligns with reference case, but limited evidence available to inform comparative effectiveness.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Aligns with reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Aligns with reference case, but no data available that applies specifically to the lines of therapy specified in the company's proposed population.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Aligns with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Aligns with reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Aligns with reference case, although some uncertainty around some of the values applied.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Aligns with reference case

PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.

4.2.2 Model structure

The company describe their de novo cost-effectiveness model in section B.3.2.2 of their submission document. It takes the form of a three-state partitioned survival model: pre-progression, progressed and dead. The structure, they note, is consistent with those used in all previous NICE appraisals in relapsed or refractory follicular lymphoma.

The PFS and OS data to inform the model comes from the relevant subgroup (treatment line 4L+) of the ZUMA-5 trial for axi-cel and from the propensity score weighted SCHOLAR-5 data for the blended comparator (section 3.3 above). Based on experience from the use of CAR-T therapies in other haematological cancers, and precedence set by previous NICE appraisals, a proportion of those alive and progression free at 5 years are assumed be long-term survivors. This proportion face no further risk of progression but do face an elevated mortality rate relative to the age/sex matched general population. An SMR of 1.09 is applied in the base case in line with appraisals of CAR-T therapies in diffuse large B-cell lymphoma (DLBCL) (TA559, TA567).[35, 36]

Model settings are in line with the NICE reference case with respect to perspective on costs and outcomes, time horizon, and discounting. A cycle length of 28 days was chosen, and a half-cycle correction appropriately applied.

The ERG is broadly supportive of the model structure but note a few structural uncertainties related to the Part SA approach and the company's long-term survivorship assumptions.

- *A fraction of those projected to be progression free at 5 years (equating to 25% of the total cohort) are assumed to be longer-term survivors at zero risk of progression from 5 years onwards. The remaining survivors face risks of progression and death based on the chosen extrapolation curves for PFS and OS. However, the extrapolation curves are fitted to observed data for the overall mITT cohort and so may not be appropriate for extrapolating*

outcomes for the subset who are not long-term survivors. The ERG asked the company to consider this at the clarification stage and explore the impact of applying higher risks of progression and death beyond five years to those not considered to be long-term survivors.

- *The fractional weighting of long-term survivors versus those who remain at risk of progression in the model is assumed to be constant over time from 5 years. In reality, the fraction of the surviving cohort that are long-term survivors should increase over time as those who are not continuing to progress and face a higher risk of mortality. The company were also asked to consider an adjustment for this in the model at the clarification stage.*
- *The Part-SA approach creates challenges with respect to validating the time spent in the progressed disease state by treatment arm of the model, and also the expected costs of subsequent lines of treatment. For example, all progressed patients, irrespective of treatment arm, are assumed to incur subsequent therapy costs equal to those incurred in the comparator arm in the first treatment line of the model. This is despite an expectation that increasing lines of therapy are associated with poorer response rates and reduced PFS and OS. Further, the selected PFS and OS curves in the model infer that patients who progress on axi-cel can expect to survive for longer following progression compared to those who progress on the comparator. The clinical plausibility of this is unclear, but if valid it might also suggest that higher subsequent therapy costs could be expected per progressed patient on axi-cel. This is not the case in the model. There is a post progression survival benefit for axi-cel without any increase in subsequent therapy costs.*

4.2.3 Population

In line with the proposed marketing authorisation and the data from ZUMA-5 used to inform the model inputs, the population considered in the model is those with [redacted] (referred to as the 4L+ population in the company submission). The starting age of the cohort is [redacted] in line with the mean age at baseline of 4L+ patients in ZUMA-5.

The ERG has no major concerns regarding the proposed population, but note it is a subset of the overall population of relapsed or refractory non-Hodgkin lymphoma as set out in the final scope for the appraisal. Correspondingly, it is a subgroup from the ZUMA-5 trial which is used to inform the model inputs. ZUMA-5 also included patients with marginal zone lymphoma and patients with relapsed or refractory disease after fewer prior lines of therapy.

4.2.4 Interventions and comparators

The intervention is axi-cel, as described in section B.3.2.3.1 of the company submission.

The company argue that there is no true standard of care for the 4L+ r/r FL population, and so consider the comparator to be a basket of treatments (blended comparator). Whist rituximab monotherapy and best supportive care (BSC) were listed as comparators in the scope, the company argue that both would be reserved for patients considered not fit enough to receive intensive active therapy, a group considered not suitable for CAR-T therapy. Therefore, both rituximab monotherapy and BSC are excluded from the blended comparator. The data used to inform the comparative efficacy of the blended comparator come from the SCHOLAR-5 study. Further discussion of the blended comparator is provided in the following sections.

The ERG's clinical expert was broadly in agreement with the company's blended comparator, and that rituximab and BSC should not be considered comparators.

4.2.5 Perspective, time horizon and discounting

The perspective on cost and outcomes is in line with the NICE reference case. The time horizon is 40 years, with a starting age of [redacted] for the modelled cohort. Given the potential for long-term survivorship for a fraction of the cohort, this seems reasonable. Shorter time horizons are explored by the company in scenario analyses.

4.2.6 Treatment effectiveness and extrapolation

Clinical inputs for axi-cel were derived from the analysis of PFS and OS data of the modified intention-to-treat population of ZUMA-5, comprising [redacted] patients with r/r FL with three or more lines of prior therapy. All analyses were based on the September

2020 data cut at the time of the main company submission, providing a median follow-up of [redacted] for PFS and [redacted] for OS. The company plans to update these analyses and utilise later data with 6 more months of follow-up during technical engagement.

Clinical inputs for the comparator arm, current 4L+ care, were derived using propensity score weighted data from the SCHOLAR-5 study discussed in section 3.3 above.[32] This external control included [redacted] patients with FL at 4L+ for comparison with the ZUMA-5 mITT population. The company noted that following propensity weighting of the SCHOLAR-5 data the effective sample size was reduced to 77 patients. However, they further clarified that due to an absence of progression dates for the index therapy in DELTA study, a sub-cohort of SHOLAR-5, these patients were excluded from the PFS analysis. Thus, there were fewer patients (n=51) to inform PFS post-weighting.

Given the unique mechanism of action of axi-cel compared to other available 4L+ treatments, the company considered it unreasonable to expect proportional hazards between treatment arms to hold, and so independently fitted parametric curves to PFS and OS data for each treatment arm (see company submission, section B.3.3.1.4). Seven standard parametric survival models were fitted for each outcome. Following NICE DSU TSD guidance, the company considered visual fit, statistical fit, and plausibility of long-term extrapolation, based on clinical opinion, to select a parametric curve for each outcome.

Axi-cel PFS

Based on consideration of visual and statistical fit, and clinical expert opinion, the company selected the most conservative Weibull curve for extrapolation of PFS. However, the company note that from interviews with clinical experts, it is reasonable to expect a proportion of r/r FL patients treated with axi-cel to have mortality hazards that are more in line with the general population after 5 years. The company base case assumes this applies to 25% of the cohort, which is approximately [redacted] of those alive and progression free at 5 years ([redacted]) in the model. They assume that this 25% face zero risk of progression from 5 years, and a risk of death which is held at 9% (SMR=1.09) above general population mortality. The remainder, who are not

considered to be long-term survivors, continue to follow the extrapolated hazard of progression or death based on the chosen Weibull PFS curve.

Axi-cel OS

With respect to OS, the company fitted the same seven standard parametric curves to the data from ZUMA-5 (see Figure 22 of the company submission, document B), and selected the Weibull curve for their base case. The company highlight the immaturity of the OS data and note that the AIC and BIC were within five points across all models with the exception of the BIC for the generalised gamma and log normal models. The more pessimistic extrapolations produced by the generalised gamma and Gompertz models were ruled out based on advice of clinical experts on the clinical plausibility of the long-term extrapolations, as was the log-normal which produced unrealistically high long-term survival. Of the remaining options, the Weibull was chosen for the company base case. This is the third most pessimistic (after the Gompertz and generalised gamma), projecting survival of [redacted] at 5 years, [redacted] at 10 years, [redacted] at 20 years, [redacted] at 30 years, and [redacted] at 40 years. However, as noted above for PFS, the company base case assumes that 25% of those treated with axi-cel are long-term survivors who face and SMR adjusted general population mortality from 5 years onwards. Therefore after 5 years, the chosen Weibull is only used to extrapolate survival of those assumed not considered to be long-term survivors. It should be further noted that there is an override in the model which ensures the extrapolated mortality never falls below the mortality hazard for the age/sex matched general population. This applies to the chosen Weibull curve from [redacted] years when it projects [redacted] survival. This is somewhat counterintuitive, as it assumes a lower mortality rate for the non-long-term survivors compared to long-term survivors from [redacted] years onwards.

ERG critique

There are clearly challenges related to the extrapolation of PFS and OS given the immaturity of the data. Further uncertainties relate to company's long-term survivor assumptions, with currently no data available to validate this in the r/r FL, 4L+ population. In addition, their approach to applying different hazards of progression and death for long-term survivors creates some inconsistencies in the model:

1. *The extrapolation curves were fitted to data for the whole mITT ZUMA-5 cohort, but from 5 years are only applied to those assumed not to be long-term survivors. There is scope for these fitted curves to overestimate the survival for this fraction of the cohort; had it been possible to fit curves separately for non-long-term survivors, more pessimistic extrapolations may have been obtained.*
2. *Related to (1) above, it is not clear if the clinical experts who validated the chosen PFS and OS extrapolation curves were aware that they were intended for projecting expected survival for only a fraction of the cohort from 5 years onwards, rather than the whole cohort.*
3. *The company's original approach to separating the hazards (from 5 years) assumed a constant proportional split between long-term and non-long-term survivors, which didn't account for the differing hazards of progression and death moving forwards. The company implemented a correction for this at the clarification stage, which had a modest downward impact on the ICER (see company response to the clarification letter, QB7).*
4. *The override to ensure the mortality hazard for non-long-term survivors doesn't fall below general population mortality, whilst assuming long-term survivors face SMR adjusted general population mortality, results in non-long-term survivors facing a lower hazard of death than long-term survivors from [redacted] years in the model.*

The above issues may contribute to the extrapolated post-progression life-year gain for axi-icel versus current 4L+ care in the company base case. Whilst there are plausible reasons why axi-cel treated patients might experience better post-progression survival than those treated with current 4L+ therapies (see company response to clarification letter, QB8), overestimating OS for the non-long-term survivor fraction could also contribute to the modelled post-progression survival benefit. Given the above, the ERG requested scenarios from the company to explore the impact of increasing the risk of progression and death for the non-long-term survivor fraction from 5 years. The company provided this by applying hazard ratios of 1.09 and 1.2 to the chosen axi-cel PFS and OS curves from five years, which had a modest upward impact on the ICER years (see response to clarification letter, QB6).

The company noted the arbitrary nature of the HR values applied given the lack of data.

A further potential issue related to the modelling of OS for axi-cel treated patients, is the acknowledgement that a number of patients in ZUMA-5 who achieved a complete or partial response at month three, but subsequently experienced disease progression, were allowed retreatment with axi-cel. This is noted to have occurred in [redacted] of the 4L+ mITT cohort. The company noted that retreatment would not be expected to occur in routine clinical practice in England and so have not accounted for these costs in the model. However, they have not made any corresponding adjustment to the Kaplan-Maier OS data. Nevertheless, the OS data is very immature, and it may be too early for any potential bias to have materialised in the observed OS data. But it perhaps should be considered when choosing extrapolation curves for OS. In the absence of providing an adjustment to OS to account for the removal of post-progression axi-cel from the ZUMA-5 data, the company have provided a scenario analysis in response to the clarification letter which includes these retreatment costs. This has a moderate upward impact on the ICER (see response to clarification letter, QB2.)

Standard 4L+ PFS

Parametric survival models were fitted to the propensity score weighted data from the SCHOLAR-5 study. As indicated above, the company noted that the timing of progression could not be determined for cohort C of SCHOLAR-5, so these patients were excluded from the analysis of PFS to inform the comparator arm. This results in substantially fewer patients (n=51) informing the PFS curve compared to the number informing the OS curve (n=77) for the blended comparator. Cohort C of SCHOLAR-5 came from the open label phase II DELTA study of patients with r/r FL treated with idelalisib.

The available PFS data was mature, with the Kaplan-Maier curve reaching [redacted] by approximately 31 months. This results in less uncertainty related to the choice of parametric curve in the comparator arm, and the company note that clinical experts they consulted suggested all the parametric curves provided plausible extrapolations.

Therefore, the company selected the exponential curve for their base case based on statistical fit (lowest AIC and BIC).

Standard 4L+ OS

OS for the comparator arm was informed by analysis of the propensity weighted data from all sub-cohorts of SCHOLAR-5. Based on AIC and BIC, the generalised gamma provided the best statistical fit to the observed OS data. However, the company note that it, along with the Gompertz, provides implausibly high long-term survival projections. The company note that based on the clinical validation interviews, the gamma curve was selected for the base case based on plausibility of the extrapolation. This provides the second most pessimistic extrapolation of OS of the available curves for the comparator arm (see Figure 24 of the company submission, document B)

ERG critique

There are several uncertainties relating to the company's approach to estimating efficacy inputs for the comparator arm of the economic model. The uncertainties inherent in constructing an external control group for the single arm ZUMA-5 trial were discussed in section 3.3 above. Accepting that the company are limited by the availability of data and the non-randomised design of ZUMA-5, the ERG identifies some further issues related to the company's approach:

- *It is potentially problematic that cohort C of SCHOLAR-5 (data from the DELTA trial) was excluded from the analysis of PFS but included for the analysis of OS. The result is that PFS in the model is informed by [redacted] fewer patients than OS, which may invalidate the use of the chosen curves for partitioning the standard care cohort. The ERG sought clarity on this issue at the clarification stage. The company noted in their response that a subgroup analysis had been conducted as part of the SCHOLAR-5/ZUMA-5 comparative analysis, in which the DELTA sub-cohort of SCHOLAR-5 had been excluded from the comparison of OS using the smaller inferential analysis set of ZUMA-5 (n=60). They note that this produced an estimated hazard ratio for OS that was very similar to the main analysis which included DELTA patients (see company response to clarification letter, QB3). However, this does not fully address the concern because: 1) the model does*

not rely on hazard ratios, but independently fitted survival curves; and 2) the model outcomes for axi-cel are based on the mITT population rather than the inferential analysis set. Given the above, the ERG believe it might have been preferable to conduct an analysis that excluded the DELTA patients from the OS curve fitting for the comparator arm. This could be further justified by the potential lack of generalisability of the DELTA cohort (treated with idelalisib) to the NHS in England, where idelalisib is not available.

- *As things stand, with the DELTA patients included in the OS curves for the comparator arm, there appears to be* [redacted]*, which may not be consistent with the PFS curve which excludes the DELTA patients. Further, the parametric curves that provide the best statistical and visual fit to the observed OS data result in implausibly high projections of long-term OS, whilst the curves that provide more plausible long-term projections of OS, according to clinical experts, provide poorer statistical and visual fit to the observed data. The company acknowledge this issue and note that they prioritised the plausibility of extrapolation during the curve selection process. The ERG acknowledges that the better fitting curves lack plausibility with respect to long-term survival but are concerned that the chosen OS curve provides a poor fit to the observed data which undermines confidence in its suitability for extrapolation.*
- *A further issue with the comparator data from SCHOLAR-5 is that it includes patients who received treatments that are not available in the NHS in England (including idelalisib). Therefore, the company reweighted the distribution of SCHOLAR-5 treatments for the purpose of calculating the blended comparator costs in the model (Table 15). However, no corresponding adjustment to efficacy was possible. The company were asked to comment on the expected direction and magnitude of any bias that this may introduce. The company response focussed on the more favourable outcomes that idelalisib would be expected to have over treatments that are used at fourth line or above in the NHS in England. Thus, they suggest that the SCHOLAR-5 curves are optimistic compared to current clinical practice in England. However, the ERG notes that those patients (19%) who received CVP alone in SCHOLAR-5*

and were redistributed to other treatments for the purpose of costing, may have experienced poorer outcomes than would be expected in the NHS. It is not clear how the 26% of patients who received experimental treatments in SCHOLAR-5 would have fared on the other treatments available in routine practice in England. Therefore, it is difficult to predict the overall direction and magnitude of bias caused by the mismatch between the SCHOLAR-5 treatment distribution and the treatment distribution used in the NHS in England. Assuming that those treated with idelalisib or experimental treatments in SCHOLAR-5 would tend to have experienced better outcomes than they otherwise would, the ERG believes that the mismatch is more likely to biases in favour of the comparator (against axi-cel). However, it remains uncertain.

Table 15: Distribution of current 4L+ care therapies [source: Table 40, Document B of the CS].

Treatment	SCHOLAR-5 distribution	Include as comparator?	Re-weighted distribution
Idelalisib	12.0%	No	**0.0%**
Bendamustine + obinutuzumab	5.3%	Yes	**13.3%**
Bendamustine + rituximab	10.7%	Yes	**26.7%**
CVP + rituximab	6.0%	Yes	**15.0%**
Radioimmunotherapy	3.0%	No	**0.0%**
Lenalidomide + rituximab	9.0%	Yes	**22.5%**
R-CHOP	9.0%	Yes	**22.5%**
CVP	19.0%	No	**0.0%**
Experimental	26.0%	No	**0.0%**
Total	100.0%	**Re-weighted total**	**100.0%**

Key: 4L+, fourth-line plus; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; CVP, cyclophosphamide, vincristine, and prednisone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone; SCT, stem cell transplant.

4.2.7 Health related quality of life

Quality of life was captured in the model by applying utility weights to pre-progression and progressed health states, with adverse event disutilities applied separately. In the company base case, no distinction was made for patients classified as long-term survivors with the chosen pre-progression utility value assumed to apply.

As quality-of-life data were not collected in the ZUMA-5 or SCHOLAR-5 studies, a systematic literature search was conducted to identify relevant utility values for use in the model. The search identified 7 studies reporting health state utility values in the r/r FL population but none of the identified studies were used in the model base case. Instead, assumptions from the NICE appraisal of lenalidomide with rituximab for previously treated FL (TA627) were used.[21] In TA627, utility values were derived from quality of life data collected in the AUGMENT study but capped to ensure the progression-free utility value remained below age-adjusted general population values. Relative utility decrements were then applied to the progressed health states. The company adopts the same approach here on the basis that these utility values were accepted by NICE in a similar patient population. The utility values used are reported in table 16 below.

Many of the studies identified in the literature search reported the same set of utility values from Wild et al 2006/Pettengell at al 2007 and these values were used in sensitivity analysis.[5, 39] As only the abstract was available for the Wild et al study, this was not included in the literature review but information from the study is reported in other published papers and relevant NICE appraisals (TA627, TA604).[21, 33] The study reported in Wild et al and Pettengell et al is from 222 patients in the UK with histologically confirmed FL. Patients completed several patient-reported outcome measures and were analysed according to five disease states: ‘active disease-newly diagnosed’, ‘active disease-relapsed’, ‘partial response’, ‘complete response’ and ‘disease free’. These health states were then grouped to form two broad health states of progression-free (partial response, complete response and disease free) and progressed disease (active disease-newly diagnosed and active disease-relapsed). Quality of life was assessed using the EQ-5D, Functional Assessment of Cancer Therapy –Lymphoma (FACT-Lym) measure and the Hospital Anxiety and

Depression Scale. The EQ-5D data were used to derive utility values for progression-free and progressed disease health states.

The CS notes that in TA604 utility values from Pettengell et al/ Wild et al were used in the model, whereas in TA627 these values were used in a scenario analysis due to the availability of EQ-5D data from the AUGMENT study. Given the lack of relevant quality of life data from the trial, the company acknowledge the chosen utility values are uncertain and explore the impact of using alternative data sources. A summary of the utility values identified in the literature and relevant NICE appraisals which were used in scenario analyses are presented in Table 16.

Table 16: Summary of relevant utility values used in the company base case and sensitivity analysis [adapted from Table 31, Document B of the CS]

Health state	Base case and TA627 FAD[21]	AUGMENT (TA627)[21]		Wild et al. (2006)[39] /Pettengell et al. (2008)[5] (TA604)[33]	GADOLIN (TA629; as reported in TA627)[21]
		R^2	R-mono		
Pre-progression	Age-matched general population 0.829 at baseline ([redacted] years)	0.847	0.840	0.805 (0.018)	On-treatment: 0.822 (0.010) Off-treatment: 0.807 (0.012)
Progressed disease	Age-matched general population (with relative decrement) 0.803 at baseline ([redacted] years)	Off-treatment: 0.821 On-treatment: 0.791	Off-treatment: 0.813 On-treatment: 0.784	0.736 (aggregated) 0.62 (0.06 – relapsed disease)	0.758 (0.024)

Key: FAD, Final Appraisal Determination; R-mono, rituximab monotherapy; R^2, lenalidomide with rituximab; TA, technology appraisal.

The ERG agrees there is uncertainty in the utility values due to a lack of quality-of-life data available in the patient population who would be eligible to receive axi-cel in clinical practice. A literature search identified a number of potentially relevant studies, but the company provided limited justification for deciding to adopt a similar approach to that used in TA627 in preference to other studies identified in the literature. While the patients in the AUGMENT study had r/r FL (or marginal zone lymphoma), the majority were enrolled at second-line (54%) with only 24% fourth-line or greater who would be comparable to the patients who would be eligible for axi-cel. As patients in the AUGMENT study are at an earlier stage in the disease pathway, clinical expert advice to the ERG indicates these patients would be expected to have a higher quality of life than patients receiving treatment at fourth line and beyond. Although it is not clear what line of treatment patients were receiving in the Wild et al/Pettengell et al study, the utility values are lower than those in AUGMENT and may better reflect the quality of life of patients at this later stage of the r/r FL treatment pathway. The values from Wild et al/Pettengell et al have been used in other appraisals in either the base case (TA604)[33] or sensitivity analysis (TA627)[21] but are also associated with some limitations. The study dates back to 2006 and is not published with only the poster abstract available. The utility values from the study are widely quoted in NICE appraisals but the ERG has been unable to verify them in a published paper. Despite these limitations, the ERG prefers the increased face validity of the Wild et al study utility values in the base case but conclude the lack of relevant quality of life data in fourth line r/r FL patients remains a key uncertainty.

Adverse events

The quality-of-life impact of adverse events was captured in the model as a one-off utility decrement. The impact of grade ≥3 treatment-related adverse events in 5% or more patients in ZUMA-5 were included. In addition, all grades of adverse events were included for those considered to be clinically important for CAR T-cell therapies. The CS notes this approach is consistent with previous NICE CAR T-cell therapy appraisals. The following adverse events were modelled for axi-cel:

- Grade ≥3 axi-cel related adverse events occurring in 5% or more of subjects in ZUMA-5 (see CS table 32 for adverse events included)

- Grade ≥3 treatment-emergent CRS occurring in ZUMA-5 ([redacted]) and any grade CRS requiring treatments with tocilizumab ([redacted])
- Patients who received immunoglobulin treatment ([redacted])

For current 4L+ care, adverse event frequencies were sourced from clinical trial data reported in TA627 for the treatments included in the basket of current care. Only grade ≥3 adverse events that occurred in 5% or more of ZUMA-5 patients were included in the model, which the CS states is a conservative assumption.

In terms of utility decrements, a one-off QALY decrement of 0.15 was applied in the first model cycle for most grade ≥3 adverse events occurring in more than 5% of patients based on a study by Guadagnolo et al in patients with Hodgkin lymphoma.[40] For grade ≥3 CRS a quality of life of 0 was applied for the duration of the event ([redacted]) and for hypogammaglobulinaemia it was assumed there would be no impact on quality of life. The approach taken is consistent with that used in NICE appraisals of CAR T-cell treatment in advanced lymphoma (TA559 and TA677).[35, 37]

The approach to adverse event disutilities is generally consistent with other relevant NICE appraisals of CAR T-cell treatments. Some simplifying assumptions have been made but in general the ERG considers these assumptions are reasonable. One potential area of uncertainty is the adverse event durations were taken from ZUMA-1 and ZUMA-2 as reported in TA677/TA559 rather than ZUMA-5. No explanation was provided for this other than maintaining consistency with other relevant NICE appraisals. This is unlikely to be a key source of uncertainty.

4.2.8 Resources and costs

The costs and resource use included in the model can be categorised as follows: axi-cel treatment-related costs, current 4L+ costs and administration, costs of subsequent treatments, health state resource use, adverse event and end-of-life costs.

Axi-cel treatment-related costs

In addition to drug acquisition and monitoring costs for axi-cel, other treatment-related costs are incurred due to axi-cel production involving patient T-cells. These include:

- Leukapheresis to extract patient T-cells
- Bridging therapy for some patients to remain stable prior to the CAR T-cell infusion
- Conditioning chemotherapy

A summary of the axi-cel costs included in the model is provided in table 17 below.

Table 17: Summary of axi-cel treatment costs (adapted from Table 39, Document B of the CS)

Axi-cel cost category	Cost	Source and assumptions
Leukapheresis	[redacted]	• Same approach as used in previous NICE appraisals for CAR T-cell therapies (TA677 and TA559).[35, 37] • Cost uses weighted average of stem cell and bone marrow harvest from NHS reference costs (2019/2020).[41] • The weighted average cost (£1,953.38) was adjusted to account for patients ([redacted]) who underwent leukapheresis but did not receive axi-cel
Bridging therapy	[redacted]	• Bridging therapy cost consisted of 1 dose of rituximab based on the ZUMA-5 trial where [redacted] required bridging therapy.
Conditioning chemotherapy	£2,880.65	Drug cost: • IV cyclophosphamide and IV fludarabine on 5th, 4th and 3rd day prior to axi-cel infusion • Drug wastage was included • Resulting costs were £17.50 and £39.51 per dose for cyclophosphamide and fludarabine respectively Hospitalisations: • Assumed administered in hospital over 3 days in elective inpatient setting, consistent with other CAR-T therapies • Cost based on weighted average malignant lymphoma, including Hodgkin lymphoma and non-Hodgkin lymphoma from NHS reference costs 2019/2020 (£7,301.52).[41] • Mean cost per day of £903.20 per day based on mean length of stay of 8.1 days for malignant neoplasms of lymphoid, haematopoeietic and related tissues
Drug acquisition	[redacted]	• The drug acquisition cost of axi-cel is [redacted] at list price.

		• A patient access scheme (PAS) of [redacted] has been agreed with NHS England reducing the acquisition cost to [redacted]. • [redacted] of the [redacted] treated patients in ZUMA-5 received re-treatment with axi-cel, but the costs are not included as thisdoes not form part of the expected marketing authorisation and is not expected to occur in practice.
Infusion and monitoring	[redacted]	Following infusion, patients are monitored in an elective inpatient setting consistent with assumptions applied in other CAR T-cell appraisals • Cost of hospitalisation based on weighted average cost for malignant lymphoma, including Hodgkin lymphoma and non-Hodgkin lymphoma from NHS reference costs.[41] • ZUMA-5 mean duration of hospitalisation is [redacted] days. Hospitalisation cost based on mean cost per day of £903.30 for [redacted] days.
Total	[redacted]	

The ERG considers the costs associated with axi-cel treatment in general have been implemented appropriately in the model and are largely consistent with the approach used in other relevant NICE appraisals for CAR T-cell therapies (TA559 and TA677).[35, 37] *One area of uncertainty relates to axi-cel retreatment. Although [redacted] of patients required retreatment in the ZUMA-5 trial, the costs of this were not included in the model on the basis that retreatment would not occur in practice. Following clarification, the company provided an analysis including retreatment costs to align with the clinical effectiveness data used in the model. This analysis included the costs associated with the elements of retreatment received by the patients in the ZUMA-5 trial and increased the total axi-cel cost to [redacted] (see response to clarification letter QB2) resulting in a moderate increase to the ICER. The ERG notes that if the marketing authorisation specifies that retreatment is not permitted then the relevant costs for the model are those treatment patients would receive in practice, ie subsequent treatment costs, rather than axi-cel drug acquisition costs and therefore this sensitivity analysis may be considered conservative from a cost perspective.*

Clinical expert advice to the ERG confirmed that retreatment was unlikely to happen in practice at least in the short term.

Current 4L+ care costs

There is no single established treatment for patients who have received 3 or more lines of treatment for r/r FL. To estimate the cost of current treatment a weighted average basket of treatments was included based on the treatments patients received in SCHOLAR-5 adjusted to reflect treatments approved for routine use in England (see Table 15 above). Wastage was included for treatments administered intravenously. For oral treatments (lenalidomide and prednisolone) the most efficient pack size was included based on the dosing schedule. Administration costs were costed using NHS reference costs according to the complexity of the procedure with oral administration assumed to incur no costs (see CS, document B tables 43 and 44). No time on treatment data are available from SCHOLAR-5 to estimate treatment durations in the model and as such treatment durations were based on the median treatment durations reported in relevant SmPCs and assumed exponential time on treatment curves were assumed to the estimated treatment durations.

Clinical expert advice to the ERG confirmed the range and proportions of treatments included for current 4L+ are broadly reasonable and likely to reflect the treatments patients receive in practice. Stem cell transplant is not included as a treatment option, and this was considered appropriate. However, the adjustments made to better reflect treatment proportions used in practice may impact on the clinical effectiveness estimates of current 4L+ care as described in section 4.2.6. The adjustment to exclude idelalisib may work in favour of the comparator arm with an arguably more effective treatment efficacy being included without the cost. Conversely, the re-weighted proportions result in higher proportions of higher cost drugs obinutuzumab and lenalidomide being included in the costings but without any corresponding adjustment for efficacy. The direction of any bias as a result of these adjustments is unclear but on balance the ERG consider any bias to be in favour of current 4L+care. Another source of uncertainty is the use of the median time on treatment from the SmPCs which results in patients receiving current 4L+ treatments beyond progression. Clinical expert advice to the ERG indicates this would not occur in practice, as

patients would stop treatment upon progression. The ERG's preferred base case, therefore, assumes patients on current 4L+ treatments receive treatment until progression, reducing the cost of the comparator arm.

Subsequent treatment costs

The approach taken to model subsequent treatment costs is similar to that outlined above for current 4L+ care. On the basis that there is no established standard of care at this stage of the treatment pathway, it was assumed that the distribution of subsequent therapies is equal in both the axi-cel and current 4L+ care arms of the model. This is applied using a one-off subsequent treatment cost at the point of progression of £45,040.02 and administration cost of £10,131.55.

The ERG notes the simplifying assumption made that subsequent therapy costs are equal in both arms of the model and considered this may not be appropriate particularly as the model estimates post-progression survival benefit with axi-cel. Furthermore, since the comparator 4L+ care costs are recycled to approximate the costs of subsequent therapy, and the company's approach to modelling current 4L+ care costs allow for treatment beyond progression, this approach will likely overestimate subsequent treatments costs. However, if time on current 4L+ treatment is capped at PFS, then the approximated cost of subsequent treatment drops accordingly. While the exact cost of subsequent therapy is uncertain, the clinical expert advice to the ERG suggested that it is not unreasonable to assume equal subsequent costs between the arms.

Health-state unit costs and resource use

Health-state resource use was applied in the model to be consistent with previous FL NICE submissions and relevant clinical guidelines. Costs were applied to the pre-progression and progressed disease health states, with pre-progression further split into induction and maintenance phases. Resources included haematologist vists, diagnostic tests and CT scans. For axi-cel, the duration of the induction phase is 6 cycles followed by maintenance until year 5. Beyond year 5, patients who are alive and progression-free in the axi-cel arm (long-term survivors) are assumed to require no further resource use. For current 4L+ care, the duration of the induction phase was

7 cycles based on a weighted average of the treatments included. The health state resource use costs applied in the model are summarised in table 18.

Table 18: Summary of health state resource use assumptions (adapted from Table 47, Document B of the CS)

Resource use	Pre-progression (induction)	Pre-progression (maintenance)	Progressed disease
Haematologist visit	1 every 1 months	1 every 3.5 months	1 every 4 weeks
Diagnostic tests	1 every 1 months	1 every 3.5 months	1 every 4 weeks
CT scans	1 every 6 months	1 every 12 months	0
Total cost/cycle	**£171.20**	**£52.85**	**£152.82**
Key:CT, computerised tomography Cost source: NHS reference costs 2019/20[41]			

The resource use costs appear low but are largely consistent with those accepted in TA627 and have been validated by the ERG clinical expert. One source of uncertainty relates to the assumption that long-term survivors require no further monitoring beyond year 5. Clinical advice to the ERG suggests practice is variable with respect to long-term follow up, and at clarification the company included the cost of a GP visit every 6 months, which had minimal impact on the ICER. However, this remains a source of uncertainty as it may be that ongoing consultant visits are more realistic which would incur a higher cost. It was also noted that haematologist visits were costed assuming non-face-to-face attendance (£95.66), whereas TA627 used the cost of a face-to-face attendance (£171.18).[21] It is likely the non-face-to-face cost was applied on the assumption that virtual appointments are more likely during the COVID-19 pandemic, but clinical advice confirmed it is more appropriate to assume in person attendance in the model, particularly for progressed patients who would be receiving ongoing treatment.

Adverse event and end-of-life care costs

Most adverse event costs were applied as one-off costs in the first model cycle as a simplifying assumption. For axi-cel treated patients, it was assumed that the treatment-related monitoring and hospitalisation costs included the cost of managing most adverse events. An additional bed day cost was included for all patients experiencing grade ≥3 AE ([redacted]). Additional costs were also included for managing hyopgammaglobulinaemia and CRS.

The cost of intravenous immunoglobulin (IVIG) to treat hypogammaglobulinaemia was included for a proportion of patients ([redacted]). As treatment for this adverse event is ongoing, costs were applied to pre-progression patients for a duration of 12 months. This is consistent with the assumptions applied in TA677 and TA559. The weighted average cost applied was [redacted] per model cycle.

In ZUMA-5 [redacted] of patients required tocilizumab to manage CRS and this cost is included in the model [redacted]). In addition, patients experiencing grade 3/4 CRS ([redacted]) are assumed to be managed in intensive care, which is consistent with the costing approach taken in TA559 and TA677. A daily ICU cost of £1,508.65 was used based on a weighted average of the costs for supporting one or two organs. Length of stay was assumed to be 4 days to be consistent with TA559 and TA677 resulting in a grade 3/4 CRS cost of [redacted]. The total cost of CRS management included in the model is [redacted]. For current 4L+ care adverse events only those experienced by 5% or more of ZUMA-5 patients were included using rates reported in TA627 weighted by the treatments received in current practice. This was considered a conservative assumption.

Finally, the cost of end-of-life care was included as a one-off cost of £6,636.83 applied upon death. This was estimated from an average cost from the Round et al (2015) study which has been used in a number of submissions to NICE.[42]

The ERG considers the approach to modelling adverse events is generally appropriate and consistent with that used in other NICE appraisals.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The company's base case ICER at the time of the main submission is outlined in Table 19. With the PAS price applied for axi-cel, and publicly available prices applied for the comparator therapies, the ICER is £48,272. axi-cel is associated with an incremental cost of [redacted] for an incremental QALY gain of [redacted] over current 4L+ therapies. A confidential appendix will be provided for the committee, which includes confidential price discounts available for comparator and subsequent treatments.

Figure 27 and Figure 28 of the company submission provide graphical representations of the Markov trace for axi-cel and current 4L+ care respectively. The Excel model provides further breakdowns of the incremental cost and QALYs. The majority of the QALY gain results from increased time spent in the progression free state. However, there is also a substantial modelled life-year gain for axi-cel in the progressive disease state, inferring that those treated with axi-cel can be expected to survive for longer following progression compared to those who progress on current 4L+ therapies. With respect to the incremental cost, this is driven primarily by the additional drug acquisition costs for the index line of therapy in the model. axi-cel is associated with a saving in subsequent treatment costs (due to delayed/averted progression), a modest increase in adverse event costs and other HCRU costs, and slightly lower discounted end of life costs.

Table 19 Company base case deterministic results (with PAS for axi-cel), adapted from Table 55, Document B of the CS)

Technologies	Total costs (£)	Total LYG*	Total QALYs	Δ costs (£)	Δ LYG*	Δ QALYs	ICER (£/QALY)
Current 4L+ care	[redacted]	[redacted]	[redacted]	-	-	-	-
Axi-cel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£48,272

Key: 4L, fourth line; Δ, incremental; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year. *Life-years undiscounted.

5.2 Company's sensitivity analyses

The company provided the results of probabilistic sensitivity analysis in Table 56 of their submission document B. The results are reproduced in Table 20 below. The incremental cost is very similar to the deterministic result, but the incremental QALY is slightly lower, resulting in a modest increase in the ICER. The company provide some further analysis which indicates that this difference is attributable to the asymmetric uncertainty surrounding correlated survival analysis parameters.

Table 20 Company base case probabilistic results (with PAS for axi-cel), adapted from Table 56, Document B of the CS)

Technologies	Total costs (£)	Total QALYs	Δ costs (£)	Δ QALYs	ICER (£/QALY)
Current 4L+ care	[redacted]	[redacted]	[redacted]	[redacted]	-
Axi-cel	[redacted]	[redacted]	[redacted]	[redacted]	£51,990
Key: 4L, fourth line; Δ, incremental; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year. *Life-years undiscounted.					

The company provide results of one-way sensitivity analysis in Figure 32 and Table 57 of their submission. Their base case ICER was most sensitive to variation in the proportion of the axi-cel treated patients considered long-term survivors, followed by the utility value for the progression free state and the utility value for the progressive disease state. Hospital length of stay for axi-cel treatment, and the percentage requiring immunoglobulin, were also both relatively important.

With respect to scenario analyses conducted by the company, covering structural uncertainties and assumptions, these are provided in Table 58 of the company submission (document B). The ICER was relatively sensitive to assumptions around long-term survivorship; both the assumed proportion it applies to and the timepoint from which it applies. Capping the time on comparator 4L+ treatments to progression free survival, rather than overall survival, also had a modest upward impact on the ICER. *The ERG is of the opinion that the latter assumption is more appropriate based on its clinical advice. It is also more consistent with the assumption that all patients who are assumed to progress receive a one-off subsequent treatment costs in line with the modelled 4L+ comparator costs*.

The company provided limited exploration of alternative OS curve extrapolations for axi-cel and current 4L+ care. This focussed on the log-logistic as a more optimistic alternative for axi-cel, and the exponential as a more pessimistic alternative for current 4L+ care. Since curves for axi-cel were fitted to the whole mITT population of ZUMA-5 but used only for the extrapolation of OS of non-long-term survivors beyond 5 years, the ERG requested some further scenarios that applied higher risks of progression and death after five years for those not considered to be long-term survivors. The company provided this by applying SMRs of 1.09 and 1.2 to their preferred PFS and OS curves for axi-cel after 5 years, which had a modest upward impact on the ICER for axi-cel. They also provided a further scenario whereby they allowed the proportion of long-term survivors to update over time based on the split progression/survival assumptions. This resulted in a modest reduction in the ICER. Finally, the company also provided additional scenarios that applied axi-cel re-treatment costs as observed in ZUMA-5, reduced subsequent treatment costs by set percentages, and included some ongoing monitoring costs for long-term survivors beyond 5 years. The results of all the additional scenarios provided by the company in response to the clarification letter are replicated in Table 21 below.

Table 21 Further deterministic cost-effectiveness scenario results provided by the company's clarification response [source: Tables 3, 5, 6, 7 and 8 of the company's clarification response].

Setting	Base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
Base case			[redacted]	[redacted]	£48,272	N/A
Increase risks of progression and death from 5 years for non-long-term survivors	Use PFS and OS curves fitted to whole mITT population of ZUMA-5.	Apply SMR of 1.09 to selected PFS and OS curve	[redacted]	[redacted]	£50,087	£1,814
		Apply SMR of 1.2 to selected PFS and OS curve	[redacted]	[redacted]	£52,326	£4,054
Dynamic updating of surviving proportion that are long-term survivors	Apply a static/fixed proportion of long-term survivors	Allow the proportion that are long-term survivors (in progression free and progressive disease states) to increase over time.	[redacted]	[redacted]	£46,105	-£2,168
Reduce subsequent treatment costs given lower expectations for PFS and OS in subsequent lines of therapy	Recycle total expected 4L+ care costs as one-off cost applied to progressed patients	Reduce subsequent treatment costs by 25%	[redacted]	[redacted]	£49,177	£905
		Reduce subsequent treatment costs by 50%	[redacted]	[redacted]	£50,081	£1,809
Regular 6 monthly GP visit applied to	No follow-up of long-term survivors from 5 years.	100%	[redacted]	[redacted]	£48,321	£48
		50%	[redacted]	[redacted]	£48,296	£24

Setting	Base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
percentage of long-term survivors		25%	[redacted]	[redacted]	£48,284	£12

5.3 Model validation and face validity check

In section B.3.10.1 of their submission document, the company describe quality assurance checks conducted on the model prior to submission. The ERG has similarly conducted its own consistency checks, using a combination of formula checking and black box tests suggested by Tappenden and Chilcott.[43] The results of the black-box tests are summarised in Table 22. No major issues were identified. Issues relating to structural inconsistencies and other uncertainties have been covered in the preceding sections.

A greater challenge is validating the survival projections produced by the model. The company acknowledge the immaturity of the PFS and OS data for the ZUMA-5 mITT population, which makes it challenging to extrapolate and validate the absolute and relative survival gains for axi-cel. There is further uncertainty regarding the long-term survivor assumptions applied in the model, and the use of the parametric PFS and OS curves (fitted to the whole mITT cohort of ZUMA-5) to model outcomes for only the non-long-term survivors from 5 years. There is potential with these assumptions to overestimate survival for axi-cel treated patients, particularly the non-long-term survivor proportion. It is worth further noting that the company base case does in fact project a substantial post progression survival gain for axi-cel, which could in part be down to unrealistic survival assumptions being applied to non-long-term survivors. However, there are plausible reasons why the introduction of axi-cel could confer a post progression survival benefit, including ongoing benefits of the CAR T-cells after progression, and the fact that it represents an additional treatment in the pathway, meaning that patients will have more of the current options available to them following progression than those in the comparator arm. On the latter point, however, it should be noted that patients in the axi-cel arm are not assumed to incur any increase in subsequent treatment costs compared to those who progress following treatment with current 4L+ therapies.

With the respect to the current 4L+ comparator, the company acknowledge the limitations of SCHOLAR-5 data for informing expected OS and PFS due to the substantial proportions that received idelalisib or experimental treatments that are not available routinely in England. The company also note that based on clinical feedback, patients with r/r FL are generally not expected to survive beyond 3 years

when treated with available 4L+ options in England. The modelling based on extrapolation of the survival data for SCHOLAR-5 does not appear to support this, despite fairly pessimistic parametric curves being selected, which suggests it may be overestimating OS compared what might be expected in the NHS in England. The ERG broadly agrees that there is potential for SCHOLAR-5 data to overestimate survival for the current 4L+ care arm, but it is difficult to verify this without actual data that is more applicable the NHS setting.

Given the above, there is considerable uncertainty surrounding the magnitude of the projected survival gain for axi-cel versus current 4L+ care.

Table 22 Summary of "black box" checks of the model carried out by the ERG

Model component	Model test	Unequivocal criterion for verification	Issues identified in company model
Clinical trajectory	Set relative treatment effect (odds ratios, relative risks or hazard ratios) parameter(s) to 1.0 (including adverse events)	All treatments produce equal estimates of total LYGs and total QALYs	Equalised the survival curve parameters on the 'PSM inputs' sheet, switched all survival curves to the exponential distribution, removed the long term survivorship assumption and equalized the QALY decrement for adverse events. This led to equal QALY and LYG for the treatment arms.
	Sum expected health state populations at any model timepoint (state transition models)	Total probability equals 1.0	No issues found.
QALY estimation	Set all health utility for living states parameters to 1.0	QALY gains equal LYGs	No issues found.
	Set QALY discount rate to 0	Discounted QALYs = undiscounted QALYs for all treatments	No issues found.
	Set QALY discount rate equal to very large number	QALY gain after time 0 tend towards zero	No issues found
Cost estimation	Set intervention costs to 0	ICER is reduced	No issues found. Incremental costs behave as expected.

Model component	Model test	Unequivocal criterion for verification	Issues identified in company model
	Increase intervention cost	ICER is increased	No issues found.
	Set cost discount rate to 0	Discounted costs = undiscounted costs for all treatments	No issues found.
	Set cost discount rate equal to very large number	Costs after time 0 tend towards zero	Minimal effect on the axi-cel arm as drug acquisition costs are applied in the first cycle.
Input parameters	Produce n samples of model parameter m	Range of sampled parameter values does not violate characteristics of statistical distribution used to describe parameter.	Sample tested. No issues found.
General	Set all treatment-specific parameters equal for all treatment groups	Costs and QALYs equal for all treatments	Not possible as several cost inputs are calculated as a one-off cost in the first cycle. Given the first test of clinical trajectory found no issues there is no concern.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook a number of further scenario analyses to address uncertainties it believes the company had not fully explored. These are outlined below, with results provided in Table 23.

Given the uncertainty surrounding the long-term extrapolations of PFS, and in particular OS (section 4.2.6 above), the ERG undertook further scenario analysis around the choice of parametric survival curves for axi-cel and current 4L+ care (scenarios 1-4).

Further, due to the uncertainty arising from using curves fitted to PFS and OS data from the whole mITT cohort of ZUMA-5, to extrapolate only for non-long-term survivor from 5 years (section 4.2.6 above), the ERG extended the company's scenarios that inflate the hazard of the extrapolated progression and mortality from 5 years (Scenario 5 below).

Noting a possible anomaly in the model with respect to the long-term mortality risk of non-long-term survivors falling below that of long-term survivors (section 4.6), the ERG implemented a fix to cap OS for non-long-term survivors to that of long-term survivors (i.e. the SMR adjusted general population mortality) – Scenario 6 below.

To further explore the possibility of longer-term secondary care-based follow-up of long-term survivors (Section 4.8), the ERG explored the impact of applying the cost of haematology follow-up every 12 months beyond year 5 (Scenario 7).

To explore the possibility of patients treated with axi-cel having more untried treatment options available to them following recurrence, and surviving for longer in the progressive disease state, the ERG assessed the impact on reducing subsequent treatment costs following progression on current 4L+ care by set percentages relative to subsequent treatment costs following axi-cel (scenario 8).

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The results of the additional scenario analyses outlined in section 6.1 are presented in Table 23 below. The greatest upward uncertainty in the ICER for axi-cel arises form more optimistic extrapolations of OS for current 4L+ care (scenario 3); more pessimistic extrapolation of OS for non-long-term axi-cel survivors (scenario 5); and relative increases in the cost of subsequent treatment for those who progress on axi-cel versus those who progress on current 4L+ care (scenario 8). The ICER for axi-cel is reduced somewhat with the selection of the more pessimistic exponential extrapolation of OS for current 4L+ care (scenario 3), and more optimistic extrapolation of PFS for axi-cel (scenario 2).

Table 23 Results of the ERG's further scenario analysis around the company base case

Setting	Company base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
Base case			[redacted]	[redacted]	£48,272	N/A
1. PFS extrapolation (4L+)	Exponential	Generalised gamma	[redacted]	[redacted]	£48,357	£84
		Lognormal	[redacted]	[redacted]	£48,385	£113
2. PFS extrapolation (axi-cel)	Weibull	Generalised gamma	[redacted]	[redacted]	£46,698	-£1,574
3. OS extrapolation (4L+)	Gamma	Lognormal	[redacted]	[redacted]	£58,745	£10,473
		Weibull	[redacted]	[redacted]	£50,898	£2,626
		Exponential	[redacted]	[redacted]	£44,530	-£3,742
4. OS extrapolation (4L+)	Weibull	No plausible less optimistic alternative available when non-long term survivors modelled as a fixed proportion.				
5. Increase risks of progression or death in non-long-term survivors	SMR = 1	SMR = 1.09	[redacted]	[redacted]	£50,087	£1,814
		SMR = 1.2	[redacted]	[redacted]	£52,326	£4,054
		SMR = 1.5	[redacted]	[redacted]	£58,552	£10,280
		SMR = 2	[redacted]	[redacted]	£69,258	£20,986

Setting	Company base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
6. OS cap for non-long-term survivors	General population survival	General population SMR adjusted survival	[redacted]	[redacted]	£48,354	£82
7. Follow up of long term survivors.	No follow-up of long-term survivors from 5 years.	Assume annual haematologist visit for all	[redacted]	[redacted]	£48,331	£59
8. Costs of subsequent therapy following progression (Axi-cel)	Costs equal between arms upon progression	Costs in 4L+ arm reduced by 10%	[redacted]	[redacted]	£49,283	£1,011
		Costs in 4L+ arm reduced by 25%	[redacted]	[redacted]	£50,799	£2,527
		Costs in 4L+ arm reduced by 50%	[redacted]	[redacted]	£53,327	£5,055

6.3 ERG's preferred assumptions

Based on the critique providing the preceding sections of this report, the ERGs preferred assumptions for its base case analysis are as follows**:**

1. Given the company's approach to assuming different risks of progression and death for long-term and non-long-term survivors from 5 years, the ERG prefers the company's amendment that allows for the OS and PFS survival to be extrapolated separately from 5 years for the two groups. This allows time dependent updating of the proportion of survivors that are long-term survivors/non-long-term survivors, so that the weighted average hazard of death or progression can be accurately calculated.
2. Because the OS and PFS curves for axi-cel were fitted for the whole mITT population of ZUMA-5, but then from 5 years only used to extrapolate outcomes for non-long-term survivors, there is a risk the chosen curves will result in upward bias of PFS and OS for this group. The ERG, therefore, believes that a downward adjustment should be applied to the PFS and OS extrapolation curves from 5 years when the modelled hazards are a split by long-term survival status. The ERG, therefore, applies an SMR of 1.2 to the chosen curves from 5 years. Accepting that the chosen SMR is arbitrary, further scenario analysis is conducted around this parameter from the ERG preferred base case.
3. Capping of overall survival of non-long-term survivors at SMR adjusted general population mortality, to avoid the risk of death in non-long-term survivors dropping below that of long-term survivors.
4. Capping current 4L+ care time on treatment to the selected PFS curve for current 4L+ care. This assumes that treatment can continue up to the point of progression but not beyond as assumed in the company base case. This is justified by clinical advice to the ERG and the company's approach to modelling subsequent treatment costs upon progression.
5. Lower utility values reported by *Wild et al and Pettengell et al* for the progression free and progressive disease state, to account for the fact the current population is more heavily treated and at a later stage in the disease pathway than the population considered in TA627.[5, 21, 39]

6. Retain PF health state utility from Wild et al. for long-term survivors from 5 years.[39] The company scenario using Wild et al assumes general population utility from 5 years for long-term survivors.

The cumulative effect of these changes on the ICER are illustrated in Table 24 below. Combined, they result in a modest increase in the ICER, to £56,332 per QALY gained. The results of probabilistic analysis from this alternative base case are provided in Table 25 and Figure 3 and 4.

Given remaining uncertainties related to the economic case for axi-cel, the ERG also conducted further scenario analysis around its revised base case (Table 26), including: alternative curve selections for PFS and OS (scenarios 1-4); an increased risk of mortality and progression in non-long-term survivors, above those projected by the curves fitted to the axi-cel cohort as a whole (scenario 5); relative reductions in the cost of subsequent treatment following progression on current 4L+ care compared to progression on axi-cel (scenario 6); changes to the assumed long-term survivor fraction (scenario 7); and increasing the SMR used to adjust the survival of long-term survivors relative of general population survival (scenario 8).

Table 24 ERG's preferred model assumptions

Preferred assumption	Section in ERG report	Incremental cost	Incremental QALY	Cumulative ICER £/QALY	Change from company base case
Company base-case		[redacted]	[redacted]	£48,272	
1. Time dependent updating of long-term survivor proportion from 5 years	*4.2.6*	[redacted]	[redacted]	£46,105	-£2,168
2. Increase progression and mortality risks by 20% after 5 years non-long-term survivors	*4.2.6*	[redacted]	[redacted]	£48,709	£437
3. Cap overall survival of non-long-term survivors at SMR adjusted general population mortality	*4.2.6*	[redacted]	[redacted]	£48,749	£477
4. Capping the current 4L+ time on treatment to the selected PFS curve for current 4L+ care	*4.2.8*	[redacted]	[redacted]	£54,736	£6,464
5. Apply Wild et al/Pettengell et al. utility values for progression free and progressive disease states.	*4.2.7*	[redacted]	[redacted]	£55,383	£7,111
6. Retain PF health state utility from Wilde et al. for long-term survival from 5 years	*4.2.7*	[redacted]	[redacted]	£56,332	£8,060

Table 25 ERG base case (probabilistic)

Technology	Total cost (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALYs)
Current 4L+ care	[redacted]	[redacted]			
Axi-cel	[redacted]	[redacted]	[redacted]	[redacted]	£58,773

Figure 3 Cost-effectiveness scatter-plot (ERG base case)

Figure 4 Cost-effectiveness acceptability curve (ERG base case)

Table 26 Additional scenario analysis around the ERG preferred base case

Setting	ERG Base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
Base case			[redacted]	[redacted]	£56,332	N/A
1. PFS extrapolation (4L+)	Exponential	Generalised gamma	[redacted]	[redacted]	£56,541	£209
		Lognormal	[redacted]	[redacted]	£56,550	£218
2. PFS extrapolation (axi-cel)	Weibull	Generalised gamma	[redacted]	[redacted]	£54,950	-£1,382
3. OS extrapolation (4L+)	Gamma	Lognormal	[redacted]	[redacted]	£67,765	£11,433
		Weibull	[redacted]	[redacted]	£59,171	£2,839
		Exponential	[redacted]	[redacted]	£52,383	-£3,949
4. OS extrapolation (4L+)	Weibull	Generalised gamma	[redacted]	[redacted]	£73,034	£16,702
5. Increase risks of progression or death in non-long-term survivors	SMR = 1.2 over selected PFS and OS curves	SMR = 1	[redacted]	[redacted]	£53,470	-£2,862
		SMR = 1.09	[redacted]	[redacted]	£54,797	-£1,535
		SMR = 1.5	[redacted]	[redacted]	£60,084	£3,752
		SMR = 2	[redacted]	[redacted]	£65,190	£8,858
6. Costs of subsequent therapy following progression (4L+)	Costs equal between arms upon progression	Costs in 4L+ arm reduced by 10%	[redacted]	[redacted]	£56,887	£555
		Costs in 4L+ arm reduced by 25%	[redacted]	[redacted]	£57,721	£1,389

Setting	ERG Base case	Scenario	Incremental costs	Incremental QALYs	ICER	Change from base case
		Costs in 4L+ arm reduced by 50%	[redacted]	[redacted]	£59,109	£2,777
7. Long term survivor proportion	25%	10%	[redacted]	[redacted]	£66,840	£10,508
		All who are alive and progression free at 5 years	[redacted]	[redacted]	£52,130	-£4,202
8. SMR applied to long term susrvivors	SMR = 1.09	SMR = 1.2	[redacted]	[redacted]	£57,142	£810

6.4 Conclusions of the cost effectiveness section

The company have provided robust and flexible model to assess the cost-effectiveness of axi-cel versus current 4L+ care for patients with r/r FL. The case is broadly in line with the final scope for the appraisal, although it focusses a sub-group of wider population defined. The cost-effectiveness case is inherently uncertain given the lack of a randomized comparator in the clinical data, and the immaturity of the PFS and OS data for axi-cel from the ZUMA-5 trial. The company acknowledge the uncertainty and consider that axi-cel would be a suitable candidate for cancer drug fund approval, so that current uncertainties can be addressed.

The ERG believe that company have provided a reasonable estimate of the ICER given the data available but suggest a number of changes may be justified which result in a modest increase in the ICER. The ERG believes that a number of uncertainties were not identified or fully explored in the original company submission. However, these issues have been addressed in further scenario analysis provided by the company in response to clarification letter and further scenario analysis undertaken by the ERG. The remaining areas of uncertainty that result in the greatest uncertainty in the ICER are:

1. the proportion of patients that can be considered long-term survivors following treatment with axi-cel, and the time point from which this applies.
2. the assumptions around overall survival extrapolation for those considered to be long-term survivors and those who considered to be non-long-term survivors.
3. The OS for patients treated with current 4L+ therapies available in the NHS
4. The costs of current 4L+ treatment based on time on treatment assumed, and whether this should be capped using the PFS curve from SCHOLAR-5 or allowed to continue beyond progression.
5. Related to point 4, the cost of subsequent treatment that is assumed to apply in the model, and whether it is reasonable to assume this is equal between treatment arms or that it could potentially be higher following progression on axi-cel.

7 End of life

The company make a case for axi-cel being considered an end-of-life drug for the current indication (see section B.2.13.14 of the company submission). The company claim that life expectancy in this cohort is usually approximately three years and refer to their SCHOLAR-5 data and clinical expert opinion. They acknowledge that this is longer than the 24 months stated in the NICE end of life criteria, but they also note that they believe clinicians would adopt axi-cel as an end-of-life treatment in NHS England - perhaps suggesting that clinicians would use it more judiciously in those with lower life expectancy at its 4L+ positioning. It is not in doubt that axi-cel can be expected to deliver gains in overall survival of more than three months.

The ERG acknowledges the company's case but would note that it is median overall survival in SCHOLAR-5 that is close to 3 years, rather than average life expectancy (which is unobserved). The extrapolation modelling for the company base case suggests a mean undiscounted life expectancy of [redacted] *years in the current 4L+ care arm. Given this, the end-of-life criteria is not strictly met.*

8 References

1. Cancer Research UK. How doctors group non-Hodgkin lymphomas. 2020. Available from: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/groupp (Accessed 1 February 2022).

2. Cancer Research UK. Survival [Non-Hodgkin lymphoma]. 2022. Available from: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/survival (Accessed 1 February 2022).

3. Haematological Malignancy Research Network. Factsheets: Follicular Lymphoma. 2021. Available from: https://hmrn.org/factsheets#follicular_lymphoma (Accessed 1 February 2022).

4. Cancer Research UK. Follicular lymphoma. 2020. Available from: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/follicular-lymphoma (Accessed 1 February 2022).

5. 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. 2007;**19**(3):570-6.

6. Pozo AL, Godfrey EM, Bowles KM. Splenomegaly: investigation, diagnosis and management. Blood Rev. 2009;**23**(3):105-11.

7. Blaes AH, Ma L, Zhang Y, Peterson BA. Quality of life appears similar between survivors of indolent and aggressive non-Hodgkin lymphoma. Leuk Lymphoma. 2011;**52**(11):2105-10.

8. Smith SK, Zimmerman S, Williams CS, Preisser JS, Clipp EC. Post-traumatic stress outcomes in non-Hodgkin's lymphoma survivors. J Clin Oncol. 2008;**26**(6):934-41.

9. Smith SK, Zimmerman S, Williams CS, Zebrack BJ. Health status and quality of life among non-Hodgkin lymphoma survivors. Cancer. 2009;**115**(14):3312-23.

10. Kite (a Gilead company). NHS clinical validation of the economic modelling approach for axi-cel in 4L+ relapsed/refractory follicular lymphoma (NICE ID1685). Data on File. Santa Monica, CA: Kite; 2021.

11. Cheung MC, Imrie KR, Friedlich J, Buckstein R, Lathia N, Mittmann N. 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.

12. Hernandez-Ilizaliturri F. Follicular Lymphoma (Non-Hodgkin Lymphoma) Staging. 2020. Available from: https://emedicine.medscape.com/article/2007038-overview (Accessed 1 February 2022).

13. Cancer Research UK. Treatment for advanced disease (stage 3 or 4). 2020. Available from: https://www.cancerresearchuk.org/about-cancer/non-hodgkin-lymphoma/types/follicular-lymphoma?_ (Accessed 1 February 2022).

14. Smith SM. Dissecting follicular lymphoma: high versus low risk. Hematology. 2013;**2013**(1):561-7.

15. Welaya K, Casulo C. Follicular Lymphoma: Redefining Prognosis, Current Treatment Options, and Unmet Needs. Hematol Oncol Clin North Am. 2019;**33**(4):627-38.

16. Casulo C, Byrtek M, Dawson KL, Zhou X, Farber CM, Flowers CR, et al. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. J Clin Oncol. 2015;**33**(23):2516-22.

17. National Institute for Health and Care Excellence. Rituximab for the first-line treatment of stage III-IV follicular lymphoma [TA243]. 2012. Available from: https://www.nice.org.uk/guidance/ta243 (Accessed 1 February 2022).

18. National Institute for Health and Care Excellence. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma [TA226]. 2011. Available from: https://www.nice.org.uk/guidance/ta226 (Accessed 1 February 2022).

19. National Institute for Health and Care Excellence. Obinutuzumab for untreated advanced follicular lymphoma [TA513]. 2018. Available from: https://www.nice.org.uk/guidance/ta513 (Accessed 1 February 2022).

20. National Institute for Health and Care Excellence. Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma (TA137). 2008. Available from: https://www.nice.org.uk/guidance/ta137 (Accessed 1 February 2022).

21. National Institute for Health and Care Excellence. Lenalidomide with rituximab for previously treated follicular lymphoma (TA627). 2020. Available from: https://www.nice.org.uk/guidance/ta627 (Accessed 1 February 2022)

22. National Institute for Health and Care Excellence. Obinutuzumab with bendamustine for treating follicular lymphoma after rituximab (TA629). 2020. Available from: https://www.nice.org.uk/ta629 (Accessed 1 February 2022).

23. National Institute for Health and Care Excellence. Treating follicular lymphoma. 2021. Available from: https://pathways.nice.org.uk/pathways/non-hodgkins-lymphoma/treating-follicular-lymphoma#:~:text=1%20October%202019%20Idelalisib%20for%20treating%20refractory%20follicular,604%29%20added%20to%20treating%20relapsed%20or%20refractory%20disease. (Accessed 23 August 2021).

24. Kite (a Gilead company). NHS clinical validation of the economic modelling approach for axi-cel in 4L+ relapsed/refractory follicular lymphoma (NICE ID1685). Data on File. 2021.

25. Batlevi CL, Sha F, Alperovich A, Ni A, Smith K, Ying Z, et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. Blood Cancer J. 2020;**10**(7):74.

26. Link BK, Day BM, Zhou X, Zelenetz AD, Dawson KL, Cerhan JR, 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.

27. Fuji S, Tada Y, Nozaki K, Saito H, Ozawa T, Kida T, et al. A multi-center retrospective analysis of patients with relapsed/refractory follicular lymphoma after third-line chemotherapy. Ann Hematol. 2020;**99**(9):2133-9.

28. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking systematic reviews in health care. University of York: 2009. Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm.

29. Kite (a Gilead company). A Phase 2 Multicenter Study of Axicabtagene Ciloleucel in Subjects with Relapsed/Refractory Indolent Non-Hodgkin Lymphoma: 18-Month Follow-Up Analysis. Clinical Study Report Addendum. Data on File. Santa Monica, CA: Kite; 2021.

30. Gopal AK, Kahl BS, de Vos S, Wagner-Johnston ND, Schuster SJ, Jurczak WJ, et al. PI3Kδ inhibition by idelalisib in patients with relapsed indolent lymphoma. N Engl J Med. 2014;**370**(11):1008-18.

31. Kite (a Gilead company). A Phase 2 Multicenter Study of Axicabtagene Ciloleucel in Subjects with Relapsed/Refractory Indolent Non-Hodgkin Lymphoma: 18-Month Follow-Up Analysis. Clinical Study Report Addendum. Data on File. 2021.

32. Kite (a Gilead company). A comparison of clinical outcomes from ZUMA-5 (axicabtagene ciloleucel) and the international SCHOLAR-5 external control cohort in relapsed/refractory follicular lymphoma. Non-interventional Final Study Report. Data on File. Santa Monica, CA: Kite; 2021.

33. National Institute for Health and Care Excellence. Idelalisib for treating refractory follicular lymphoma (TA604). 2019. Available from: https://www.nice.org.uk/guidance/ta604 (Accessed 1 February 2022).

34. National Institute for Health and Care Excellence. Obinutuzumab with bendamustine for treating follicular lymphoma refractory to rituximab (TA472). 2017. Available from: https://www.nice.org.uk/guidance/ta472 (Accessed 1 February 2022).

35. National Institute for Health and Care Excellence. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies (TA559). 2019. Available from: https://www.nice.org.uk/guidance/ta559 (Accessed 1 February 2022).

36. National Institute for Health and Care Excellence. Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies (TA567). 2019. Available from: https://www.nice.org.uk/guidance/ta567 (Accessed 1 February 2022).

37. National Institute for Health and Care Excellence. Autologous anti-CD19-transduced CD3+ cells for treating relapsed or refractory mantle cell lymphoma (TA677). 2021. Available from: https://www.nice.org.uk/ta677 (Accessed 1 February 2022).

38. Hettle R, Corbett M, Hinde S, Hodgson R, Jones-Diette J, Woolacott N, et al. The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal. Health Technol Assess. 2017;**21**(7):1-204.

39. Wild D, Walker M, Pettengell R, Lewis G. PCN62 Utility elicitation in patients with follicular lymphoma. Value Health. 2006;**9**.

40. Guadagnolo BA, Punglia RS, Kuntz KM, Mauch PM, Ng AK. Cost-effectiveness analysis of computerized tomography in the routine follow-up of patients after primary treatment for Hodgkin's disease. J Clin Oncol. 2006;**24**(25):4116-22.

41. National Health Service. 2019/20 National Cost Collection data. 2021. Available from: https://www.england.nhs.uk/national-cost-collection/ (Accessed 1 February 2022).

42. 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.

43. Tappenden P, Chilcott JB. Avoiding and identifying errors and other threats to the credibility of health economic models. Pharmacoeconomics. 2014;**32**(10):967-79.