Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after 1 or more multi-agent chemotherapy regimens [ID1557]

A Single Technology Appraisal

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School South Cloisters St Luke's Campus Heavitree Road Exeter EX1 2LU
Authors	**Dr Maxwell S. Barnish**, Research Fellow[1] **Mr Brian O'Toole**, Research Fellow[1] **Mr David Packman**, Graduate Research Assistant[1] **Mr Madhusubramanian Muthukumar**, Research Fellow[1] **Mr Justin Matthews,** Research Fellow[1] **Ms Naomi Shaw**, Information Specialist[1] **Dr Claudius E Rudin**, Consultant Haematologist[1,2] **Ms Louise Crathorne**, Senior Research Fellow[1] **Prof G.J. Melendez-Torres**, Professor[1] [1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter [2] Cwm Taf Morgannwg University Health Board, Wales
Correspondence to	Dr Maxwell S. Barnish 3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; m.s.barnish@exeter.ac.uk
Date completed	02/12/2020
Source of funding	This report was commissioned by the NIHR Systematic Reviews Programme as project number **13/21/84**.
Declared competing interests of the authors	None.

Acknowledgments	The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG).
Rider on responsibility for document	The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.
This report should be referenced as follows	Barnish MS, O'Toole B, Packman D, Muthukumar M, Matthews J, Shaw N, Rudin CE, Crathorne L, Melendez-Torres GJ. Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after 1 or more multi-agent chemotherapy regimens [ID1557]: A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2020.

Author Contributions:

Maxwell S. Barnish	Project managed the ERG team and critiqued the clinical effectiveness evidence.
Brian O'Toole	Critical appraisal of the economic evidence, writing and editorial input, and co-supervised the final report
David Packman	Critical appraisal of the economic evidence, checked and re-analysed the economic model, carried out further scenario analyses, and drafted economic sections of the report
Madhusubramanian Muthukumar	Critical appraisal of the economic evidence, checked and re-analysed the economic model, carried out further scenario analyses, and drafted economic sections of the report
Justin Matthews	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report
Naomi Shaw	Critical appraisal of the search methods and editorial input.
Claudius Rudin	Clinical advice and review of draft report.
Louise Crathorne	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report
G.J. Melendez-Torres	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report. Guarantor of the report.

Table of Contents

List of issues

List of tables

List of Figures

Abbreviations

AEs	Adverse events
AFT	Accelerated failure time
AIC	Akaike's information criterion
ASaT	All Subjects as Treated
ASCT	Autologous stem cell transplantation
AUC	Area under the curve
-ben	Bendamustine
BIC	Bayesian information criterion
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
BV	Brentuximab vedotin
CAA	Commercial access agreement
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptability curve
ChiVPP	Chlorambucil with vinblastin, procarbazine and prednisolone
CI	Confidence interval
CR	Complete response
CS	Company submission
DHAP	D – dexamethasone; HA – high dose Ara C, also known as cytarabine; P – cisplatin
ECOG	Eastern Cooperative Oncology Group
EORTC-QLQ-C30	European Organization for Research and Treatment of Cancer quality of life questionnaire
EQ-5D	EuroQol 5 dimensio
ERG	Evidence Review Group
ESHAP	E – etoposide; S – solu-medrone (also called methylprednisolone); HA – high dose cytarabine, also known as Ara C; P – cisplatin
ESR	Erythrocyte sedimentation rate
ESS	Effective sample size
GDP	G – gemcitabine, d - dexamethasone, C – cisplatin
-Gen Gam / -GG	Generalised gamma
HL	Hodgkin's lymphoma
HR	Hazard ratio

HRQoL	Health-related quality of life
HTA	Health technology assessment
ICE	I – ifosfamide; C – carboplatin; E – etoposide
ICER	Incremental cost-effectiveness ratio
IGEV	I – ifosfamide; G – gemcitabine; V – vinorelbine
IPD	Individual patient data
ITT	Intention to treat
IWG	International Working Group
KM	Kaplan-Meier
LY	Life year(s)
MAIC	Matched adjusted indirect comparison
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
-nivo	nivolumab
NMA	Network meta-analysis
NR	Not reported
NS	Not stated
NSCLC	Non-small-cell lung cancer
ORR	Objective response rate
OS	Overall survival
OWSA	One-way sensitivity analysis
PAS	Patient access scheme
PD	Progressed disease
PD-1	Programmed cell death protein 1
PD-L1	Programmed death ligand 1
pembro	pembrolizumab
PET	Positron emission tomography
PET-CT	Positron emission tomography and computed tomography (scan)
PFS	Progression free survival
PICOS	Population Intervention, Comparator, Outcomes, Study Design
PR	Partial response
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services

PSSRU	Personal Social Services Research Unit
QALY	Quality adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
R/R	Relapsed/refractory
R/RcHL	Relapsed/refractory classic Hodgkin's lymphoma
SCT	Stem cell transplantation
SCT-2L	Patients with R/RcHL who did not have at least two prior therapies when autologous stem cell transplant is not a treatment option
SCT+3L+	Patients with R/RcHL who are at least third line with prior stem cell transplant.
SCT-3L+	Patients with R/RcHL who are at least third line when autologous stem cell transplant is not a treatment option
SG	subgroup
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of Product Characteristics
TA	Technology appraisal
TAG	Technology Appraisal Guidance
TLR	Targeted literature review
ToT	Time on treatment
TTO	Time trade off
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
vs	Versus
WTP	Willingness to pay

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, 1.4, 1.5, and 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 ERG's key issues

A brief overview of the key issues identified by the ERG in their appraisal of the company submission (CS) is provided in Table 1. Further detail of the issues is provided in Sections 1.3, 1.4, 1.5, and 1.6.

Broadly speaking the key clinical issues related to immaturity of overall survival data, the matched-adjusted indirect comparison (MAIC) and the generalisability to UK clinical practice of the intention to treat (ITT) analyses. In terms of cost effectiveness issues, the ERG noted uncertainty surrounding the extrapolation of OS and PFS, estimation of base case utility values (particularly the PD health state), inclusion of SCT rates, assumptions relating to subsequent treatment usage and calculation of time on treatment (ToT) costs as well as health state resource use costs for the PD health state.

Table 1: Summary of key issues

ID	Summary of issues	Report sections
Key Issue 1: Immaturity of overall survival data	The immaturity of OS data in the key trial, meaning no directly observed comparative OS data were available for use in the economic model	Section 3.2.5.1
Key Issue 2: How reliable is the comparison of pembrolizumab with standard of care made by the MAIC	The matched adjusted indirect comparison (MAIC) analysis was only conducted with regard to one potential 2L salvage chemotherapy regimen (IGEV) and is therefore not generalizable to the full range of	Section 3.3

for the SCT-2L subgroup?	regimens used in clinical practice in the UK	
Key Issue 3: Generalisability of the intention to treat (ITT) population to UK clinical practice	The intention to treat (ITT) analysis is not generalizable to the UK treatment pathway, since there are three clear subgroups (SCT-2L, SCT-3L+ and SCT+3L+), not all of which have BV as a relevant comparator.	Section 3.2.1
Key Issue 4: Uncertainty in PFS estimation in the SCT-2L subgroup	There are no head-to-head data comparing pembrolizumab to chemotherapy within this subgroup. The company has therefore conducted a MAIC to estimate clinical effectiveness.	Section 4.2.6.2
Key Issue 5: Uncertainty in the maintenance of PFS benefit associated with pembrolizumab after treatment discontinuation in Year 2	The incremental QALY gain associated with pembrolizumab was driven by the difference in PFS between treatments. A key assumption (which is applied in all subgroups) is that after treatment discontinuation (Year 2), PFS will not be affected i.e. the proportion of patients in the PFS health state will continue to follow the chosen extrapolation curve over time.	Sections 3.2.5.2 and Section 6.2.1.3
Key Issue 6: Utility values used in the progressed disease (PD) health state for pembrolizumab	There is uncertainty surrounding the base case pembrolizumab PD health state utility value, which appears to lack plausibility.	Sections 4.2.7 and 6.2.1.1
Key Issue 7: Uncertainty in subsequent treatments and assumed proportions in the company's base case analysis	There is uncertainty surrounding the company's base case assumptions with respect to subsequent treatment usage.	Sections 4.2.8.3 and 6.2.1.13
Key Issue 8: Gopal et al. (2015) should not be used as the primary source of OS for all subgroups	It was assumed that OS from Gopal et al. (2015)[1] was generalisable to all subgroups. However, given that patients in Gopal et al. (2015), were those who had a prior SCT (reflecting the SCT+3L+ subgroup) there was some concern surrounding the generalisability of OS estimates to the subgroups.	Section 3.2.5.1
Key Issue 9: Time on treatment (ToT) for BV in SCT-3L+ and SCT+3L+ subgroups	The company assumed that patients treated with BV will receive the same maximum ToT as pembrolizumab (35 cycles). However, as per the SmPC for BV, treatment should be provided for a maximum of 16 cycles.	Section 4.2.8.2

Abbreviations: BV, brentuximab vedotin; ITT intention to treat; MAIC, matched adjusted indirect comparison; OS, overall survival; PD, progressed disease; PFS, progression free survival; SCT, stem cell transplant; ToT, time on treatment

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are outlined in Table 2.

Table 2: Key differences between the company's preferred assumptions and ERG's preferred assumptions

	Company's preferred assumption	ERG preferred assumption	Report Sections
Population	The company has presented an ITT analysis as the base case for consideration (with subgroup analyses results provided for information)	The ERG preferred to individually appraise each subgroup.	Sections 4.2.3 and 4.2.4
OS	The company prefer to use one clinical study (Gopal et al., 2015)[1] to estimate OS for all subgroups.	The ERG preferred to use Balzarotti et al. (2016)[2] for SCT-2L and SCT-3L+, and Gopal et al. (2015)[1] for SCT+3L+	Section 4.2.6.1
PFS	The company preferred to model PFS using a 52-week cut point (ITT population and SCT-3L subgroups).	The ERG preferred to model PFS using a 26-week cut point.	Section 4.2.6.2
Utilities	The company prefer to use treatment specific QoL data from KEYNOTE-204[3] to estimate both the PFS and PD health state utilities.	The ERG preferred to assume no difference in PD utility between treatments (applying the same value to both treatment arms).	Section 4.2.7
ToT	The company preferred to model ToT using an 80-week cut point.	The ERG preferred to model ToT using a 26-week cut point.	Section 4.2.8.2
Maximum number of treatment cycles	The company preferred to assume that BV would require a similar maximum number of treatment cycles to pembrolizumab (35 cycles).	The ERG preferred the SmPC estimate of a maximum of 16 cycles to be used for BV.	Section 4.2.8.1
SCT rates	The company preferred to use SCT rates from KEYNOTE-204.[3]	The ERG preferred to remove differences in SCT rates between treatments from the model.	Section 4.2.8.4

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ITT intention to treat; OS, overall survival; PD, progressed disease; PFS, progression free survival; QoL, quality of life; SCT, stem cell transplant; SmPC, summary of product characteristics; ToT, time on treatment

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:

- Improving the quality of life of patients in both the PFS and PD health states. The model estimates that patients receiving pembrolizumab have a higher utility value in both the PFS and PD states compared to the comparator (BV). The incremental QALY gain associated with pembrolizumab is therefore due to a higher proportion of patients remaining progression free and the associated higher quality of life with being in both the PFS and PD health state, versus the comparator.
- Increasing the proportion of patients in the PFS health state. The model estimated a higher proportion of patients on pembrolizumab would remain progression free compared to those receiving the comparator treatment (brentuximab vedotin [BV]).
- The ERG noted that the model does not estimate pembrolizumab to have an effect on OS, compared to the comparator treatment (BV). Due to the OS modelling approach adopted by the company, whereby a single OS curve was assumed to apply to both treatments, pembrolizumab does not result in an incremental life year (LY) gain versus the comparator treatment.

Overall, the technology is modelled to affect costs by:

- Lowering medicine acquisition costs, compared to BV, in ITT, SCT-3L+ and SCT+3L+ subgroups. The model therefore assumes that, at list price, pembrolizumab as a treatment strategy will be cheaper than BV.
- Including a two-year stopping rule for pembrolizumab which assumes that patients do not continue on treatment after this time point. Treatment costs are therefore capped at two years in the model.
- Subsequent treatment usage. Modelled results are sensitive to subsequent treatment assumptions.

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

- Base case utility values.
- The distribution of subsequent treatments, which may vary between clinical practice, treatments that are relevant for this appraisal (e.g. CDF-only treatments) and trial data. For SCT-3L+, pembrolizumab is positioned as a subsequent treatment but is a CDF-only drug and is thus not routinely commissioned.
- The assumption of a long-term PFS benefit for pembrolizumab, in interaction with utility values. A key model assumption relates to the maintenance of pembrolizumab treatment benefit (with respect to PFS state membership) over time, despite treatment discontinuation at Year 2.

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

The ERG reviewed the approach of the company to addressing the NICE decision problem for this appraisal, and identified no key issues with the decision problem.

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

The ERG reviewed the clinical effectiveness and safety evidence presented in the CS, and identified the following key issues for consideration by the committee.

Key Issue 1: Immaturity of overall survival data

Report sections	**3.2.5.1**
Description of issue and why the ERG has identified it as important	No mature OS data were provided from the pivotal KEYNOTE-204[3,4] trial since median OS had not been reached. This meant that no directly comparative OS data for pembrolizumab and BV were available to inform the economic model.
What alternative approach has the ERG suggested?	The ERG conducted additional scenario analyses using OS data from published studies including KEYNOTE 087,[5] Balzarotti et al. (2016)[2] and Gopal et al. (2015)[1]
What is the expected effect on the cost-effectiveness estimates?	The impact of these scenario analyses on the ICER was minimal, given that the same data are used to model OS for both pembrolizumab and comparator treatment arms (see Section 3.2.5.1).
What additional evidence or analyses might help to resolve this key issue?	Mature OS data from KEYNOTE-204[3,4] will be key to resolving this uncertainty.

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; OS, overall survival

Key Issue 2: How reliable is the comparison of pembrolizumab with standard of care made by the MAIC for the SCT-2L subgroup?

Report sections	3.3
Description of issue and why the ERG has identified it as important	The company carried out unanchored MAIC for SCT-2L subgroup for pembrolizumab vs salvage chemotherapy. However, this analysis is susceptible to bias arising from any missing prognostic factors or effect modifiers and is limited by a small effective sample size and the inclusion of only one salvage chemotherapy regimen.
What alternative approach has the ERG suggested?	The ERG has not carried out additional MAIC analyses given the limitations of the analysis and the available data.
What is the expected effect on the cost-effectiveness estimates?	The impact on cost-effectiveness estimates is uncertain.
What additional evidence or analyses might help to resolve this key issue?	An analysis that draws on a richer data set with larger sample size, for example routinely collected data, may produce a more robust analysis and resolve remaining uncertainty in the impact of pembrolizumab as compared to salvage chemotherapy regimens.

Abbreviations: ERG, Evidence Review Group; MAIC, matched adjusted indirect comparison; SCT, stem cell transplant; SoC, standard of care

Key Issue 3: Generalisability of the intention to treat (ITT) population to UK clinical practice

Report sections	3.2.1
Description of issue and why the ERG has identified it as important	The company presented intention to treat (ITT) results from KEYNOTE-204[3,4] as the primary clinical effectiveness data to inform its economic model. The ITT analysis included SCT-2L, SCT-3L+ and SCT+3L+ patients. These three patient groups do not have a common comparator – since salvage chemotherapy is the relevant comparator for the SCT-2L group and BV is the relevant comparator for the other 2 groups. This means that the ITT population does not generalise to the UK treatment pathway in clinical practice.
What alternative approach has the ERG suggested?	Due to the concern surrounding the plausibility of an overall ITT population, the ERG was of the opinion that each subgroup should be assessed individually.
What is the expected effect on the cost-effectiveness estimates?	The company has provided cost effectiveness results for each subgroup. The ICER presented for each subgroup differs to the ITT ICER due to

	differences in comparator, clinical effectiveness data and subsequent treatment usage.
What additional evidence or analyses might help to resolve this key issue?	Additional clinical advice to confirm the generalisability of the trial and its subgroups to UK clinical practice would resolve uncertainty. In addition, clinical evidence targeted at subgroups relevant to UK clinical practice (e.g. for SCT-2L) would reduce uncertainty about generalisability.

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention to treat; OS, overall survival; PFS, progression free survival

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

The ERG reviewed the economic model and cost-effectiveness evidence presented in the CS, and identified the following key issues for consideration by the committee.

Key Issue 4: Uncertainty in PFS estimation in the SCT-2L subgroup

Report sections	**4.2.6.2**
Description of issue and why the ERG has identified it as important	There are no head-to-head data comparing pembrolizumab to chemotherapy within this subgroup. The company has therefore conducted a MAIC to estimate clinical effectiveness. The ERG noted that the PFS benefit associated with pembrolizumab is being driven by an imprecise HR, due to the small sample size of patients within the MAIC. There is considerable uncertainty surrounding the pembrolizumab treatment effect within this subgroup.
What alternative approach has the ERG suggested?	The ERG noted that although the clinical effectiveness results are highly uncertain, the company appears to have used best available evidence to generate treatment effect for this subgroup. The ERG acknowledged that a scenario analysis which removes the pembrolizumab PFS benefit could be conducted. However, given that a conservative assumption has already been adopted by the company with respect to OS modelling, this scenario would be considered overly pessimistic.
What is the expected effect on the cost-effectiveness estimates?	A scenario which assumed no difference in PFS between treatments would result in a cost minimisation analysis, given that the incremental QALY gain associated with pembrolizumab stems from improved PFS alone. However, pembrolizumab would not be considered a cost

	saving treatment in this scenario. The ERG did not consider this to be a plausible scenario.
What additional evidence or analyses might help to resolve this key issue?	Conducting a cost minimisation analysis would address uncertainty surrounding the long-term benefit of pembrolizumab with respect to PFS, however, the scenario analysis lacks validity. Therefore, the ERG considered that the issue should be noted as an area of significant uncertainty and that the results for the SCT-2L subgroup should be interpreted with caution. Additional, more robust clinical evidence considering the range of salvage chemotherapies and additional sources of real-world data would assist in resolving this uncertainty.

Abbreviations: ERG, Evidence Review Group; HR, hazard ratio; MAIC, matched adjusted indirect comparison; OS, overall survival; PFS, progression free survival

Key Issue 5: Uncertainty in the maintenance of PFS benefit associated with pembrolizumab after treatment discontinuation in Year 2

Report sections	**3.2.5.2 and 6.2.1.3**
Description of issue and why the ERG has identified it as important	The ERG noted that the incremental QALY gain associated with pembrolizumab was driven by the difference in PFS between treatments. A key assumption (which is applied in all subgroups) is that after treatment discontinuation (Year 2), PFS will not be affected i.e. the proportion of patients in the PFS health state will continue to follow the chosen extrapolation curve over time. The ERG considered this assumption to be highly uncertain, given a lack of long-term clinical effectiveness data supporting this assumption.
What alternative approach has the ERG suggested?	Clinical opinion to the ERG has noted that it may be plausible for some patients to continue receive PFS benefit after stopping treatment, however the extent of this benefit in terms of duration is not clear. The ERG requested the company provide a scenario analysis which incorporated a waning in pembrolizumab treatment effect from Year 3, until no difference was assumed between treatments in Year 5. The company did not provide this analysis citing a lack of precedent for this type of scenario and that a conservative approach had already been adopted in the base case analysis with respect to OS. As an exploratory analysis the ERG has conducted this scenario.

What is the expected effect on the cost-effectiveness estimates?	This scenario analysis resulted in an increased ICER for pembrolizumab, given that PFS, in interaction with utility values, is a driver of the incremental QALY gain within the model.
What additional evidence or analyses might help to resolve this key issue?	Longer term data are required to address uncertainty surrounding maintenance of treatment effect.

Abbreviations: ERG, Evidence Review Group; OS, overall survival; PFS, progression free survival; QALY, quality adjusted life year

Key Issue 6: Utility values used in the progressed disease (PD) health state for pembrolizumab

Report sections	**4.2.7 and 6.2.1.1**
Description of issue and why the ERG has identified it as important	The ERG considered utility values were uncertain due to the following; • Small patient numbers and limited QoI data collection with respect to the estimation of PD values. • Clinical opinion to the ERG, outlined that the value used in the pembrolizumab PD health state was somewhat high and lacked face validity. Furthermore, patients in this health state have a higher quality of life than those on BV, who are progression free. Due to the uncertainty surrounding the pembrolizumab PD utility value, the incremental QALY gain associated with pembrolizumab appears to be overestimated.
What alternative approach has the ERG suggested?	The ERG conducted a scenario analysis that applies the BV PD utility value ([redacted]) to both treatment arms. See Section 4.2.7.
What is the expected effect on the cost-effectiveness estimates?	This scenario analysis resulted in a reduction in incremental QALYs for pembrolizumab (and increased ICER).
What additional evidence or analyses might help to resolve this key issue?	Additional data and more robust estimation of utility values post-progression, alongside a clear clinical rationale for differential utilities post-progression, would assist in resolving this uncertainty.

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; PD, progressed disease; QALY, quality adjusted life year; QoL, quality of life

Key Issue 7: Uncertainty in subsequent treatments and assumed proportions in the company's base case analysis

Report sections	4.2.8.3 and 6.2.1.13
Description of issue and why the ERG has identified it as important	• The ERG did not consider an ITT population to be appropriate for decision making therefore the subsequent treatments and proportions used for this analysis should be interpreted with caution. • For the SCT-3L+ subgroup, the company assumed that patients who failed on BV go on to receive pembrolizumab. The ERG noted that pembrolizumab is a CDF treatment, therefore it is not routinely commissioned and this assumption is not appropriate. • For the SCT+3L+ subgroup, the company assumed that 100% of patients who failed on pembrolizumab go on to receive BV. However, the ERG understood that nivolumab is the most appropriate subsequent treatment for use. Therefore, the company's base case assumption potentially underestimates costs for pembrolizumab. ERG preference for Nivolumab as subsequent therapy in this subgroup was based on the current treatment pathway, however clinical opinion to the ERG noted that BV could potentially be used. • There were some discrepancies between modelled subsequent treatments and those reported in the CS.
What alternative approach has the ERG suggested?	The ERG undertook scenario analyses using alternative subsequent treatment assumptions. See Section 6.2.1.13.
What is the expected effect on the cost-effectiveness estimates?	Altering subsequent treatments had a substantial impact on the subgroup results, resulting in increased ICERs for pembrolizumab.
What additional evidence or analyses might help to resolve this key issue?	In subgroups where subsequent treatments are poorly understood, routinely collected data could inform more realistic assumptions.

Abbreviations: BV, brentuximab vedotin; CDF, Cancer Drugs Fund; CS, company submission; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention to treat

Key Issue 8: Gopal et al. (2015) should not be used as the primary source of OS for all subgroups

Report sections	3.2.5.1
Description of issue and why the ERG has identified it as important	It is assumed that OS from Gopal et al. (2015)[1] is generalisable to all subgroups. However, given that patients in Gopal et al. (2015), were those had a prior SCT (reflecting the SCT+3L+ subgroup) there was some concern surrounding the generalisability of OS estimates to the subgroups. Furthermore, based on clinical opinion to the ERG (and clinical opinion provided to the company), it may be reasonable for OS to differ according to subgroup.
What alternative approach has the ERG suggested?	The ERG has sought to validate the company's modelled base case OS estimates via clinical expert opinion. See Section 4.2.6.1 In addition, the ERG proposed that the following sources be used to estimate OS within the submission: • SCT-2L: Balzarotti et al. (2016)[2] • SCT+3L+: Gopal et al. (2015)[1] • SCT-3L+: Balzarotti et al. (2015)[2] The ERG was aware that OS data from KEYNOTE 087[5] are available and have therefore conducted additional scenario analyses using this study (see Section 6.2.1.10).
What is the expected effect on the cost-effectiveness estimates?	Given that a conservative approach to modelling OS has been adopted the use of alternative data sources for OS as outlined by the ERG may not have a material impact on the ICER, but may improve the plausibility of estimates for life-years gained and thus QALYs gained. PFS is the key driver in this submission.
What additional evidence or analyses might help to resolve this key issue?	Mature OS data from KEYNOTE-204 along with clinical validation of OS estimates would address outstanding uncertainty surrounding OS extrapolation. In the absence of mature OS data, exploration of larger and more robust datasets (e.g. routinely collected data) that could inform OS assumptions may inform a more appropriate range of scenarios for OS.

Abbreviations: ERG, Evidence Review Group

Key Issue 9: Time on treatment (ToT) for BV in SCT-3L+ and SCT+3L+ subgroups

Report sections	4.2.8.2
Description of issue and why the ERG has identified it as important	The company assumed that patients treated with BV receive the same maximum ToT as pembrolizumab (35 cycles). However, as per the SmPC for BV, treatment should be provided for a maximum of 16 cycles. The ERG considered the company's base case assumption to be inappropriate and leads to an overestimation of BV treatment costs.
What alternative approach has the ERG suggested?	Assuming a maximum number of treatment cycles of 16 is the ERG's preferred assumption. The ERG conducted this scenario analysis. For completeness, the ERG also conducted a number of ToT scenarios including use of KM data only (no extrapolation) and the use of alternative extrapolation points (26 weeks and 52 weeks).
What is the expected effect on the cost-effectiveness estimates?	Assuming a maximum number of 16 cycles (for BV) will result in lower acquisition costs for BV and an increased ICER for pembrolizumab.
What additional evidence or analyses might help to resolve this key issue?	Data reflecting the use of BV in clinical practice, including in terms of 'real-world' utilisation of BV by subgroups relevant to UK clinical practice, would inform more realistic ICER estimates.

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ToT, time on treatment

1.6. Other key issues: summary of the ERG's views

No other key issues were identified.

1.7. Summary of ERG's preferred assumptions and resulting ICER

A summary of ERG's preferred assumptions and resulting ICER is provided for each subgroup in Table 3 (SCT-2L), Table 4 (SCT-3L+), and Table 5 (SCT+3L+).

Table 3: Summary of ERG's preferred assumptions and ICER (SCT-2L)- includes pembrolizumab PAS

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case deterministic	[redacted]	[redacted]	£53,581
ERG corrected company base case (deterministic)	[redacted]	[redacted]	£53,099

ERG corrected company base case (probabilistic)	[redacted]	[redacted]	£56,446
ERG corrected company base case used as start point for ERG analyses, below			
Scenario 14: Balzarotti et al (2016) used as the data source for estimating OS for both pembrolizumab and chemotherapy (IGEV) - Key Issue 8	[redacted]	[redacted]	£41,007
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms - Key Issue 6	[redacted]	[redacted]	£94,319
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	£89,930
Scenario 5: No difference in SCT rates between treatment arms (apply pembrolizumab allo-SCT and auto SCT rate to both arms)	[redacted]	[redacted]	£109,876
Scenario 6: Dose intensity for pembrolizumab assumed to be 100%	[redacted]	[redacted]	£112,387
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	£112,284
Scenario 11: 26-week data cut point for ToT	[redacted]	[redacted]	£202,428
ERG base case (deterministic)*	[redacted]	[redacted]	£202,428
ERG base case (probabilistic)	[redacted]	[redacted]	£176,859

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Notes: * ERG base case combines all preferred scenarios

Table 4: Summary of ERG's preferred assumptions and ICER (SCT-3L+)- includes pembrolizumab PAS

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case	[redacted]	[redacted]	Dominant (-£33,316)
Company base case used as start point for ERG analyses, below			
Scenario 14: Balzarotti et al (2016) used as the data source for estimating OS for both pembrolizumab and chemotherapy- Key Issue 8	[redacted]	[redacted]	Dominant (-£24,450)
Scenario 22: Semi parametric approach to modelling PFS (cut point for PFS set at 26 weeks)	[redacted]	[redacted]	Dominant (-£27,163)
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms - Key Issue 6	[redacted]	[redacted]	Dominant (-£61,670)
Scenario 18: Subsequent treatment assumed to reflect UK practice: 100% of patients who fail pembrolizumab go on to receive BV AND 100%	[redacted]	[redacted]	£24,265

of patients who fail on BV go on to receive bendamustine alone - Key Issue 7			
Scenario 19: Maximum ToT for brentuximab set to 16 cycles (not 35 as per base case) - Key Issue 9	[redacted]	[redacted]	£52,006
Scenario 11: Cut-off for ToT to reflect PFS data cut point (26 weeks)	[redacted]	[redacted]	£79,232
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	£67,399
Scenario 5: No difference in SCT rates between treatment arms (pembrolizumab allo-SCT and auto-SCT rate to both arms)	[redacted]	[redacted]	£62,226
Scenario 6: Dose intensity for pembrolizumab 100%	[redacted]	[redacted]	£65,018
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	£64,124
ERG base case (deterministic)*	[redacted]	[redacted]	£64,124
ERG base case (probabilistic)	[redacted]	[redacted]	£58,738

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Notes: * ERG base case combines all preferred scenarios

Table 5: Summary of ERG's preferred assumptions and ICER (SCT+3L+)- includes pembrolizumab PAS

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case	[redacted]	[redacted]	Dominant (-£73,896)
Company base case used as start point for ERG analyses, below			
Scenario 22: Semi parametric approach to modelling PFS (cut point for PFS set at 26 weeks)	[redacted]	[redacted]	Dominant (-£57,940)
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms - Key Issue 6	[redacted]	[redacted]	Dominant (-£79,339)
Scenario 19: Maximum ToT for brentuximab set to 16 cycles (not 35 as per base case) - Key Issue 9	[redacted]	[redacted]	Dominant (-£68,202)
Scenario 11: Cut-off for ToT to reflect PFS data cut point (26 weeks)	[redacted]	[redacted]	Dominant (-£49,001)
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	Dominant (-£61,514)

Scenario 5: No difference in SCT rates between treatment arms (pembrolizumab allo-SCT and auto-SCT rate to both arms)	[redacted]	[redacted]	Dominant (-£66,889)
Scenario 6: Dose intensity for pembrolizumab 100%	[redacted]	[redacted]	Dominant (-£64,127)
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	Dominant (-£63,904)
Scenario 18: Subsequent treatment assumed to reflect UK practice: 100% of patients who fail pembrolizumab go on to receive nivolumab AND 100% of patients who fail on BV go on to receive nivolumab - Key Issue 7	[redacted]	[redacted]	Dominant (-£33,849)
ERG base case (deterministic)*	[redacted]	[redacted]	Dominant (-£33,849)
ERG base case (probabilistic)	[redacted]	[redacted]	Dominant (-£34,156)

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Notes: * ERG base case combines all preferred scenarios

Modelling errors identified and corrected by the ERG are described in Section 5.3. For further details of the exploratory and sensitivity analyses done by the ERG, see Sections 5.2 and 6.2.

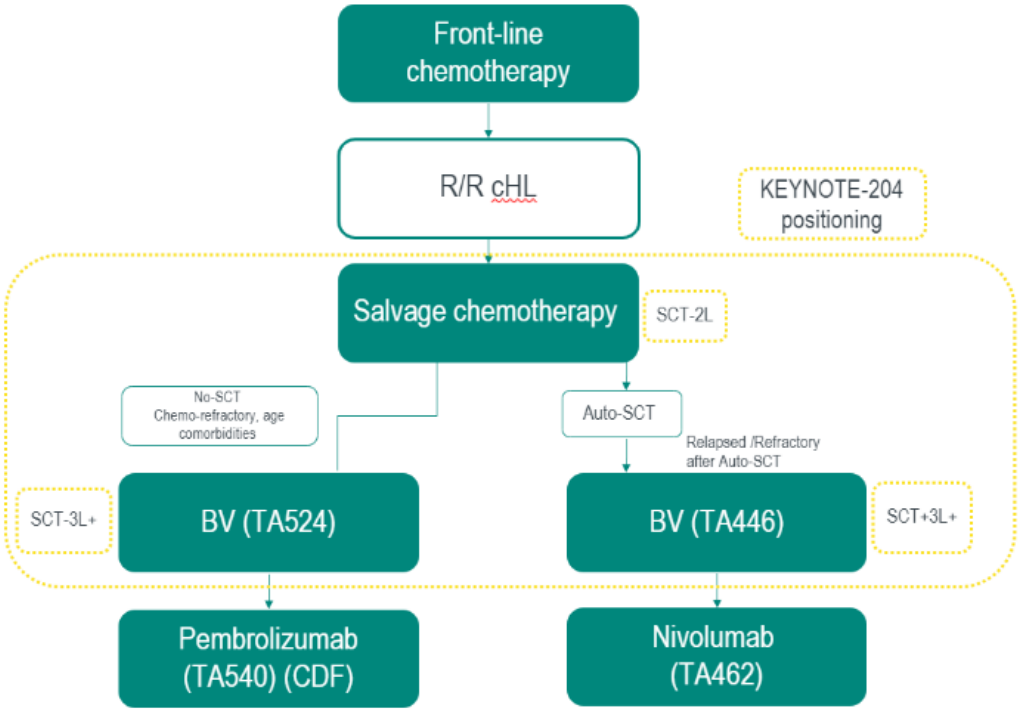
2. INTRODUCTION AND BACKGROUND

2.1. Introduction

Hodgkin's lymphoma (HL) is a form of cancer of the lymphatic system, which is an important component of the immune system. HL accounts for around 20% of all lymphomas.[6] A rare malignant proliferation of cells from the lymphoreticular system, HL mainly affects lymph node tissues, spleen, liver and bone marrow.[7] Survival with HL in England between 2013 and 2017 was 90.6% at one year and 75% at 10 years.[8] However, those considered to be relapsed or refractory (R/R) have considerably worse prognosis than the wider HL population.[9,10] The majority (59%) of HL cases occur in males and the condition is associated with a bimodal age distribution with the first peak between 20 and 24 years and the second peak between 75 and 79 years.[11] The Evidence Review Group (ERG) considered that the Company Submission (CS) offered an acceptable description of the condition; its pathophysiology, natural course and epidemiology; and the current treatment options available.

No National Institute for Health and Care Excellence (NICE) clinical guideline for the management of HL was cited in the CS, and the ERG did not identify a relevant NICE guideline. Instead, the CS depicts a treatment algorithm summary for relapsed or refractory classic HL (R/RcHL) in the UK, which is reproduced in Figure 1.

Figure 1: Treatment algorithm summary for patients with R/RcHL



Source: CS, Document B, Figure 2, p.19

The CS also outlines the relevant NICE-approved comparators for this indication (CS Document B.1.3, p.21):

- BV is recommended as an option for treating CD30‑positive HL in adults with R/R disease,[12] only if:
  - They have already had autologous stell cell transplant (ASCT) or
  - They have already had at least two previous therapies when ASCT or multi-agent chemotherapy are not suitable and,
  - The company provides BV according to the commercial agreement.
- Nivolumab is recommended, within its marketing authorization, as an option for treating R/RcHL in adults after ASCT and treatment with BV.[13]
- Pembrolizumab is recommended, within its marketing authorization, for use within the Cancer Drugs Fund as an option for treating R/RcHL in adults who have had BV and cannot have ASCT.[14]

2.2. Background

Pembrolizumab is a monoclonal antibody of the IgG4/Kappa isotope designed to exert dual ligand blockade of the programmed cell death protein 1 (PD-1) pathway by directly blocking the interaction between PD-1 and its ligands, programmed death-ligand 1 (PD-L1) and programmed death-ligand 2 (PD-L2), which appear on antigen-presenting or tumour cells. Pembrolizumab is currently used for a range of other cancer indications in current practice. The ERG considered that the company's intended positioning, as compared to current standard of care, was appropriate and generally well-described.

The company's intended positioning for pembrolizumab can be conceptualised as three specific sub-populations:

- Patients with R/RcHL who did not have at least two prior therapies when autologous stem cell transplant is not a treatment option (SCT-2L)
- Patients with R/RcHL who are at least third line with prior stem cell transplant. (SCT+3L+)
- Patients with R/RcHL who are at least third line when ASCT stem cell transplant is not a treatment option (SCT-3L+)

For the SCT+3L+ and SCT-3L+ groups, this is the position in the treatment pathway currently occupied by brentuximab vedotin (BV), while for the SCT-2L group, this is the position in the treatment pathway currently occupied by salvage chemotherapy. Clinical advice to the ERG was that these were broadly the appropriate comparators, although the company's use of exclusively IGEV (ifosfamide; gemcitabine; vinorelbine)[2] as a chemotherapy regimen in the economic modelling did not reflect the diversity of regimens used in clinical practice. Clinical advice to the ERG indicated that various combination regimens have some evidence of efficacy, although the regimens have not been compared head-to-head. This means it is difficult to determine whether it is appropriate to assume comparable efficacy between treatments. Furthermore, the selection of chemotherapy regimen is largely a matter of centre and clinician preference.

2.3. Critique of company's definition of decision problem

The ERG considered that the company's definition of the decision problem generally matched the decision problem in the NICE scope.[15]

Table 6: Summary of 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	People with relapsed or refractory classical Hodgkin lymphoma who have received: • autologous stem cell transplant or • at least one prior therapy when autologous stem cell transplant (ASCT) is not a treatment option	As per final scope	Not applicable	The ERG considered that the company decision problem was generally well matched to the NICE scope. However, the ERG noted that the company systematic literature review (SLR) specified patients should be at least 3 years of age, whereas the company economic model excluded the paediatric population. Therefore, the company submission was narrower in age range than the company decision problem. Whereas the NICE scope said patients should have received ASCT, the company submission (CS) specified patients should have failed ASCT not solely received it.
Intervention	Pembrolizumab	As per final scope	Not applicable	As per the scope for the appraisal.
Comparator(s)	Brentuximab vedotin (BV) For people who did not have at least two prior therapies when autologous stem cell transplant is not a treatment option	As per final scope	Not applicable	The ERG agreed that BV and chemotherapy are the comparators of interest in this appraisal. The ERG, however, noted that the company SLR listed BV monotherapy, nivolumab

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• Chemotherapy regimens			monotherapy, standard of care chemotherapy regimens and ASCT as interventions as opposed to comparators. It also listed placebo or best supportive care, any intervention of interest, any treatment that facilitates an indirect comparison and no intervention as comparators.
Outcomes	The outcome measures to be considered include: • overall survival (OS) • progression-free survival • response rates • proportion receiving subsequent stem cell transplant • adverse effects of treatment • health-related quality of life.	[redacted]	[redacted]	The ERG agreed that the outcome measures are comparable between the NICE final scope and company submission. However, it is important to note that OS measures used in the company model were not directly observed from an included trial, and instead modelled from BV data in the Gopal et al. (2015)[1] study.
Economic analysis	If the evidence allows the following subgroups may be considered • people who could have a subsequent stem cell transplant (autologous or allogeneic) if they respond to treatment	Post-hoc efficacy analyses for PFS and ORR are presented for 3 subpopulations; second line subjects with no prior stem cell transplant ("SCT-2L") subjects who are at least third line with no prior SCT ("SCT-3L+")	Patients who were considered ineligible for auto SCT included patients who could have a subsequent stem cell transplant if they respond to treatment and patients whom stem cell transplant is contraindicated because of comorbidities and age.	The ERG agreed that the economic subgroup analyses presented are aligned with the reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• people for whom stem cell transplant is contraindicated because of comorbidities	subjects who are at least third line with prior stem cell transplant ("SCT+3L+")		
Subgroups	If the evidence allows the following subgroups may be considered • people who could have a subsequent stem cell transplant (autologous or allogeneic) if they respond to treatment • people for whom stem cell transplant is contraindicated because of comorbidities	Post-hoc efficacy analyses for PFS and ORR are presented for 3 subpopulations; second line subjects with no prior stem cell transplant ("SCT-2L") subjects who are at least third line with no prior SCT ("SCT-3L+") subjects who are at least third line with prior stem cell transplant ("SCT+3L+")	Patients who were considered ineligible for auto SCT included patients who could have a subsequent stem cell transplant if they respond to treatment and patients whom stem cell transplant is contraindicated because of comorbidities and age.	The ERG considered that the sub-groups in the company decision problem to be appropriate and clinically relevant, although specified differently than in the NICE final scope. The ERG considered the fact that the 'third-line' subgroups included patients who were at least third-line rather than solely third-line to be a minor issue in terms of generalizability, but to be reasonable in the circumstances.
Special considerations including issues related to equity or equality	NS.	MSD does not envisage any equality issues with the use of pembrolizumab for the treatment of R/RcHL who have received: ASCT or at least one prior therapy when ASCT is not a treatment option.	NA.	NA.

Abbreviations ASCT, Autologous stem cell transplant; BV, brentuximab vedotin; CS, Company submission; ERG, Evidence Review Group; NICE, National Institute for Health and Care Excellence; NA, Not applicable; NS, Not stated; OS, Overall survival; SLR, Systematic literature review.

Source: CS, Document B, Table 1, p.13; CS, Document B, Section 1.4, p.20.

3. CLINICAL EFFECTIVENESS

The sections below discuss the evidence submitted by the company in support of the clinical effectiveness of pembrolizumab for [redacted] The ERG reviewed the details provided on:

- Methods implemented to identify, screen, data extract and assess the risk of bias in relevant evidence
- Clinical efficacy of pembrolizumab
- Safety profile of pembrolizumab
- Assessment of comparative clinical effectiveness of pembrolizumab against relevant comparators

A detailed description of an aspect of the CS is only provided where the ERG disagreed with the company's assessment or proposal, or where the ERG identified a particular area of concern that the ERG considered necessary to highlight for the Committee.

The ERG identified three key issues in the clinical effectiveness evidence:

- The immaturity of OS data in the key trial meaning no directly observed comparative OS data were available for use in the economic model.
- The matched adjusted indirect comparison (MAIC) analysis was only conducted with regard to one potential 2L salvage chemotherapy regimen (IGEV) and is therefore not generalisable to the full range of regimens used in clinical practice in the UK.
- The intention to treat (ITT) analysis is not generalizable to the UK treatment pathway, since there are three clear subgroups (SCT-2L, SCT-3L+ and SCT+3L+), not all of which have BV as a relevant comparator.

3.1. Critique of the methods of review(s)

The company undertook a systematic review to identify relevant publications on the efficacy and safety of pembrolizumab monotherapy, compared to BV monotherapy, nivolumab monotherapy, standard of care chemotherapy regimens, ASCT, BSC and placebo, for adult and paediatric patients aged three years or older with R/RcHL who have failed ASCT or following at least one prior therapy when ASCT is not a treatment option. The company considered BV and, in the

case of 3L+ ASCT-ineligible patients: R/RcHL patients who have not had an ASCT and received more than 1 prior line of therapy, standard of care chemotherapy regimens, to be the most relevant comparators.

In total, 98 publications (describing 45 unique trials) were included in the SLR. Most studies identified in the SLR were single arm and therefore offered no comparative effectiveness data for pembrolizumab. One open-label phase III RCT (KEYNOTE-204)[3,4] was identified that included the target population and formed the pivotal trial for this appraisal. There were two further single-arm studies (KEYNOTE-087[5] and Gopal et al. (2015)[1]) that the company included as clinical effectiveness sources in the economic model. The identified evidence, with a focus on the pivotal trial, is critiqued in Section 3.2.

Table 7: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D.1.1.2	The ERG was broadly satisfied with the search methods but noted the following limitation: the SIGN RCT filter applied to database searches may not have retrieved all relevant single-arm prospective studies, Despite this limitation, the ERG was satisfied that the clinical effectiveness searches identified all relevant trial evidence.
Inclusion criteria	Appendix D.1.1.2	The ERG was generally satisfied with the robustness of the inclusion criteria. There were some potential limitations. Differences in aspects of how the population, interventions and comparators were defined are outlined above in Table 6. A total of 98 publications were included, representing 45 unique trials. The ERG was satisfied that important trials are likely to have been identified.
Screening	Appendix D.1.1.3	The ERG was satisfied with the screening process. Two independent reviewers were used with a third reviewer to adjudicate disagreements.
Data extraction	Appendix D.1.1.3	The ERG was satisfied with the data extraction process. Two independent reviewers were used with a third reviewer to adjudicate disagreements. Standardised extraction forms were used.
Tool for quality assessment of included study or studies	Appendix D.1.2.3	The ERG was satisfied with the risk of bias assessment. The Newcastle-Ottawa scale was used for single-arm studies, and the NICE risk of bias tool (a modification of the

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		Cochrane tool) was used for comparative studies.
Evidence synthesis	Document B.2.8; Document B.2.9; Appendix D.1.2	No meta-analysis of pembrolizumab trials was conducted since there was only one Phase III RCT. The ERG considered this to be appropriate. The ERG's critique of the matched adjusted indirect comparison (MAIC) is found in Section 3.4.

Abbreviations: CS, Company submission; ERG, Evidence Review Group

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

Of 45 studies included in the SLR, only one study (KEYNOTE-204[3,4]) – an open-label RCT compared pembrolizumab with BV directly – and therefore forms the pivotal trial in the clinical effectiveness evidence.

3.2.1. Study design

The key trial included from the company's SLR, and the only source of directly comparative evidence to inform the economic model, is a Phase III, open label RCT (KEYNOTE-204[3,4]) evaluating pembrolizumab in patients with R/RcHL who have previously received at least one multi-agent chemotherapy regimen from countries including the UK, USA, Japan, Italy, Sweden, Australia, Poland and Russia (although details on UK sites were not provided). The clinical effectiveness data in the CS are principally from the intention-to-treat (ITT) population, although post-hoc subgroup results from the three subgroups as outlined in Section 2.2 are also provided in CS Appendix L. The ERG considered the subgroups as opposed to the ITT population to be appropriate for decision making (Section 4.2.3), since the population comprises three subgroups (SCT-2L, SCT-3L+ and SCT+3L+) which do not all share a common relevant comparator in the UK treatment pathway. The company has presented a cost effectiveness scenario analysis using clinical effectiveness inputs from one other pembrolizumab trial (KEYNOTE-087[5]), although this was single-arm in nature and was not used in the company base case. The ERG considered this to be appropriate and therefore did not present further critique of this study. The ERG critique of clinical effectiveness results therefore focuses on KEYNOTE-204.[3,4]

The population, intervention and outcomes presented in KEYNOTE-204[3,4] were broadly consistent with the NICE decision problem, although it is important to note that mature OS data

were not available from KEYNOTE-204[3,4] and were therefore mapped from a single-arm BV study (Section 4.2.6.1).[1]

No specific dose of pembrolizumab was stated in the NICE decision problem for this appraisal. At the clarification stage, the company clarified to the ERG that the doses included in the CS for adult patients – 200 mg administered every three weeks and 400 mg administered every six weeks [redacted]. However, the key trial (KEYNOTE-204[3,4]) utilised only the 200 mg every three weeks dose (CS, Document B, Table 3, p.23), and this was therefore the dose used in the company base case economic model. The 400 mg every six weeks dose was considered separately in a scenario analysis.

Clinical advice to the ERG indicated that the doses of pembrolizumab and BV were appropriate with regard to UK clinical practice. However, for the SCT-2L sub-group, the company's economic model did not consider a full range of salvage chemotherapy regimens, and instead focused on IGEV, which the clinical advisor to the ERG considered to be only one of multiple potential chemotherapy regimens in clinical practice. There is likely to be some regional and/or centre-level variation in terms of chemotherapy regimen use. Clinical advice to the ERG indicated a preference for bendamustine-based regimens, whereas the clarification response from the company indicated that clinical advice received by the company did not support including such regimens on the standard of care list.

3.2.2. Randomisation stages and protocol amendments

The KEYNOTE-204[3,4] trial involved the randomisation of patients (1:1) to either pembrolizumab monotherapy (200 mg every three weeks) or BV. The ERG considered that randomisation was carried out appropriately. It was stratified by prior auto-SCT status and HL status.

KEYNOTE-204[3,4] was subject to seven protocol amendments (CS, Document B, Table 55, p.118). However, the ERG did not identify any protocol amendments that it considered likely to have introduced a high risk of bias in addition to the potential bias inherent in an open-label trial design.

3.2.3. Quality assessment of the trials of the technology of interest

The company reported a generally favourable assessment of study quality for KEYNOTE-204[3,4] as well as for the single-arm pembrolizumab studies KEYNOTE-013,[16] KEYNOTE-087[5] and KEYNOTE-051,[17] of which KEYNOTE-087[5] was used to inform a scenario analysis in the economic model. These three single-arm studies did not inform the MAIC. The complete quality assessment is available in Appendix D of the CS (Tables 29 and 30). The company

acknowledged appropriately the limitations of the open-label nature of KEYNOTE-204.[3,4] The company evaluated RCTs using the NICE Risk of Bias Tool, which is a modified version of the Cochrane tool, and evaluated single-arm studies using the Newcastle-Ottawa Scale,[18] which the ERG considered to be appropriate for this purpose. The ERG considered risk of bias using the published literature as well as the data presented in the CS and accompanying documents specifically for the outcomes from KEYNOTE-204[3,4] that informed the economic model (primarily PFS, response rates, proportion receiving subsequent transplant, adverse events and health-related quality of life).

While the ERG noted some strengths of trial quality such as appropriate randomisation and broadly similar baseline characteristics across arms, the ERG notes the limitations associated with the open label nature of KEYNOTE-204,[3,4] whereby neither investigators nor patients were blinded to the treatment allocation. However, the different mode of administration for pembrolizumab as an immunotherapy versus BV as a chemotherapy would make blinding difficult to achieve. While ITT analysis is typically a strength of trials in terms of internal validity, in the context of this appraisal it has substantial limitations in terms of external validity given the existence of three clear sub-groups (SCT-2L, SCT-3L+ and SCT+3L+), not all of which have BV as a relevant comparator. Additionally, the ERG identified a risk of attrition bias in the KEYNOTE-204[3,4] trial given [redacted]. A further limitation to the external validity of the KEYNOTE-204[3,4] trial in the context of this appraisal is the immaturity of OS data, precluding the use of directly observed comparative OS data as a clinical effectiveness input to the economic model.

3.2.4. Baseline characteristics

Baseline characteristics for patients included in the KEYNOTE-204[3,4] study were reported in the CS (Document B, Table 7, pp.33-36) for the ITT population. Baseline characteristics were not provided in the CS for the subgroup populations (SCT-2L, SCT-3L+ and SCT+3L+) that the ERG considered to be most relevant for decision-making. Considering the ITT population, the ERG agreed with the company's assertion that the baseline characteristics in KEYNOTE-204[3,4] were generally well-balanced between the pembrolizumab and BV arms. While the ERG noted a tendency for ECOG score of 1 and high-risk features such as bulky disease, baseline B symptoms and baseline bone marrow involvement to be more prevalent in the pembrolizumab arm than the BV arm, the ERG considered there to be no major baseline imbalances between the two arms of the KEYNOTE-204[3,4] trial. The ERG noted that in the company base case economic model the patient characteristics from European sites only were used for some

variables rather than the international population in an attempt to better reflect the UK population, while for other variables the full international ITT population was used. The ERG however considered that the international population may be more suitable, given the population of Europe as a whole is less ethnically diverse than the UK population. Baseline characteristics for selected variables for the European population in KEYNOTE-204[3,4] were presented in the CS, Document B, Table 106, p.177. Ethnicity was not reported in the European population, however age and gender appeared comparable with the international ITT population.

3.2.5. Clinical effectiveness results

Data in the target population were presented for PFS, response rates, proportion of patients receiving subsequent stem cell transplant, health-related quality of life and adverse events. It is important to note that no OS data were available from the KEYNOTE-204[3,4] trial. Statistical analyses were broadly appropriate. The primary analysis population in the CS was the ITT population for all efficacy outcomes and the All Subjects as Treated (ASaT) population for safety outcomes. The ERG has explained above how the ITT population has generalisability problems in the context of the UK treatment pathway, and that sub-group analyses are preferable for decision-making. Therefore, the clinical effectiveness efficacy results that the ERG considered to be most relevant are those presented in Section 3.2.5.5.

3.2.5.1. Overall survival

Mature OS data were not available from the KEYNOTE-204[3,4] trial (Section 4.2.6.1). Therefore, clinical effectiveness inputs for OS parameters in the company economic model were not based on directly observed comparative data.

3.2.5.2. Progression-free survival

PFS was assessed per IWG 2007[19] by blinded independent central review. Statistical analysis was conducted using the stratified Log-rank test for testing and a stratified Cox model with Efron's tie handling method for estimation. The main analysis used the primary censoring rule (CS, Document B, Table 16, p.66) for handling missing data. PFS curves were estimated using the non-parametric Kaplan-Meier (KM) method.

In the ITT population, based on a median (range) follow-up time of [redacted] months, median PFS of 13.2 (95% CI 10.9, 19.4) months in the pembrolizumab arm compared favourably with median PFS of 8.3 (95% CI 5.7, 8.8) months in the BV arm, with a hazard ratio (HR) of 0.65 (95% CI 0.44, 0.88), one-sided Log-rank test p=0.00271.

3.2.5.3. Response rate

Objective response rate (ORR) was assessed per IWG 2007[19] by blinded independent central review. Statistical analysis was conducted using the stratified Miettinen and Nurminen method.[20] Participants with missing data were classed as non-responders.

In the ITT population, there was a numerical difference in ORR in favour of pembrolizumab (ORR 65.6%, 95% CI [redacted]) over BV (ORR 54.2%, 95% [redacted]), although the difference was [redacted].

3.2.5.4. Health-related quality of life

Health-related quality of life (HRQoL) was assessed in KEYNOTE-204[3,4] using two measures - EORTC-QLQ-C30 questionnaire (version 3.0)[21] which was used to assess cancer-related quality of life, as well as the generic health status measure, EQ-5D-3L.[22] Questionnaires were completed at several time points within KEYNOTE-204: pre-dose at Cycle 1 (baseline), Cycle 3 (Week 6), Cycle 5 (Week 12), Cycle 7 (week 18), and Cycle 9 (Week 24) and then every 12 weeks until PD or up to one year while the subject is receiving study treatment. Questionnaires were also collected at discontinuation and at the 30-day safety follow-up visit.

EQ-5D-3L is the standard HRQoL measure for NICE appraisals, and following the NICE reference case, HRQoL data were reported directly from patients using the EQ-5D-3L questionnaire and the utility of the changes in QoL in the company base case economic model was based on public preferences using a choice-based method.

There was a statistically significant benefit for pembrolizumab over BV in terms of EQ-5D-3L utility scores of [redacted] points, 95% CI [redacted], [redacted], at 24 weeks. There was a statistically significant benefit for pembrolizumab over BV in terms of EQ-5D-3L visual analogue (VAS) scores of [redacted] points, 95% CI [redacted], [redacted], at 24 weeks.

3.2.5.5. Subgroup analyses

The CS reports both pre-specified and post-hoc subgroup analyses for the pivotal KEYNOTE-204 trial.[3,4]

Pre-specified subgroup analyses were conducted to assess efficacy within each category of the following classification variables:

- Prior ASCT

- Disease status following first-line therapy (refractory vs relapsed within 12 months vs relapsed after 12 months)
- Sex
- Age (binary split at 65)
- ECOG status (0 vs 1)
- Geographic region
- Prior BV status (Yes vs No)

Post-hoc subgroup analyses were conducted, dividing the population into three cohorts:

- SCT-2L
- SCT-3L+
- SCT+3L+

The results of the pre-specified and post-hoc subgroup analyses can be found in CS (Appendix L). The ERG considered the factors selected by the company for consideration in subgroup analysis to be appropriate. However, the ERG considered that the three cohorts considered in the post-hoc subgroup analysis should have been pre-specified analyses, given their relevance to clinical treatment pathways and decision-making.

In the post-hoc subgroup analysis, results in the primary analysis favoured pembrolizumab over BV for both PFS and ORR in all three cohorts. However, p-values or confidence intervals for the between-arm difference were not reported. This made it difficult for the ERG to comment on the robustness of the efficacy of pembrolizumab in each of these three cohorts. PFS in the pembrolizumab arm was highest in the SCT-2L subgroup ([redacted], 95% CI [redacted] vs [redacted], 95% CI [redacted] for BV). Notwithstanding the lack of information regarding statistical significance, the mean difference between arms was also highest in the SCT-2L subgroup, and lowest in the SCT+3L+ subgroup (pembrolizumab [redacted], 95% CI [redacted]; BV [redacted], 95% CI [redacted]).

3.2.5.6. Adverse effects

Adverse events (AEs) in the KEYNOTE-204 trial[3,4] were reported in the CS B.2.10. AEs were considered in the ASaT population, which formed the primary safety analysis population. Overall, the ERG agreed with the company that pembrolizumab had an acceptable safety

profile. AEs were very common with nearly all participants experiencing at least one AE and the majority in each treatment arm experiencing treatment-related AEs. The ERG agreed with the company that the incidence of AEs both overall and in specific AE categories was comparable between the treatment arms. The ERG agreed with the company that the biggest difference was noted with regard to serious adverse events (SAEs, pembrolizumab [redacted] vs BV [redacted]), and accepts the company's explanation of this in terms of differing duration of exposure (pembrolizumab median [redacted] days vs BV median [redacted] days).

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

As stated in Section 2.2 the appropriate comparator for the SCT-2L subgroup is 'standard of care' (SoC), which is salvage chemotherapy but not BV. The pivotal trial (KEYNOTE-204[3,4]) did not contain a head-to-head comparison of pembrolizumab vs SoC, and therefore the company carried out an indirect comparison, with adjustment for known prognostic or effect-modifying covariates by MAIC.

The company identified a retrospective study of UK clinical practice (Eyre et al., 2017[23]) and received clinical advice (CS, Appendix D1.2.1, pp35-36), and thereby identified the following SoC regimens in the UK with associated trial evidence (CS, Document B, Table 58): GDP (gemcitabine, dexamethasone, cisplatin), IGEV, ICE (ifosfamide; carboplatin; etoposide), ICE + panobinostat, DHAP (dexamethasone; high dose Ara C [cytarabine]; cisplatin), ESHAP (etoposide; solu-medrone [also called methylprednisolone]; high dose Ara C [cytarabine]; cisplatin).

Investigational regimens and combinations with other agents were excluded on the basis that these were 'not considered representative of SoC in the UK' (CS, Document B, p36).

Clinical advice received by the ERG suggested that there were local preferences for these SoC regimens in different UK centres and each had some track record of efficacy, but they had not been compared head-to-head. They also commented that the company's selection seemed comprehensive, other than the omission of bendamustine or bendamustine-containing regimens. This omission was raised in clarification and the company explained that these were not included by their clinical advisors for clinical practice, nor were they suggested by guidelines or a retrospective study (see clarification response A5). Furthermore, the company stated that a MAIC analysis of bendamustine regimens would not have been feasible with the information available (clarification response A16).

The company carried out a targeted literature review (TLR), to identify potential prognostic or effect-modifying variables, viz.: 1.) disease status (early relapse vs late relapse vs refractory), 2.) age, 3.) ECOG 0 vs 1, 4.) presence of bulky disease, 5.) prior radiotherapy, 6.) sex, and 7.) presence of B symptoms. The company also stated that "The following patient characteristics were considered as potential prognostic factors but were either considered to have significant overlap with the aforementioned covariates, or were deemed to be less relevant from a clinical perspective: refractory relapse vs sensitive relapse, serum albumin levels, haemoglobin levels, white cell count, and lymphocyte count." (CS, Appendix D1.2.3 p52). The clinical advisor to the ERG assessed the company's list of variables and was largely satisfied, though suggested the addition of erythrocyte sedimentation rate (ESR).

With the six selected trials believed to represent UK practice the company inspected the available outcomes (Precision report table 13)[24] and available covariates (CS, Appendix D, Table 23). On this basis, and in particular because only Balzarotti et al. (2016)[2] provided KM data for PFS assessment, the company did not present results from the other five trials. The company further elaborated (clarification question A14) that the study populations were not comparable in terms of ASCT ineligibility, though Balzarotti et al. (2016)[2] was retained (further details below). The ERG did not receive any evidence that IGEV is not a suitable proxy for SoC in the UK, but generalisability of the results is not assured.

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

The company assembled six MAIC analyses for a pembrolizumab vs SoC comparison. However, as detailed in Section 3.3, the company presented only one of these (pembrolizumab vs IGEV using Balzarotti et al. (2016)[2]) in the CS, now described.

The company base case analysis for this subgroup was "pembrolizumab vs. IGEV in second line subjects without prior stem cell transplant (SCT) based on the KEYNOTE-204 and Balzarotti 2016 studies." (CS, Document B, pp125-6). The base case analysis was restricted to patients aged <65 years to conform with the IGEV population, but a sensitivity analysis was also described without this restriction.

The company matched pembrolizumab (from KEYNOTE-204, individual patient data [IPD] available) to IGEV (Balzarotti et al. (2016)[2]) aggregate data available for some covariates, and pseudo-IPD data for PFS) with an 'unanchored' MAIC since there was no common comparator between these studies. The numbers of participants available in each trial was low ([redacted], KEYNOTE-204 pembrolizumab arm; [redacted], Balzarotti et al. trial IGEV arm) and under matching

the effective sample size (ESS) was lower again ([redacted]). The company provided a histogram of MAIC weights in response to clarification question A15. There was only one observation with zero or very small weight, indicating very substantial overlap between the two samples (but with a known *lack of overlap in age already accounted for* by exclusion of >65 year olds when forming the base case).

The company presented MAIC-adjusted results for CR, PR, OR (CR or PR) and PFS. The last of these is relevant from an economic perspective. OS was not analysed because it is not yet available from the KN204 trial. After matching, the estimated base case PFS [redacted] (CS, Document B, Table 61). Under the sensitivity analysis (which did not restrict the age of participants in KEYNOTE-204) the estimate for PFS was [redacted], Doc B Table 66). The result is not significant under either analysis (base case or sensitivity) though the directions of the point estimates differ. The estimates were made with Cox regression but the ERG questions whether proportional hazards would be supported in the company's base case (Doc B fig 27). For clinical outcomes, in the base case the company reported significantly improved PR (RR= [redacted]) (CS, Document B, Table 64) but the result for CR (RR= [redacted]) (CS, Document B, Table 63) was not significant.

The purpose of the MAIC in this instance is to reweight participants in the pembrolizumab trial (KN-204) so that its aggregate covariate values match those of the IGEV trial. Characteristics before and after matching are shown in Tables 60 and 65 (CS, Document B), showing that the MAIC correctly adjusted for these covariates. Nevertheless, the ERG notes that the interpretation of the resulting estimate is of the effect of pembrolizumab vs IGEV *in the population of the IGEV trial*. The IGEV trial was carried out in specialist centres in Italy, and it is important to consider whether this is a suitable representation of the 'target population', SCT-2L in UK clinical practice. For example, the Balzarotti et al. (2016)[2] sample contained no patients over 65 (even though "age was not specifically an exclusion criterion in the comparator study", Doc B p126) compared with [redacted] of the pembrolizumab (KN-204) sample. The ERG suggests this may indicate a less age-diverse study population in the IGEV trial than in KEYNOTE-204 or UK clinical practice.

Because the base case MAIC is unanchored, an assumption must be made that all effect modifiers and prognostics have been accounted for. The company acknowledged this was a strong assumption and in the ERG's view correctly warned of a potential for bias. ECOG score was a known prognostic variable that, because it was not reported in the comparator study, could not be adjusted for in the MAIC. Another important prognostic not incorporated to the

company's MAIC was SCT eligibility. In the KEYNOTE-204 subgroup no participants had received prior ASCT, and these were treated as ASCT-ineligible, whereas "none of the comparator studies explicitly limited enrolment to ASCT-ineligible patients". The company indicated that information was limited on this and relevant patient characteristics "for example comorbidities was not well-described in publications beyond a requirement for 'adequate organ function'" (CS, Document B, p137). The clinical advisor explained to the ERG that ASCT eligibility can be a dynamic characteristic in some patients. The company outlined (in the CS and clarification A14) that there were differences in the proportions of patients from subgroup SCT-2L who subsequently went on to receive SCT: much lower in KN-204 ([redacted] and [redacted] in each arm) compared to Balzarotti et al. (2016)[2] (at least 81%).

Conclusion: Only one SoC regimen was available for MAIC analysis with respect to PFS (IGEV). The ERG noted that a number of other salvage treatments are used in clinical practice but these could not in the company's view be analysed by MAIC. On the other hand, the ERG did not receive any evidence that IGEV was an unsuitable proxy for SoC. The ERG agreed with the company that the results of this unanchored MAIC (Pembrolizumab vs IGEV) should be treated with caution. The MAIC accounted for a number of important prognostic/effect-modifying variables, but may contain residual bias from others unadjusted for, and in particular was known not to adjust for SCT eligibility or ECOG. Furthermore, the ESS was low, leading to estimates with poor precision. Finally, the estimate of effect is with reference to the population in the IGEV trial rather than UK clinical practice.

3.5. Additional work on clinical effectiveness undertaken by the ERG

None.

3.6. Conclusions of the clinical effectiveness section

The ERG considered that the company had identified all relevant clinical evidence for this appraisal. Data were not available for the OS outcome included in the NICE final scope for this appraisal.[15] Requisite information regarding the methodology and outcomes for clinical effectiveness was available in the CS, and was generally reasonably described.

There was one pivotal RCT comparing pembrolizumab and BV (KEYNOTE-204[3,4]) that could provide directly comparative evidence for the base case economic model. A further single arm pembrolizumab trial (KEYNOTE-087[5]) informed a company scenario analysis. While there were several strengths to the KEYNOTE-204 trial,[3,4] the open-label nature of the trial was a key limitation, although the extent to which blinding could be achieved was limited by the different

modes of administrative of the immunotherapy pembrolizumab and the chemotherapy BV. The ERG was satisfied that there was evidence of a benefit for pembrolizumab over BV in terms of PFS and ORR. In the absence of directly comparative evidence for pembrolizumab versus salvage chemotherapy (the relevant comparator for the SCT-2L subgroup), MAIC analysis was conducted. The base case MAIC that informed the economic model included two trials. Limitations of the MAIC included the fact that it was unanchored.

The three key issues in the clinical effectiveness evidence are as follows:

- The immaturity of OS data in the key trial meaning no directly observed comparative OS data were available for use in the economic model
- The matched adjusted indirect comparison (MAIC) analysis was only conducted with regard to one potential 2L salvage chemotherapy regimen (IGEV) and is therefore not generalizable to the full range of regimens used in clinical practice in the UK
- The intention to treat (ITT) analysis is not generalizable to the UK treatment pathway, since there are three clear subgroups (SCT-2L, SCT-3L+ and SCT+3L+), not all of which have BV as a relevant comparator.

4. COST-EFFECTIVENESS

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

The company conducted a single systematic literature review with the overall objective being to identify and summarize a) the published cost-effectiveness analysis, b) health-related quality of life associated with the treatment healthcare costs, and c) and resource requirements of patients with R/RcHL.

Table 8: Summary of ERG's critique of the methods implemented by the company to identify health economic evidence: Cost-effectiveness studies

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix G.1 and Appendix G.5	The ERG was broadly satisfied with the search methods.
Inclusion criteria	Appendix G.2, Appendix G, Table 32	Appropriate. Studies including adults and children with R/RcHL were eligible for inclusion. No restriction was placed in respect of pharmacological interventions other than line of therapy, second- or later line therapies (although the latter distinction was not noted in the PICOS table but in the supporting narrative). Study designs specified were relevant for the objective of the review (economic evaluations). Only full texts available in English language were included. Included studies were grouped: UK and non-UK setting. A total of 16 studies met the eligibility criteria for the review: of these, 2 were UK-specific. In addition, 7 UK-specific HTA submissions (4 NICE and 3 SMC). The company noted that two were conducted in a UK setting and no studies compared pembrolizumab versus brentuximab or chemotherapy in the population of interest in the UK setting.
Screening	Appendix G.3	Appropriate. Studies were dual screened independently at title/abstract and full-text screening stages.
Data extraction	Appendix G.4	Appropriate. Data extraction was completed by two reviewers independently and checked by a third reviewer.
QA of included studies	Not reported	Quality appraisal of identified studies reporting economic evaluations was not reported. Given the absence from the CS, the ERG assumed that QA of included studies was not undertaken by the company.

Abbreviations: cHL, classical Hodgkin lymphoma; CS, Company Submission; ERG, Evidence Review Group; QA, quality assessment; R/R, relapsed, refractory

Table 9: Summary of ERG's critique of the methods implemented by the company to identify health economic evidence: Health-related quality of life

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix H	The ERG was broadly satisfied with the search methods.
Inclusion criteria	Doc B, Section B.3.4.3, Appendix H (cross references detail in Appendix G.2, Appendix G, and Table 32)	Broadly appropriate. Studies including adults and children with R/RcHL that reported HRQoL using disease-specific and generic instruments or directly reported health state utility values were eligible for inclusion. No restriction was placed in respect of pharmacological interventions other than line of therapy, second- or later line therapies (although the latter distinction was not noted in the PICOS table but in the supporting narrative). Only full texts available in English language were included. Included studies were grouped: UK and non-UK setting. A total of 21[a] studies (in 37 publications) were identified in the review. Of these, the company reported in detail on 5 of the studies as directly relevant to the submission. Of the 5 studies, 4 were relevant to the UK setting and 1 was conducted from a US perspective but had evaluated pembrolizumab. In addition, 7 previous HTAs were identified (4 NICE and 3 SMC). The company discussed the included studies and commented on the utility estimates identified in context of the KEYNOTE-204 data.
Screening	Appendix H cross references detail in Appendix G.3	Appropriate. Studies were dual screened independently at title/abstract and full-text screening stages.
Data extraction	Appendix H cross references detail in Appendix G.4	Appropriate. Data extraction was completed by two reviewers independently and checked by a third reviewer.
QA of included studies	Not reported	Quality appraisal of identified studies reporting HRQoL/utility data was not reported. Given the absence from the CS, the ERG assumed that QA of included studies was not undertaken by the company.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; HTA,s, health technology assessment; NICE, National Institute of Health and Care Excellence; QA, quality assessment; SMC, Scottish Medicines Consortium

Notes:

a 18 studies (in 33 publications) were identified in the review and an additional 3 studies (in 4 publications) were identified as relevant from the cost-effectiveness review.

Table 10: Summary of ERG's critique of the methods implemented by the company to identify health economic evidence: Healthcare resource use and costs

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix I	The ERG was broadly satisfied with the search methods.
Inclusion criteria	Appendix I (cross references detail in Appendix G.2, Appendix G, and Table 32)	Broadly appropriate. Studies including adults and children with R/RcHL that reported healthcare costs and/or resource use were eligible for inclusion in the review. No restriction was placed in respect of pharmacological interventions other than line of therapy, second- or later line therapies (although the latter distinction was not noted in the PICOS table but in the supporting narrative). Only full texts available in English language were included. Included studies were grouped: UK and non-UK setting. A total of 25[a] studies were included. Of these, the company considered that two of the studies were UK specific. The company did, however, also tabulate findings from the included non-UK specific studies. In addition, 7 previous HTAs were identified (4 NICE and 3 SMC). Identified evidence relevant to the UK setting was used to inform model parameters with the exception of Parker (2017)[25] (Scotland).
Screening	Appendix I cross references detail in Appendix G.3	Appropriate. Studies were dual screened independently at title/abstract and full-text screening stages.
Data extraction	Appendix I cross references detail in Appendix G.4	Appropriate. Data extraction was completed by two reviewers independently and checked by a third reviewer.
QA of included studies	Not reported	Quality appraisal of identified studies reporting healthcare resource use and cost data was not reported. Given the absence from the CS, the ERG assumed that QA of included studies was not undertaken by the company.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HTAs, health technology assessment; NICE, National Institute of Health and Care Excellence; QA, quality assessment; SMC, Scottish Medicines Consortium

Notes:

a 21 studies were identified in the literature search and four studies identified as eligible for inclusion from the review of cost-effectiveness analyses

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

4.2.1. NICE reference case checklist

Table 11: NICE reference case checklist

Attribute	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Only direct health effects were captured in the model. Carer disutility and wider societal benefits were not considered. The company's appraoch seems reasonable.
Perspective on costs	NHS and PSS	An NHS perspective was adopted as appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The company submitted a cost utility analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The time horizon used in the base case was 40 years. At this time point [redacted] of patients were still alive in the model (in both treatment arms). The ERG considered using a longer time horizon within their preferred base case.
Synthesis of evidence on health effects	Based on systematic review	For the base case analysis (ITT population) and subgroup analyses (SCT-3L+ and SCT+3L+), treatment efficacy with respect to PFS was derived directly from KEYNOTE-204.[3,4] For the SCT-2L subgroup, treatment efficacy was based on a MAIC. OS for all subgroups were based on a published study by Gopal et al.[1]
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs were used as appropriate. The EQ-5D-3L was used, which is considered the preferred health related quality of life measure in adults.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Values were elicited directly from patients in KEYNOTE-204.[3,4]
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Dolan et al. (1997)[26] was used, which is considered a valid source.

Attribute	Reference case	ERG comment on company's submission
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	There were no equity concerns.
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	PSSRU (2018/19) and NHS reference costs were used as appropriate.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and benefits were discounted at 3.5%, as appropriate.

Key: EQ-5D, EuroQol 5 dimension; ERG, Evidence Review Group; HRQoL: health-related quality of life; ITT, intention to treat; NHS, National Health Service; PSS, Pseronal Social Services; QALY: quality-adjusted life year; TA: technology appraisal

4.2.2. Model structure

The company submitted a partitioned survival model, also known as an area under the curve (AUC) model which consisted of three mutually exclusive health states, Progression free survival (PFS), progressed disease (PD) and death. Patients entered the model in the PFS health state and the proportion of patients remaining progression free over time was determined by the slope of the PFS curve. Membership in the PD health state was estimated based on the difference between the OS and PFS curves. The ERG acknowledged that AUC models are frequently used within the area of oncology. Clinical opinion to the ERG has confirmed that PFS and OS are considered the key outcomes for patients with RR cH/L.

The ERG noted that the company's modelling approach differed to previous models submitted to NICE for Hodgkin's lymphoma with respect to SCT. Within the current submission the company confirmed that pembrolizumab would not be used as a bridge to transplant, where the aim is to control the disease, and possibly elicit a disease response to allow for SCT. The company also stated that within this submission SCT was not modelled as an explicit health state, but rather as a model input due to the study design of KEYNOTE 204[3,4] and paucity of data. However in previous NICE TAs, including TA540[14] and TA462,[13] treatments were modelled as bridge to transplant and included survival, cost and Qol implications associated with SCT. The ERG noted that the current model for pembrolizumab only includes costs associated with SCT, which represents a departure from prior modelling approaches. Furthermore, based on TA524,[12] the ERG understood that pembrolizumab has the potential to be used by clinicians as a bridge to transplant in 'fitter' patients.

The cycle length used in the model was one week, which appeared to sufficiently capture progression and clinically important events. Given that pembrolizumab and brentuximab are administered on a three-weekly basis, a longer modelled cycle length (reflecting frequency of administration) could have been considered by the company. However, weekly cycles were considered appropriate.

4.2.3. Population

The company presented base case results for the ITT population in KEYNOTE-204[3,4] which included second-line patients (SCT-2L) and patients who were third-line or higher (SCT-3L+ and SCT+3L+). Several patient characteristics used within the model including weight and body surface area (BSA) were based on European patient characteristics, whilst age and sex reflected the entire ITT population. The company did not provide rationale as to why separate characteristics were used for certain model parameters. However, the ERG noted that the company's model had a function which allowed for characteristics to be changed to reflect the ITT characteristics only. During the clarification process, the company provided updated results using ITT patient characteristics only, however this did not impact on the ICER.

The ERG note that cost-effectiveness results were not provided for a paediatric population, [redacted]. As such it is unclear whether the results reported in Section 5.1 are generalisable to a paediatric population.

4.2.4. Interventions and comparators

For the SCT-2L subgroup, the company assumed the comparator most likely to be displaced is salvage chemotherapy (specifically IGEV). The company assumed that the clinical efficacy associated with IGEV (from the MAIC) is representative of other chemotherapy regimens. The ERG noted that this assumption is uncertain and has not been supported by clinical evidence. Furthermore, based on clinical opinion to the ERG, other potentially relevant treatment regimens appear to have been omitted, including bendamustine. The company was asked to comment on why the regimen was omitted and noted that clinical opinion and published literature searches did not identify bendamustine as a plausible treatment. The ERG did not consider the company's rationale to be accurate or reasonable, given that clinical response to the ERG has outlined a strong preference for using bendamustine.

With respect to the SCT-3L+ and SCT+3L+ subgroups the company assumed BV to be the comparator most likely displaced. The ERG considered this to be appropriate based on current treatment algorithm depicted in Section 2.1.

Within the ITT base case economic analysis BV was selected as the primary comparator. However, based on the clinical treatment pathway for R/RcHL patients (Section 2.1), comparators differed according to whether patients are being treated second-line or third-line. BV does not appear to be the most appropriate comparator for the SCT-2L population, given that salvage chemotherapy is the appropriate comparator for this subgroup.[15] The ERG did not consider there to be a single comparator applicable to all patient subpopulations, therefore each subgroup is assessed by the ERG separately within this appraisal. The ERG considered that the ITT analysis and results should be interpreted with caution and that the subgroup analyses results should be considered most relevant for decision making.

4.2.5. Perspective, time horizon and discounting

A 40-year time horizon was used in the economic model. The company justified the time horizon on the basis that most of the modelled patients are estimated to have died by this time point (with [redacted] of patients alive at 40 years). Based on a review of NICE TA540[14] (pembrolizumab for relapsing or refractory classical Hodgkin's Lymphoma), the ERG preferred a 50-year time horizon. However, a shorter time horizon (40 years) has been used and accepted previously in TA524[12] (BV for CD30- positive Hodgkin's Lymphoma). In order to ensure that all patients have died in the model, an additional scenario analysis was conducted in which the time horizon was increased to 50 years (Section 6.2.1.8). However, given the small proportion of patients alive at this time point, this did not have a major impact on results (see Section 6.2.1.8)

No issues were identified with respect to perspective or discouting. An NHS perspective was adopted which is considered appropriate. Costs and benefits were discounted at 3.5%, as per NICE guidance.

4.2.6. Treatment effectiveness and extrapolation

As previously mentioned, the key driver of pembrolizumab incremental QALYs within the model was PFS and associated assumptions surrounding this parameter. Given that the company assumed no difference in OS between pembrolizumab and IGEV in the SCT-2L subgroup or versus BV in the SCT-3L+ and SCT+3L+ subgroups, pembrolizumab was not associated with an incremental LY gain. Although this approach may be considered somewhat conservative, there were limitations surrounding the company's handling of OS within the model, which are discussed below.

4.2.6.1. Overall survival

OS data from KEYNOTE-204[3,4] were not mature. Therefore, for the base case ITT analysis the company estimated OS for both pembrolizumab and BV based on BV Kaplan Meier data from a published study (Gopal et al. (2015)[1]). The ERG acknowledged that eligible patients within this study were those who were 12 years or older with relapsed or refactory HL after prior auto-SCT. Median OS within the study was [redacted] months.

ITT population and subgroup analyses

OS data for BV were assumed to be representative of OS for pembrolizumab patients; i.e. OS was the same in both treatment arms. In order to model long-term survival estimates, the company extrapolated OS using a fully parametric modelling approach, whereby a log normal curve was fitted to the Gopal et al. (2015)[1] KM data (see Figure 2). The company justified the use of the log normal curve on the basis that it produced the lowest AIC/BIC statistics and produced plausible long term survival estimates. Based on this approach, five-year OS was estimated to be [redacted] for both pembrolizumab and BV.

Figure 2. Modelled OS (ITT population and subgroups)

Abbreviations: BV, brentuzimab vedotin; OS, overall survival

The ERG acknowledged that assuming no difference in survival between treatment arms may be considered a conservative assumption and could potentially underestimate the impact of pembrolizumab on OS. However, as noted above, the company assumed that OS from Gopal et al. (2015)[1] would be generalizable to all subgroups. Given that patients in Gopal et al. (2015)[1] were those had a prior SCT (reflecting the SCT+3L+ subgroup) the ERG consider that there was some concern surrounding the generalisability of OS estimates to the subgroups. Furthermore, based on clinical opinion to the ERG (and clinical opinion provided to the company), it may be reasonable for OS to differ according to subgroup. Therefore, in order to

estimate more plausible OS estimates for each subgroup, the ERG conducted additional scenario analyses whereby Balzarotti et al. (2016)[2] was used to estimate OS for the SCT-2L and SCT-3L+ subgroups. As patients in Balzarotti et al. (2016)[2] were considered to be more representative of these subgroups given that they had not received prior SCT, this source has been selected for use within the ERG preferred base case for this subgroup. See Section 6.2.1.14 for results.

In order to explore the impact of using an alternative OS data source on the ICER, the company also carried out 'Alternative approach 1' (CS Document B section 3.3.3.1, company scenario 5) in which the same assumption of equal OS between arms was made, but OS data were instead derived from KEYNOTE-087[5], a non randomised, phase II, single arm study which assessed the effectiveness of pembrolizumab in patients with R/R cHL. It should be noted that KEYNOTE-087[5] was the only other alternative data source used in the model to estimate OS. The study included 3 cohorts of R/R cHL patients. Cohort 1 were patients who failed to achieve a response or progressed after autologous stem cell transplant (auto-SCT) and relapsed after treatment with, or failed to respond to treatment with BV. Cohort 2 were patients who were unable to achieve a complete response or partial response to salvage chemotherapy and did not receive auto-SCT, but relapsed after treatment with, or failed to respond to treatment with BV. Cohort 3 were patients who failed to achieve a response to, or progressed after, auto-SCT, and had not received BV after auto-SCT and did or did not, receive BV as part of primary treatment or salvage treatment. The ERG noted that the scenario analysis results were based on the KEYNOTE-087[5] ITT patient population (results were not provided using OS rates from each individual cohort).

The ERG understood that this scenario analysis (which was provided for the ITT population only) caused the incremental QALY gain for pembrolizumab to increase, as pembrolizumab resulted in higher 5 year OS compared to 5 year OS for BV as reported in Gopal et al. (2015)[1] ([redacted] versus [redacted] respectively). Therefore the use of Gopal et al. (2015)[1] for the ITT population in the base case, could be considered somewhat conservative. Overall the ERG found KEYNOTE-087[5] to lack robustness given that it is a non randomised and single arm study, however it was useful to determine the impact of using an alternative OS data source on the ICER. As the company did not provide results for each subgroup the ERG subsequently conducted scenaro analyses for the SCT-2L, SCT-3L+ and SCT+3L+ subgroups using KEYNOTE-087[5] as the relevant OS source for both treatment arms. It was not possible for the ERG to obtain OS data for each individual cohort within KEYNOTE-087[5], therefore results are based on the ITT population within KEYNOTE-087[5]. Results are outlined in section 6.2.2.

A further approach investigated by the company ('Alternative approach 2', Doc B section 3.3.3.1, company scenario 6) was to use a predictive equation (predicting OS from PFS). The hazard ratio of OS:PFS from Gopal et al. (2015)[1] was applied to the PFS hazard in KEYNOTE-204 to obtain estimates of the OS hazard in KEYNOTE-204. The full details of the approach were not supplied e.g. use of both a ratio of hazards and a ratio of cumulative hazards are mentioned. The company indicated that a previous appraisal (TA524[12]) accepted the plausibility of an association between PFS and OS. However the company also acknowledged that Gopal et al. (2015)[1] most closely generalises to a subgroup (SCT+3L+) of the KEYNOTE-204 population only, and furthermore (Document B p205) "since the Gopal et al. publication, a variety of subsequent treatments have been introduced in the R/RcHL pathway, like immune checkpoint inhibitors". The ERG anticipates there may be further inconsistencies between the populations that might require adjustment. The ERG agrees with the company that this approach lacked face validity.

Additional limitations surrounding the modelling of OS

- The company did not provide sensitivity analysis using alternative parametric fits for OS extrapolation. As such it is unclear what impact alternative fits have on the ICER. In order to address uncertainty, the ERG conducted additional scenario analysis for the SCT-3L- and SCT-3L+ subgroups using the log logistic curve for extrapolation in both treatment arms. See Section 6.2.1.21 and 6.2.2.

- A single set of distribution parameters informs the OS curves in both treatment arms and, as a result, these curves are varied in exactly the same way in the probabilistic sensitivity analysis. The ERG noted that this may not adequately reflect uncertainty surrounding the OS parameters: this uncertainty would be captured better by using two sets of OS parameters, one for each arm. These sets contained identical values in the deterministic analysis, but were varied separately in the ERG probabilistic analysis (see Section 6.2.3.1).

4.2.6.2. Progression free survival

SCT-2L subgroup

In the company's base case, PFS for chemotherapy (IGEV) was based on pseudo-IPD, obtained from a digitized Kaplan-Meier curve in Balzarotti et al. (2016)[2] using a method developed by Guyot et al. (2012).[27]

Parametric distributions were fully-fitted to the pseudo-IPD, for the purpose of interpolation and extrapolation in the chemotherapy arm. The ERG noted an inconsistency here when compared

with the piecewise approach favoured by the company in PFS modelling elsewhere. The following parametric distributions were considered: the exponential, Weibull, Gompertz, log-logistic, log-normal and generalised gamma.

The relative statistical fit of the distributions was assessed using AIC and BIC scores, with the log-normal providing the best fit. PFS in the pembrolizumab arm was then modelled by applying the hazard ratio obtained from the MAIC to the PFS curve for IGEV. The hazard rate from the MAIC is obtained by Cox regression, with an implicit assumption of proportional hazards. The ERG noted that the inferred survival function for pembrolizumab (Appendix N.1 p202) also depends on an assumption of proportional hazards. The Weibull is a proportional hazards model but the log-normal is not. Also, the Weibull provided the second-best fit to the pseudo-IPD after the log-normal with only a slightly reduced relative fit (difference in AIC of [redacted]) (Appendix N table 77). The ERG considered the use of the fully-fitted Weibull distribution for modelling PFS in a scenario.

The ERG noted potential concerns surrounding the use of clincal effectiveness data from the MAIC to generate cost-effectiveness results for this subgroup. As noted in Section 3.4 the MAIC was associated with several limitations which introduce uncertainty and imprecision surrounding the reported HR for pembrolizumab (imprecision expressed by the wide confidence interval). Furthermore, it was unclear whether the assumption of proportional hazards held. The company acknowledged this (see p203 and p204 of the company's Appendices document) and therefore conducted a scenario using clinical data derived from a post hoc subgroup analysis of KEYNOTE-204.[3]

The ERG further noted that the company had assumed the clinical effectiveness of IGEV is generalisable to all chemotherapies, however clinical data was not supplied to support this assumption. Due to these uncertainties, the base case cost effectiveness results for this subgroup should be interpreted with caution.

As noted previously, a trial-based scenario analysis was also presented, with clinical data derived from a post hoc subgroup analysis of KEYNOTE-204.[3] For this scenario analysis, independent semi-parametric models were fitted to each arm, with BV used as a proxy for the chemotherapy comparator. The same method was used to identify break-points as for the other subgroups, with a break-point at Week 26 chosen based on visual inspection of the cumulative hazards plot. The best fitting distribution to the data beyond Week 26, the exponential, was not chosen for the parametric extrapolation, because the hazards were not found to be constant.

The log-normal, the second-best fitting distribution for the comparator arm, was chosen for the trial-based scenario.

Although this scenario anlaysis was useful, the ERG noted that assuming comparable efficacy between BV and IGEV was a major simplifying assumption that was not underpinned by clinical clincal data, and therefore preferred the company's base case proportional hazards approach, despite its limitations.

SCT-3L+ and SCT+3L+ subgroups

Clinical data used to estimate PFS for both pembrolizumab and BV were derived from post-hoc subgroup analyses of KEYNOTE-204.[3] The ERG noted that small patient numbers within each of these subgroups may introduce uncertainty in the results, however direct comparative data versus a relevant comparator was considered a strength.

In order to model long-term PFS, the assumption of proportional hazards was assessed. The log-cumulative hazards for each arm were plotted and the ratio of hazards was not found to be constant with respect to time. Hence, the company opted to fit independent semi-parametric models, where data from a Kaplan-Meier curve was used up to a cutpoint after which a parametric curve was employed, to each treatment arm, an approach discussed by e.g. Latimer et al. (2011).[28]

Chow tests were conducted at multiple time points to detect structural changes in PFS. The ERG noted that while the Chow test can be used to assess whether a single structural-break occurs at a known time point, it is not recommended for detecting time points at which structural-breaks may occur.[29] The break-points were identified through visual inspection of the test statistics plotted against time for each treatment. As the degrees of freedom or reference lines were not shown on the plots, the ERG could not determine whether the test statistics were statistically significant. Prominent changes in the plotted test statistics were identified at Weeks 26 and 52 for the pembrolizumab arm and at Week 52 for the BV arm.

Cumulative hazards plots were reviewed before the break-points were selected. The ERG noted a substantial increase in hazard around Week 12 in both arms, with smaller increases approximately every subsequent 12 weeks. This may be due to the dates of the tumour imaging data assessment (the first of which occurs around eight to 10 weeks after the initial dose), and subsequent checks for sustained response, rather than periodic increases in the proportion of patients with progressed disease. It could be argued that a smoother modelling approach for the trial period would be preferable in order to prevent sudden steep drops in modelled PFS, which would be unlikely to occur in clinical practice with this patient population.

A delayed treatment effect was suspected for pembrolizumab based on prior immunological knowledge, with the full benefit believed to well-established within the first six months, and so break-points of less than 24 weeks were avoided. However, an investigation of the hazards in the first six months might have indicated a time point by which the treatment effect had become fully-established, with a break-point of less than 24 weeks separating the time period in which the effect was fully established from that in which it was not. The Kaplan-Meier plots were also reviewed to ensure that at least 10 events occurred following the potential break-points.

A semi-parametric piecewise modelling approach was used in the company's base case and SCT-3L+ & SCT+3L+ subgroup analysis, with a break-point at Week 52. The KEYNOTE-204 Kaplan-Meier estimators were used to model PFS directly until Week 52, with a log-normal distribution fitted to the data beyond Week 52 used for parametric extrapolation. Scenarios with break-points at Week 0 or Week 26 (Scenarios 3 and 4, respectively) were considered, along with the following alternative distributions: the exponential, Weibull, Gompertz, log-logistic and generalised gamma.

The ERG noted that, with a break-point at either Week 26 or Week 52, the differences among the distributions in AIC and BIC scores were small, indicating that there was little difference in the relative statistical fit to the data (according to Burnham and Anderson (2002)[30] a rule of thumb is that models with differences in AIC of less than 2 cannot be distinguished). But the results of scenario analyses using alternative distributions for modelling PFS were not presented in the submission.

A break-point at Week 0 or at Week 26 was not selected for the company's base case involving the KEYNOTE-204 ITT population, since the modelled five-year PFS obtained with the best fitting distribution at those break-points (prior to convergence of the generalized gamma for the BV arm) was lower than expected: the clinical experts consulted by the company gave five-year PFS estimates of 15% for patients with prior ASCT and 10% for those ineligible for transplant. These estimates were higher than the five-year PFS estimate of 5% for third-line patients, suggested by the clinical expert consulted by the ERG.

Week 52 was chosen as the break-point in the SCT-3L+ and SCT+3L+ subgroup analysis by the company for consistency with the company's base case. This means that, while the Kaplan-Meier estimators were used directly for a longer period, fewer trial data informed the parametric extrapolation than would have been the case had an earlier break-point been selected.

The log-normal was the second-best fitting distribution to the data beyond Week 52 based on AIC and BIC scores. The best fitting distribution, the exponential, was not chosen for the

parametric extrapolation, because the company stated that the hazards were not found to be constant. This was also the case for the data beyond Week 26.

In the ERG base-case, the log-normal distribution fitted to the data beyond Week 26 was selected. The ERG regarded that this was a reasonable and appropriate choice of distribution across both 3L+ subgroups, balancing parsimony in model fit and across subgroups and accounting for the pattern of hazards following the cutpoint. The earlier break-point means that more of the trial data inform the parametric extrapolation, which may introduce less uncertainty. Since these data were collected after the first six months of the trial, the treatment effect should have been well-established (company response A8a).

The semi-parametric piecewise modelling approach was used in the ERG base case, as the log-cumulative hazard plots for the SCT-3L+ & SCT+3L+ subgroups could not be well approximated by straight lines (see Figure 3 and Figure 4).

Figure 3. Log-cumulative hazard plot for the SCT+3L+ subgroup

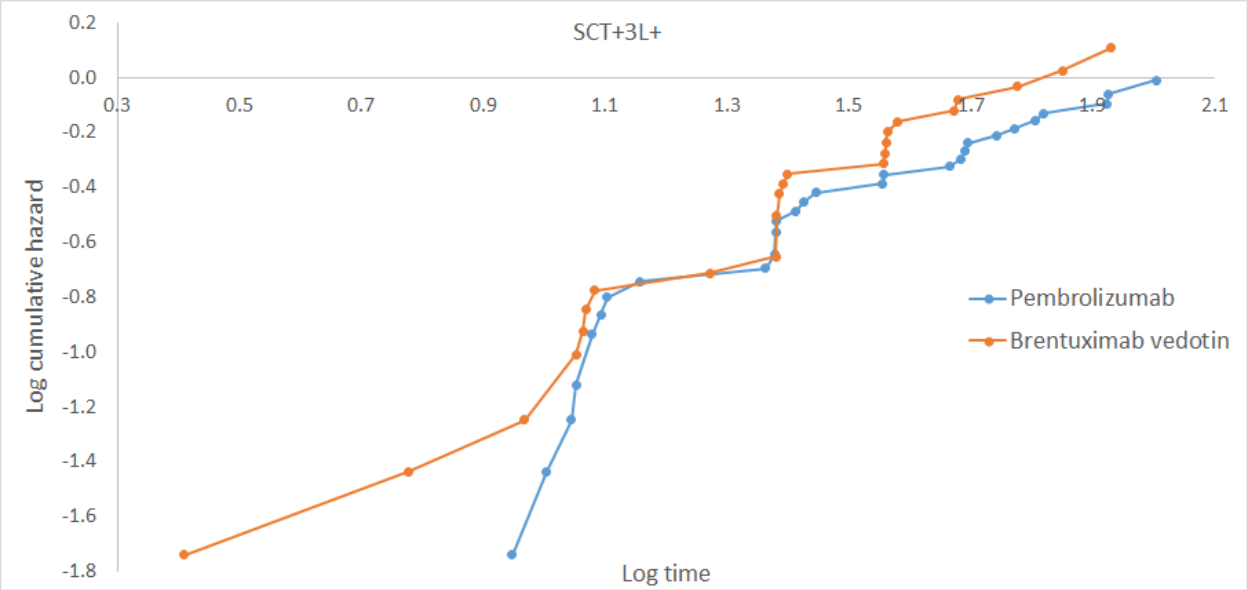
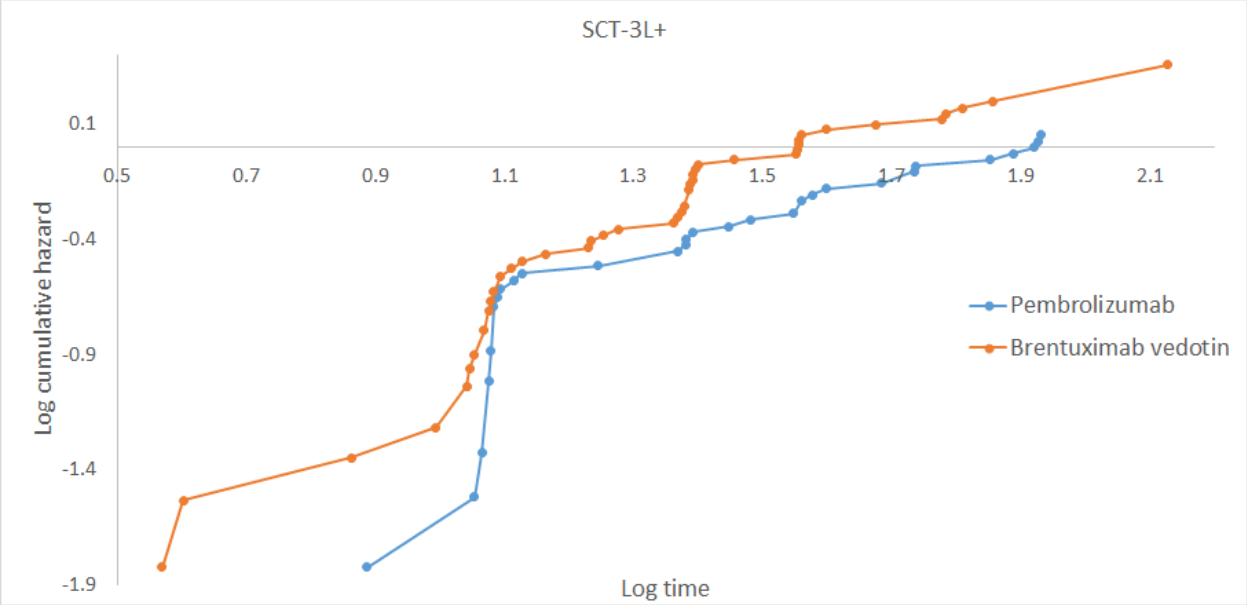


Figure 4. Log-cumulative hazard plot for the SCT-3L+ subgroup



An alternative approach would be to use a more flexible model, such as the distributions proposed by Jackson et al. (2010),[31] of which the generalized gamma is a special case. The generalized gamma provided a much better statistical fit to the full SCT-3L+ subgroup data than the other parametric distributions considered (as assessed by AIC and BIC statistics), as well as the best fit to full SCT+3L+ subgroup data for the pembrolizumab arm. The ERG conducted a scenario with PFS modelled by full-fitted generalized gamma distributions for each arm (i.e. with a break-point at Week 0).

In the company's SCT-3L+ subgroup analysis, the modelled five-year PFS was [redacted] and [redacted] for pembrolizumab and BV, respectively. For the SCT+3L+ subgroup, the modelled five-year PFS was [redacted] and [redacted] for pembrolizumab and BV, respectively.

Uncertainty surrounding the maintenance of pembrolizumab treatment effect on PFS

In the base case the company assumed that after patients discontinue treatment in Year 2, PFS for pembrolizumab would be maintained i.e. efficacy did not diminish after stopping treatment. Due to the lack of clinical data supporting this assumption, the ERG asked the company to provide a scenario analysis which incorporated a waning in PFS treatment effect for pembrolizumab after treatment discontinuation (from Year 3) until no difference in PFS was observed between treatments by Year 5. A similar approach had been used in NICE TA655[32] for assessing uncertainty surrounding OS, given limited long term clinical evidence.

However, the company did not provide this analysis, and stated that the scenario would not be appropriate to conduct on the basis that a highly conservative approach was already adopted in the modeling of OS. The ERG acknowledged that the company's base case approach of assuming no diffference in OS could be considered conservative, however, as noted above there is uncertainty surrounding the maintenance of pembrolizumab PFS benefit after patients stop treatment. The ERG was of the opinion that exploratory analyses incorporating a waning in pembrolizumab PFS treatment effect would be useful and therefore have conducted this scenario analysis for each subgroup (see Section 6.2.1.3).

4.2.7. Health-related quality of life

The company's base case analysis included disutilities associated with grade 3-5 adverse events, which are outlined on p212 and p213 of the CS. Due to the absence of disutility data from KEYNOTE 204,[3,4] the list of events and durations were based on previous NICE TAs and published literature. Disutilities for several adverse events including anaemia, diarrhoea and neutropenia were based on the average of values reported accross different data sources. To derive treatment specific disutility for both pembrolizumab and the comparator, disutilities associated with each adverse event were multiplied by the treatment specific rates from the ITT population in KEYNOTE 204 (see Table 118 in the CS). For the SCT-2L subgroup, chemotherapy (IGEV) adverse event rates were derived from NICE TA462[13] for Nivolumab, based on published study by Santoro (2007).[33] Santoro et al. (2007)[33] was an Italian prospective study designed to assess response rates, toxicity and stem cell mobilisation in 91 patients with refractory or relapsed Hodgkins Lymphoma.

The ERG noted that the company had applied the ITT adverse event rates to the SCT-3L+ and SCT+3L+ subgroups. Given the availability of subgroup data, it could be argued that these data should have been used. The ERG noted that adverse event rates were broadly similar between pembrolizumab and BV (based on subgroup data provided by the company during the clarification process), although for the SCT+3L+ subgroup, patients on pembrolizumab appeared to experience more infections and infestations compared to those on BV ([redacted] vs [redacted] respectively).

Overall, adverse events and associated disutilities did not appear to be a key driver of incremental QALYs within this submission, due in part to the 'front loading' of disutilities, whereby they were applied to Cycle 0 only. The company justified this approach on the basis that it has been used previously in NICE TA462[13] and TA540.[14] During the clarification process, the company noted an error surrounding the application of adverse events within base case the economic model and therefore provided updated results which are reflected in 5.1.1.

The company undertook a systematic literature review to identify studies reporting health-related quality of life or utility values (Section 4.1). However, determined the use of utility values from the KEYNOTE-204[3,4] study to be most aligned with the NICE reference case. Utility values used in the company's base case were derived from the ITT population within the KEYNOTE-204 study (Table 12). Values were elicited directly from patients using the EQ-5D-3L, which is considered an appropriate quality of life measure and reflects NICE guidance. Questionnaires were completed every 12 weeks from Cycle 1 (baseline) until disease progression or up till one year whilst the patient is on treatment. The valuation set used to convert the EQ-5D-3L health states into a single summary index (utility value) was based on UK public preferences using the time trade off (TTO) method from Dolan et al. (1997),[26] which elicited values from 3,395 members of the UK population.

Table 12: Base case utility values

Treatment	PFS utility	PD utility
Pembrolizumab	[redacted]	[redacted]
BV	[redacted]	[redacted]
Pooled utilities	[redacted]	[redacted]

Abbreviations: BV, brentuximab vedotin; PD, progressed disease; PFS, progression-free survival

The ERG acknowledged that using utility values elicited directly from patients within KEYNOTE-204[3,4] (as opposed to published literature sources) may be considered a strength; however, there are several uncertainties surrounding the approrpiateness of the progressed disease utility values which should be highlighted. These include the following:

- Utility values for progressed disease based on only two time points within 30 days:

  As patients in KEYNOTE-204 completed EQ-5D-3L questionnaires up to one year or until progression, it was unclear how the company captured utility for those in the progressed disease health state. The company was asked to comment and subsequently noted that patient reported outcomes (PRO's) were obtained at discontinuation and at the 30-day safety follow up visit. The ERG noted that the 30-day time frame used to estimate PD utility is short and unlikely to sufficiently capture changes in QoL.

- PD utility values were derived from fewer patients than the progression free health state:

  Values for the progressed disease health state were based on [redacted] patients in the pembrolizumab arm and [redacted] patients in the BV arm. This is considerably lower than the patient numbers used to estimate values for the progression free health state ([redacted] patients

and [redacted] patients in the pembrolizumab and BV arms respectively). As such, due to the relatively small patient numbers, utility values for the progressed disease health state may be associated with increased uncertainty.

- The progressed disease utility value for pembrolizumab appears to lack face validity:

  Clinical opinion to the ERG noted that the progressed disease utility value for pembrolizumab did not appear to be plausible. It was acknowledged that the utility decrement of moving from the progression free to the progressed disease health state is likely to be considerably higher than ([redacted]). Therefore, the value may not reflect the true quality of life burden associated with disease progression.

Given the concerns outlined above surrounding the progressed disease utlity value for pembrolizuamb, the ERG was of the opinion that the base case incremental QALY gain associated with pembrolizumab was subject to uncertainty and likely to be overestimated. The company provided a scenario analysis for the ITT population, which used the pooled value of [redacted] for the progressed disease health state (applied to both treatment arms). This resulted in a [redacted] incremental QALY gain reduction for pembrolizumab (from [redacted]). Although this served as a useful analysis, scenario analyses were not provided for the individual subgroups, which was considered a major limitation. Furthermore, the pooled value may not be appropriate to use for both treatments given that the progression free utility value for BV is [redacted]. This suggests a minimal reduction in quality of life upon disease progression for BV patients, which lacks plausibility.

The ERG noted that in NICE TA524[12] a lower utility value was used for the estimation of PD i.e. 0.38, which was derived from a published study by Swinburn et al. (2015).[34] Within this study utility was estimated for patients with R/R Hodgkin lymphoma and anaplastic large cell lymphoma. The ERG found that the PD utility value estimated by Swinburn et al. (2015),[34] was not particularly robust, given that they were not elicited directly from patients but rather from a relatively small sample of the UK population (n=100) using vignettes. Therefore the company's decision to not use Swinburn et al. (2015),[34] within their base case analysis seems justifiable. In TA524,[12] a scenario analysis was provided which estimated PD utility based on the Checkmate205[35] study, a single-arm study of nivolumab in patients with cHL following failure of ASCT. Within the study QoL data were collected from nivolumab treated patients using the EQ-5D. The PD value for these patients was estimated to be 0.715 and is outlined in SMC 1240/17.[36]

Given the limitations surrounding the PD utility value for pembrolizumab and in order to adequately test uncertainty, the ERG suggested a more reasonable approach was to remove the

difference in PD utility between treatments (whilst retaining in the PFS health state) (Section 6.2.1.2). This approach retained the treatment specific utility associated with pembrolizumab and BV in the PFS health state (observed in KEYNOTE-204[3,4]), whilst addressing uncertainty surrounding the PD state value. For this scenario the BV value for PD (was applied to both treatment arms) as it appeared to better reflect the QoL of patients whose disease had progressed and is similar to the value reported in SMC 1240/17.[36]

As an exploratory analysis, the ERG conducted an additional scenario anlaysis for each subgroup which removed treatment specific utility differences from the model (by applying BV utilities from KEYNOTE 204[3,4] to both treatment arms). This analysis was considered to be somewhat pessimistic given that direct QoL trial data are ignored (see Section 6.2.1.1).

4.2.8. Resources and costs

4.2.8.1. Medicine acquisition costs

The company noted that pembrolizumab is supplied in 100 mg vials and the list price per vial is £2,630. The ERG confirmed that this was reflective of BNF pricing. Treatment costs in the model were based on a fixed dose of 200 mg every three weeks resulting in a cost of £5,260. The dosing schedule appeared to be in line with pembrolizumab dosing in KEYNOTE-204[3,4] and the SmPC.

A patient access scheme (PAS) was submitted by the company, which reduced the price of pembrolizumab by [redacted]. The company stated that the current CAA discount in place for pembrolizumab is [redacted], however, as the discount will increase to [redacted], on TAG publication of pembrolizumab ID1140 for untreated metastatic or unresectable recurrent squamous cell head and neck cancer, the [redacted] will be used and is considered appropriate. The cost of pembrolizumab per three-week treatment cycle (with PAS) was therefore estimated to be [redacted].

The company provided scenario analysis results for the ITT population using a dose of 400 mg administered every six weeks, which did not have a meaningful impact on the ICER (see p250 of the CS). The company stated this this alternative dose forms part of draft SMPC, which has yet to receive CHMP opinion. For completeness, the ERG has considered this alternative dose within a scenario analysis for each subgroup (see Section 6.2.1.7).

Within the economic model, treatment costs were further adjusted to reflect the dose intensity within KEYNOTE-204 (98%). This was applied to both the pembrolizumab and BV treatment arms. For completeness the ERG conducted a scenario analysis for each subgroup assuming 100% dosing intensity in both treatment arms (Section 6.2.1.6).

For BV, the list price was estimated to be £2,500 per 50mg vial, as per the BNF. The company estimated the cost per cycle based on the sum product of the number of vials used and cost per vial. The ERG understood that a patient would therefore require 3 vials (administered at 1.8 mg/kg and assuming patient weight of 77 kg). The cost per treatment cycle used in the economic analysis was estimated to be £7,365, when adjusted for trial based dose intensity (see p220 of the CS). The company assumed drug wastage in the model, which was considered reasonable. Based on clinical opinion to the ERG, it was noted that vial sharing is unlikely to reflect current practice given concerns surrounding treatment shelf life/storage and small patient numbers.

For the SCT-2L subgroup, the company acknowledged that there is range of multi agent chemotherapy agents available for use within this subgroup of patients and that frequency of use is likely to differ across UK centres. The company therefore used a published study by Eyre et al. (2017)[23] to inform the list of potential regimens. The ERG note that this study was relatively recent (2017) and UK based, which is considered a strength. However, patients included in the study had two prior lines of chemo therapy and had received BV. It is therefore unclear whether treatment regimens from this study are fully generalisable to the SCT-2L population (see Table 82, p210 of the company's Appendices document for list of chemotherapy regimens used in the company's base case).

The proportion of patients receiving each treatment regimen was based on Eyre et al. (2017),[23] but amended using clinical opinion to reflect recent changes in treatment use (see Table 83 on p211 of the company's Appendices document). Clinical opinion to the ERG noted that bendamustine is used in the UK within this patient population. However, the company did not include this as a plausible treatment option, which somewhat limits the validity of the company's treatment list. Treatment acquisition costs were derived from the drugs and pharmaceutical electronic Market Information Tool (eMIT) and seemed to be largely accurate. The ERG noticed a minor error with respect to the cost of vinorelbine, which the company estimated to be £3.67 per 10 mg/1ml solution; however, the price in eMIT was £36.71. For completeness the ERG has amended this cost to reflect eMIT pricing, which is included in the ERG preferred base case (see Section 5.3). Furthermore, it was not possible to verify the cost of chlorambucil (£1.71) using eMIT. When crosschecked with the BNF, the price was higher (£42.87).

As noted in 4.2.8.3, for the ITT analysis, the unit costs of subsequent treatments were included and derived using eMIT (see p224 of the CS). Overall, costs were largely accurate though several costs could not be validated using eMIT. The ERG considered that potential variation in unit cost estimates for chemotherapy treatments, may not be a primary concern given the minor nature of these costs (with respect to the relatively high acquisition cost of pembrolizumab) and that the

ITT analysis and list of subsequent treatments is not considered to be reflective of each subgroup. See Section 4.2.8.3 for further commentary on subsequent treatment use.

4.2.8.2. Time on treatment (ToT)

According to the SmPC for pembrolizumab,[37] treatment should be continued until progression or unacceptable toxicity. However, it is worth noting that the economic model incorporates a two-year stopping rule, whereby all patients were assumed to discontinue treatment after two years. The company highlighted that this was in line with the KEYNOTE-204 protocol, where treatment was mandated to stop at 35 cycles/105 weeks. The ERG noted that this assumption was used in previous NICE technology appraisal guidance including TA428 (pembrolizumab for PD-L1 NSCLC after chemotherapy).[38] Within TA428, clinical experts commented that the decision to stop treatment would be made between the clinician and the patient, and that the number of patients likely to have treatment after two years would be small. Clinical opinion to the ERG advised that the stopping rule is likely to be adhered to in practice, given that it is part of the marketing authorization for pembrolizumab. Overall, the inclusion of a two-year stopping rule appeared to be consistent with previous NICE technology appraisals and clinical opinion.

SCT-2L subgroup

The same approach was used for modelling ToT in the pembrolizumab arm of the SCT-2L subgroup as for the other subgroups. Due to the lack of ToT data for chemotherapy (IGEV), ToT was set equal to PFS for the comparator. Given the lack of available evidence, the ERG considered this to be reasonable and used the same approach for the ERG base case.

SCT-3L+ and SCT+3L+ subgroups

ToT was modelled separately for the pembrolizumab and BV arms using a semi-parametric approach, which allowed ToT to be extrapolated beyond that observed in KEYNOTE-204[3,4] until the maximum duration of treatment (assumed to be 35 cycles/100 weeks). The company used the same modelling approach for the SCT-3L+ and SCT+3L+ subgroup analyses as for the ITT population, differing only in the portion of the data used. The KEYNOTE-204 Kaplan-Meier estimators were used directly until Week 80, with an exponential distribution fitted to the data beyond Week 80 used for parametric extrapolation.

The break-point at Week 80 was chosen as KM data for the ITT population was available until Week 88: the company wished to use the KM estimators to model ToT directly for as long a period as possible, while ensuring that what was considered to be an adequate number of events remained for fitting a parametric distribution for extrapolation. The ERG noted that while KM data

were available at least until Week 88 for the SCT-3L+ subgroup, the last recorded event in the BV arm for the SCT+3L was at 82.6 weeks.

The company selected the exponential distribution for extrapolation on the basis that it produced the lowest AIC & BIC statistics. Information on the assessment of hazards for ToT was not available in the CS.

Uncertainty surrounding the company's ToT modelling approach

The company assumed a maximum treatment duration of 35 cycles (105 weeks) for both pembrolizumab and BV in the SCT-3L+ and SCT+3L+ subgroups, which does not appear appropriate. Although the use of 35 cycles was consistent with the two-year pembrolizumab stopping rule, based on the SmPC for BV, the maximum number of cycles that treatment should be given is 16. Assuming 35 cycles for BV therefore potentially overestimates the medicine costs. The company did conduct a scenario analysis which assumed a maximum treatment duration of 16 cycles for BV (this was applied to costs only as efficacy was assumed to be maintained for 35 cycles (see p253 in the CS), The ERG considered that 16 cycles should be used and therefore this assumption forms part of the ERG's base case (Section 6.2.1.19 and Section 6.2.2).

It was recognised that using KM data to Week 80 may reduce extrapolation uncertainty; however, in order to be consistent with the company's PFS modelling approach, the ERG considered that estimating ToT using a 26-week cut point preferable. This was because ToT should be largely coterminous with PFS, as progression would often trigger a change in treatment. In order to determine the impact of alternative ToT assumptions on the results, the ERG conducted additional scenario analyses whereby ToT is based on KM data from KEYNOTE-204, as well as using an alternative cut point at 26 weeks. Parameters for ToT with cut point at 26 weeks was provided only for the ITT population during clarification; still, the ERG regarded that this would present more reliable and appropriate estimates (Section 6.2.1.10, Section 6.2.1.11 and Section 6.2.2).

The company did not provide scenario analyses using alternative distributions for ToT. Although the exponential distribution selected by the company exhibited the lowest AIC/BIC score, there was minimal difference between the scores for each distribution. For completeness, the ERG conducted additional scenario analyses using alternative ToT distributions (Section 6.2.1.12 and Section 6.2).

4.2.8.3. Subsequent treatment costs

The ERG noted that subsequent treatment costs were likely to have an impact on the ICER for pembrolizumab (see Sections 6.2.1.18 and 6.2.2). Subsequent treatment costs were included in the model and were assumed to apply once patients entered the PD health state. For the ITT analysis, the list of subsequent treatments was based on the ten most commonly used subsequent treatments within KEYNOTE-204.[3,4] The list of subsequent treatments and proportion of patients receiving each are outlined in Table 13. When estimating subsequent treatments and proportions for each subgroup, the company assumed these to reflect UK clinical practice (see Table 14). Overall, the ERG noted several concerns surrounding the company's base case subsequent treatment estimates which introduce uncertainty and may not reflect appropriate treatments provided in within current clinical practice.

Table 13. Base case subsequent treatments (ITT analysis)

Subsequent treatment(s)	After failing	
	Pembrolizumab	BV
BV	[redacted]	[redacted]
Nivolumab	[redacted]	[redacted]
Pembrolizumab	[redacted]	[redacted]
Bendamustine	[redacted]	[redacted]
Bendamustine + BV	[redacted]	[redacted]
Etoposide+melphalan	[redacted]	[redacted]
Cyclophosphamide + fludarabine phosphate	[redacted]	[redacted]
Bendamustine + gemcitabine + vinorelbine tartrate	[redacted]	[redacted]
Cisplatin + cytarabine + dexamethasone	[redacted]	[redacted]
Carmustine + cytarabine + etoposide + melphalan	[redacted]	[redacted]
None	[redacted]	[redacted]

Abbreviations: BV, brentuximab vedotin; ITT, intention to treat

Table 14: Base case subsequent treatment assumptions (subgroup analyses)

	Subsequent treatments
SCT-2L	
Pembrolizumab	100% receive BV
IGEV	
SCT-3L+	

	Subsequent treatments
Pembrolizumab	100% BV
BV	100% pembrolizumab
SCT+3L+	
Pembrolizumab	100% BV
BV	100% nivolumab

Abbreviations: BV, brentuximab vedotin

The ERG highlighted uncertainties surrounding subsequent treatments, as follows:

- As noted previously in Section 4.2.3, the ERG did not consider an ITT population to be appropriate for decision making, therefore the subsequent treatments and proportions used for this analysis should be interpreted with caution.

- The ERG noted discrepancies between the subsequent treatment assumptions applied in the model and those outlined in the CS for two key subgroups (SCT-2L and SCT-3L+), which led to differences between the modelled results and those reported in the CS. The company was asked to comment and noted that the default results for SCT-2L and SCT-3L+ in the model did not match the CS results as subsequent therapies were changed manually in the model, before copying results. The ERG considered the company's response helpful and that it clarifies the disparity between results.

- There was concern surrounding the use of pembrolizumab as the subsequent treatment for patients who fail on BV in the SCT -3L+ subgroup. As pembrolizumab at this line of treatment is included within the Cancer Drugs Fund (CDF), clinical opinion was sought to determine what treatment would be given to patients who did not have the option to be treated with pembrolizumab. It was suggested that further chemotherapy (typically with a regimen that does not contain an anthracycline) should be considered. Such options included bendamustine alone, bendamustine+gemcitabine+vinorelbine, gemcitabine with Cis- or carboplatin and dexamethasone, ChlVPP (chlorambucil with vinblastin, procarbazine and prednisolone) or similar combinations. The ERG conducted a scenario analysis for this subgroup, which assumed that 100% of patients who failed BV went on to receive 'bendamustine' only (Section 6.2.1.18).

- The ERG noted that the handling of subsequent treatment in the SCT+3L+ subgroup appeared to be inappropriate, as the company assumed that 100% of patients who failed on pembrolizumab went on to receive BV, whilst 100% of patients who failed BV went on to

receive nivolumab. Based on a review of the treatment pathway for this subgroup, patients in both treatment arms should receive nivolumab as subsequent treatment (Section 6.2.1.18).

4.2.8.4. Monitoring, administration and resource use costs

The ERG acknowledged that monitoring and resource use were not considered to be a key cost driver within this submission. However, there were concerns surrounding the estimation of resource use for the PD health state, which requires comment.

The company stated that data pertaining to resource use for patients with R/RcHL were limited and therefore estimates were derived from a previously published NICE appraisal TA446 for BV.[39] Resource use costs were valued using 2018/19 NHS reference costs, which was an appropriate source. However, the ERG considered the company's PET scan cost (£775.51) to be higher than the cost quoted in the NHS reference cost guidance, which was estimated to be £506. Using a lower PET scan cost is unlikely to have any material impact on the ICER, and is therefore not a key concern.

Annual resource use for patients in the PFS health state was based on clinical expert opinion. Estimates therefore may be subject to some degree of uncertainty. The total cost per cycle was £64.27 (see Table 16 below for granularity). In TA446[39] resource use for the PD health state was assumed to be the same as for the PFS health state. The company has adopted the same assumption within the current submission, therefore the cost per cycle associated with progressed disease is also estimated to be £64.27. The ERG considered this to be a simplifying assumption which may not reflect current practice. Clinical opinion to the ERG noted that PD health state costs would be expected to be higher due to deterioration in quality of life and requirement for additional monitoring.

The company acknowledged this limitation within the CS and provided a scenario analysis which assumed patients in the PD health state would require higher resource use, based on clinical opinion to the company (see p232 of the CS). However, the ERG noted that results were provided for the ITT analysis only and not for each subgroup. Furthermore, the scenario analysis assumed that resource use would also decrease simultaneously for patients in the PFS state. Although health state resource use was not considered a key driver of the ICER, the ERG considered that the company's scenario analysis potentially underestimates monitoring and resource use costs for pembrolizumab, whilst overestimating these costs in the comparator arm. For completeness, the ERG conducted a scenario analysis for each subgroup, which applied higher resource estimates to the PD state only.

Table 15: Base case PFS and PD health state costs

Resource	Unit cost (£)	Unit cost source (NHS reference costs 2018-2019 code)[40]	Weekly usage	Cost per cycle	Resource use source
Outpatient attendance	173.39	303: Clinical Haematology, Consultant led follow-up attendance, non-admitted face to face	0.20	34.56	NICE TA446 Committee papers,[39] ERG Table 95 (p210)
Blood count	2.79	DAPS05: Haematology	0.20	0.56	
Biochemistry	1.10	DAPS04: Clinical Biochemistry	0.20	0.22	
CT scan	115.56	RD26Z: Computerised Tomography Scan, three areas with contrast	0.06	6.64	
PET scan	775.51	RN03A: Positron Emission Tomography with computed Tomography (PETCT) of more than three areas, 19 years and over	0.03	22.29	
Total cost per week (£)				**64.27**	

Abbreviations: CT, computed tomography; ERG, Evidence Review Group; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PET, positron emission tomography; TA, technology appraisal

Administration costs

For the SCT-2L subgroup, the company has outlined the unit costs for chemotherapy administration in Table 84 on p211 of the Appendices document. The ERG considered the unit costs to be reflective of NHS reference costs 2018/19 and appropriate for use.

Administration costs were calculated in the model by multiplying the number of administrations for each treatment regimen (accounting for both the first and subsequent administrations per cycle) by the relevant cost per administration. For the SCT-3L+ and SCT+3L+ subgroups, the company assumed that both pembrolizumab and BV were administered via IV infusion over 30 minutes (as per the SmPC for each treatment) and used the National Tariff of Chemotherapy Regimens List and NHS reference costs 2018/19, to estimate costs associated with administration. Overall, the company's handling of administration costs within the CS seemed reasonable.

Adverse event costs

On p235 of the CS, the company state that subgroup specific grade 3-5 AEs from KEYNOTE 204[3,4] (with an incidence of ≥2% in any arm) were used to estimate adverse event costs in the base case. The complete list of adverse events are outlined in Table 134 on p236 of the CS. NHS reference costs 2018/19 were used as appropriate to estimate the unit cost of each event,

however NICE TA462[13] was used to estimate the cost associated with nausea vomiting and weight increase.

During the clarification stage the company noted that several AE costs within the model including pneumonia, pneumonitis, rash, thrombocytopenia, vomiting and increased weight) were different from those specified in the CS (Document B, Section B.3.5.6, pp.235-36). The company presented corrected results for the ITT population in response to clarification question A13. Overall, the ERG noted that adverse event costs were only applied to cycle 0 in the model and therefore did not have a material impact on results.

Stem Cell Transplant

In terms of stem cell transplant (SCT), patients in the PFS health state were eligible to undergo either auto-SCT or allo-SCT, based on treatment specific rates from the pivotal study KEYNOTE-204.[3,4] SCT rates used in the base case analysis were derived from subgroup data and are outlined in Table 16. The ERG noted that patient numbers within each subgroup were small, therefore the rates may be subject to uncertainty. The cost associated with an auto-SCT and allo-SCT was estimated to be £22,368 and £114,234 respectively. Costs were based on a published study by Radford et al. (2017),[41] which reported cost and resource use in 40 cHL patients who had failed after auto SCT. Radford et al. (2017)[41] was considered to be the preferred source in TA462[13] for nivolumab. The ERG also noted that this study has been used previously in NICE TA540[14] for pembrolizumab. Costs were inflated to reflect 2018/19 prices as appropriate.

Table 16: Base case SCT rates (derived from KEYNOTE 204 subgroup data)

	Auto-SCT	Allo-SCT
SCT-2L		
Pembrolizumab	[redacted]	[redacted]
IGEV (assumed to equal BV)	[redacted]	[redacted]
SCT-3L+		
Pembrolizumab	[redacted]	[redacted]
BV	[redacted]	[redacted]
SCT+3L+		
Pembrolizumab	[redacted]	[redacted]
BV	[redacted]	[redacted]

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; SCT, stem cell transplantation

Due to the uncertainty surrounding the validity of the base case SCT rates, the ERG considered it pertinent to undertake further sensitivity analysis. SCT rate is a notable, but not central, factor

affecting ICERs. A scenario analysis has therefore been conducted for each subgroup which sets SCT rates equal between groups (see Section 6.2.1.5). Given that base case SCT rates are subject to uncertainty and are associated with high costs, the ERG's preferred base case was to set these rates equal to each other between arms. See Section 6.2.1.5 for further discussion on how this scenario analysis impacts incremental costs and QALYs in each subgroup.

Terminal care costs

The company applied a once off cost of £4,462 to each death event in the model. The cost was based on a published study by Brown et al. (2013)[42] and represents the weighted average of hospital, hospice and home setting costs. Brown et al has been used to estimate terminal care costs previously in NICE TA540[14] and have been updated. The ERG noted that terminal care costs were not considered a key driver within the model.

5. COST-EFFECTIVENESS RESULTS

5.1. Company's cost-effectiveness results

5.1.1. Base case results

5.1.1.1. ITT population

Results of the company's base case analysis were presented as an ICER for pembrolizumab compared to BV. The results presented in the CS (Document B, Section B.3.7.1, p240) were based on incorrect costs for the AEs: the costs incurred for AEs (pneumonia, pneumonitis, rash, thrombocytopenia, vomiting and increased weight), applied in the model were different from those specified in the CS (Document B, Section B.3.5.6, pp.235-36). The company presented corrected results for the ITT population in response to clarification question A13. The model version submitted to the ERG following this correction is referred to as "revised model" in the sections below. The total and incremental costs, life years (LYs), and QALYs, and the ICER were replicated in Table 17 below. A patient access scheme (PAS) of [redacted] was applied to the acquisition cost of pembrolizumab.

Table 17: Company base case deterministic results

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	[redacted]	-	
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year

Source: Company "revised model" (clarification response 5 November 2020)

Based on the results, pembrolizumab was considered the dominant treatment when compared to BV resulting in an incremental QALY gain of [redacted] and incremental savings of [redacted]. Incremental savings were mainly due to lower medicines acquisition costs associated with pembrolizumab. As noted throughout this report, pembrolizumab was not associated with a survival gain, therefore incremental QALYs versus BV stem primarily from a higher proportion of patients remaining in the progression free health state.

5.1.1.2. Subgroup results

The results for the subgroups following model revision, were also presented by the company in an appendix to its response to clarification questions.

SCT-2L

For the SCT-2L subgroup, results of the company's base case analysis were presented as an ICER for pembrolizumab compared to salvage chemotherapy (IGEV). Total and incremental costs, life years (LYs), and QALYs were presented in the CS (Document B, Section B.3.9.1, p257); however, they were subsequently updated as per the company's "revised model" as replicated in (Table 18) below. A PAS of [redacted] was applied to the acquisition cost of pembrolizumab.

Table 18: Company base case deterministic results: SCT-2L

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	Lys	QALYs	
Company base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	[redacted]	-	
IGEV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£53,581

Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year.

Source: Company "revised model" (clarification response 5 November 2020)

As noted above, pembrolizumab resulted in an ICER of £53,581 compared to salvage chemotherapy based on incremental costs of [redacted] and an incremental QALY gain of [redacted]. Incremental costs were mainly due to higher medicines acquisition costs associated with pembrolizumab. Pembrolizumab was not associated with a survival gain, therefore incremental QALYs versus IGEV stem primarily from a higher proportion of patients remaining in the progression free health state.

SCT-3L+

For the SCT-3L+ subgroup results, the company's base case analysis were presented as an ICER for pembrolizumab compared to BV. Total and incremental costs, life years (LYs), and QALYs were presented in the CS (Document B, Section B.3.9.3, p260), and were subsequently updated as per the company's "revised model" as replicated in (Table 19) below. A patient PAS of [redacted] is applied to the acquisition cost of pembrolizumab.

Table 19: Company base case deterministic results: SCT-3L+

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	Lys	QALYs	Costs (£)	LYs	QALYs	
Company base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	[redacted]	-	
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Key: BV, brentuximab vedotin; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year.

Source: Company "revised model" (clarification response 5 November 2020)

For this subgroup pembrolizumab was considered to dominate BV resulting in incremental savings of [redacted] and an incremental QALY gain of [redacted]. Incremental savings were mainly due to lower medicines acquisition costs associated with pembrolizumab. Pembrolizumab was not associated with a survival gain, therefore incremental QALYs versus BV stem primarily from a higher proportion of patients remaining in the progression free health state.

SCT+3L+

For the SCT+3L+ subgroup results, results of the company's base case analysis were presented as an ICER for pembrolizumab compared to BV. Total and incremental costs, life years (LYs), and QALYs were presented in the CS (Document B, Section B.3.9.2, p260); however, they were subsequently updated as per the company's "revised model" as replicated in (Table 20) below. A PAS of [redacted] is applied to the acquisition cost of pembrolizumab.

Table 20: Company base case deterministic results: SCT+3L+

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	[redacted]	-	
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Key: BV, brentuximab vedotin; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year

Source: Company "revised model" (clarification response 5 November 2020)

For this subgroup pembrolizumab was considered to dominate BV resulting in incremental savings of [redacted] and an incremental QALY gain of [redacted]. Incremental savings were mainly due to lower medicines acquisition costs associated with pembrolizumab. Pembrolizumab was not associated with a survival gain, therefore incremental QALYs versus BV stem primarily from a higher proportion of patients remaining in the progression free health state.

5.2. Company's sensitivity analyses

In addition to exploring the role of parameter uncertainty on the model results, the CS also reported several sensitivity analyses which explored the impact of alternative settings and assumptions. These are discussed further below.

Overall, the ERG considered the approach taken for sensitivity analysis to be appropriate.

5.2.1. One-way sensitivity analysis

The company conducted a deterministic one-way sensitivity analysis (OWSA) with the included parameters as presented in CS (Document B, Table 139). The CS stated that where data were available, parameters were varied using 95% confidence intervals, otherwise upper and lower bounds were varied by a standard error of 10% of the mean (base case) value.

A tornado plot was used to present the OWSA results in the CS (Document B, Figure 55) for pairwise comparison of pembrolizumab vs. BV for the ITT population, with the ICER as the outcome of interest. The plot showed the results were most sensitive to the PFS and PD health state utility values of pembrolizumab and BV and disount rate for outcomes. However, the OWSA results for the subgroups were not presented in the CS.

The ERG noted that the OWSA results were not impacted by the changes to the company's revised model.

5.2.2. Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) to explore the impact of parameter uncertainty when the model parameters' were varied as per the respective distributions (CS, Document B, Table 137). The PSA was run for 1,000 iterations.

The company's "revised model" presented updated PSA results provided in Table 21.

Table 21: Company PSA

Arm	Totals		Incremental		ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs	
Company presented probabilistic base case					
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	
BV	[redacted]	[redacted]	[redacted]	[redacted]	34,540
Company probabilistic base case – using correct model settings					
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	
BV	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-39,266)

Key: BV, brentuximab vedotin; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year

The ERG noted that the incremental costs were positive in the PSA and not aligned with the deterministic base case results. The ERG investigated the PSA macro but did not identify any errors and assumed that it might be due to incorrect model settings while running the PSA. Therefore, the PSA was re-run by the ERG using the correct settings and the results following the re-run are provided in (Table 21) above.

Further, as per the revised model, the company stated that at a willingness-to-pay threshold of £30,000 per QALY gained, the probability of pembrolizumab being cost-effective versus BV for the ITT population was 40%. However, the ERG noted that, when re-running the PSA with the correct model settings applied (as indicated above), the probability of pembrolizumab being cost-effective versus BV for ITT population changed to 92%.

In addition, the ERG noted that the PSA results were not presented for the subgroups in the CS. Details on the PSA for subgroups carried out as part of ERG additional analyses are given in Section 6.2.3.

5.2.3. Company's scenario analyses

The company conducted several scenario analyses to assess the impact of alternative settings and assumptions and the structural uncertainties on the base case results. Scenario analysis results were provided in the CS (Document B, Table 140).

Scenarios with alternative OS data increased the incremental QALYs of pembrolizumab vs BV whereas the scenario with pooled post-progression utility decreased the incremental QALYs. Scenarios with no vial sharing, alternative maximum number of cycles with BV and subsequent treatments based on KEYNOTE-204[3,4] excluding pembrolizumab increased the incremental costs whereas the scenarios with alternative resource use, subsequent treatments based on UK

market shares and alternative dosing for pembrolizumab decreased the incremental costs. In all the scenarios presented pembrolizumab remained dominant versus BV, in line with the base case.

The scenario analyses presented were limited in number and focused on the ITT population, with none exploring the differences in modelling the PFS and OS across the subgroups considered. The results of the scenario analyses did, however, highlight the influence of the data used to model and extrapolate overall survival, alternative assumptions on utilities and subsequent treatments.

5.3. Model validation and face validity check

The ERG found the company's cost-effectiveness model to be mostly free of errors, however some minor issues were noted; for example, use of inconsistent labelling of the Cholesky matrices, duplication of a parameter for Weibull fit, non-convergence with generalized gamma. These errors were either fixed by the company during clarification and were incorporated in the "revised model" provided in the clarification response or were found not to have any impact on the model results.

Briefly, the errors corrected are listed below:

- An error in the chemotherapy (IGEV) PFS meant that the proportion of patients progression-free in each arm at each time point did not correspond to the hazard at that time point. This error affected the SCT-2L subgroup analysis.
- An error in the maximum treatment cycle reference for BV meant that the maximum treatment cycle for pembrolizumab was used for both arms, regardless of the model settings. This affected the ITT population and SCT-3L+ and SCT+3L+ subgroups. However, the results for the company's base case and subgroup analysis remain unchanged, since the same maximum number of cycles was selected for both arms. Hence, these fixes are not shown in Table 22 below.
- As noted in Section 4.2.8.1, a minor error was noted in the company's model with respect to unit cost for vinorelbine. This error affected the SCT-2L subgroup analysis.

Table 22: ERG corrections to the company's subgroup analysis case

Preferred assumption	ICER when applied individually	Cumulative ICER £/QALY
SCT-2L subgroup (pembrolizumab compared to salvage chemotherapy (IGEV))		
Company base case	£53,581	£53,581
Error in chemotherapy PFS	£53,276	£53,276
Amended vinorelbine cost	£53,403	£53,099
ERG corrected company base case	£53,099	–

Key: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1. Exploratory and sensitivity analyses undertaken by the ERG

The ERG carried out a number of exploratory and sensitivity analyses. Table 23 summarises the scenario analyses as applied to each of the three subgroups: (SCT-2L, SCT-3L+, and SCT+3L+).

Table 23. Summary of scenario analyses by subgroup

		Subgroups		
#	**Scenario**	**SCT -2L**	**SCT -3L+**	**SCT +3L+**
1	Utility value for the PD health state	●	●	●
2	Equal PFS and PD utility values	●	●	●
3	Waning of pembrolizumab PFS treatment effect	●	●	●
4	Higher resource use in the PD health state	●	●	●
5	No difference in SCT rates between treatment arms	●	●	●
6	Dose intensity for pembrolizumab assumed to be 100%	●	●	●
7	Pembrolizumab administered 400 mg (every six weeks)	●	●	●
8	Time horizon increased to 50 years	●	●	●
9	KEYNOTE-087 source for OS data (pembro & comparator[a])	●	●	●
10	ToT for pembrolizumab based on KM data only	●	●	●
11	Alternative cut points for modelling ToT (26 wks)	●	●	●
12	Alternate parametric fit (log normal) for ToT (pembro & comparator[a])	●	●	●
13	Subsequent Tx based on subgroup data from KEYNOTE-204	●	●	●
14	Balzarotti et al. (2016) used to estimate OS (pembro & comparator[a])	●	●	NA
15	Balzarotti et al. (2016) for OS + alternative parametric fit	●	●	NA
16	Alternative parametric fit for PFS, applied to both pembro and IGEV	●	NA	NA
17	Combined analysis: PFS (fully parametric) and OS (KEYNOTE 087)	●-W	●-GG	●-GG
18	Subsequent treatments assumed to reflect UK practice	NA	●-ben[b]	●-nivo[c]
19	Reduction in maximum number of cycles of BV	NA	●	●
20	Fully parametric approach to model PFS (generalised gamma curve)	NA	●	●
21	Log-logistic parametric fit for Gopal et al. (2015) OS data (pembro & BV)	NA	●	●
22	Model PFS using different data cut point (26 weeks)	NA	●	●

Abbreviations: -ben, bendamustine; BV, brentuximab vedotin; CTx, chemotherapy; -GG, generalised gamma; NA, not applicable; -nivo, nivolumab; OS, overall survival; pembro, pembrolizumab; PFS, progression free survival; SCT, stem cell transplantation; ToT, time on treatment; Tx, treatment; -W, Weibull; wks, weeks

Notes: a Comparator: SCT-2L = IGEV; SCT-3L+ & SCT+3L+ = BV; b 100% of patients who fail pembro go on to receive BV AND 100% of patients who fail BV go on to receive bendamustine alone; c 100% of patients who fail pembro go on to receive BV AND 100% of patients who fail on BV go on to receive bendamustine alone

The following adjustments were relevant to the PSA and are not associated with a deterministic ICER:

- OS modelled separately for both pembrolizumab and the comparator treatment
- PFS HR varied using the 95% confidence interval from the MAIC.

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

The scenario analyses described in Section 6.1 are described in turn below. The impact on the ICER (Section 6.3) refers to the company's base case ICER including the ERG corrections detailed in Section 5.3.

6.2.1. Scenario analyses

6.2.1.1. Scenario 1: Utility value for the PD health state

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

The company's base case utility value for the pembrolizumab PD state was associated with uncertainty and considered implausibly high (see Section 4.2.7). This scenario analysis removes the difference in treatment specific values in the PD health state by applying the BV PD health state value ([redacted]) to both pembrolizumab and the comparator (IGEV [SCT-2L] and BV [SCT-3L+ and SCT+3L+]). The ERG considered this value to better reflect the quality of life for patients with PD. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.2. Scenario 2: Equal PFS and PD utility values

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

In this scenario it was assumed that pembrolizumab and the comparator treatment (IGEV [SCT-2L] and BV [SCT-3L+ and SCT+3L+]), were associated with the same PFS and PD utility values i.e. pembrolizumab was not associated with a treatment specific utility gain. Utilities were based on BV values within KEYNOTE-204.[3,4] The ERG recognised that this assumption may be highly conservative given that QoL data reported within the pivotal study detected treatment specific differences in utility. However, given the uncertainties surrounding these trial-based utilities (Section 4.2.7), and the sensitivity of the ICER to changes in utility, the ERG considered this

scenario analysis would address further uncertainty. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.3. Scenario 3: Waning of pembrolizumab PFS treatment effect

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.6, due to the lack of long-term clinical effectiveness data, there is some uncertainty surrounding the maintenance of pembrolizumab treatment effect with respect to PFS (after patients discontinue). The ERG noted that in the absence of long-term clinical efficacy data, scenario analyses which incorporate a waning in treatment effect are helpful to address uncertainty, although this assumption appears to have only been applied to OS previously (NICE TA655[32] and TA428[38]). For this scenario analysis, a waning in pembrolizumab PFS treatment effect was applied at Year 3 until no difference in hazards was assumed by Year 5. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.4. Scenario 4: Higher resource use in the PD health state

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

Based on clinical advice, the ERG considered it was implausible for patients to have identical costs in both the PFS and PD health state (Section 4.2.8.4). This scenario analysis applied higher resource use assumptions to the PD health state, which were derived from clinical opinion to the company reported within the CS (see Section 4.2.8.4). As such, weekly outpatient visits, blood count and biochemistry tests increased from 0.20 in the company's base case to 0.32, whilst weekly CT and PET scan usage increased from 0.06 and 0.03 respectively to 0.07. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.5. Scenario 5: No difference in SCT rates between treatment arms

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.8.4, SCT is not a central model driver as pembrolizumab is not being used as a bridge to transplant, though it may be associated with meaningful shifts to the ICER. However, SCT rates are associated with considerable uncertainty given that they are based on small patient numbers. This scenario analysis assumed no difference in SCT rates between

treatment arms i.e. the pembrolizumab allo-SCT and auto SCT rates are applied to both arms). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.6. Scenario 6: Dose intensity for pembrolizumab assumed to be 100%

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

As highlighted in Section 4.2.8.1, the company estimated the base case dose intensity for pembrolizumab to be 98%. However, the ERG understood that the dose intensity in practice could potentially be higher than the 98% witnessed in KEYNOTE-204,[3,4] and therefore were interested in determining whether assuming a 100% dosing intensity is likely to impact on the ICER. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.7. Scenario 7: Pembrolizumab administered 400 mg (every six weeks)

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

The company's base case used the licensed dose of 200 mg (every three weeks) which is appropriate (Section 4.2.8.1). However, given the availability of an alternative dose (400 mg administered every six weeks), the ERG conducted an analysis to determine the impact of using this alternative dosing option on the ICER. Given that treatment acquisition costs are a key driver of costs, this scenario is unlikely to have a material impact on the ICER. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.8. Scenario 8: Time horizon increased to 50 years

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

The model predicted that a small proportion of patients remained alive at 40 years (Section 4.2.5). For completeness the ERG considered the model should be run until all patients have died. This scenario is unlikely to have a material impact on results given the small proportion of patients alive at 40 years. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.9. Scenario 9: KEYNOTE-087 as the source for OS data in both treatment arms

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

Given the availability of OS data within KEYNOTE-087[5] (Section 4.2.6.1), the ERG was interested in using available data from this single arm study of pembrolizumab in order to generate OS for both treatment arms (SCT-2L: pembrolizumab and IGEV; SCT-3L+ and SCT+3L+: pembrolizumab and BV). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.10. Scenario 10: ToT for pembrolizumab based on KM data only

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.8.2, there was some uncertainty regarding the company's base case approach to modelling treatment costs. In order to reduce extrapolation uncertainty, the ERG considered that estimating costs using relevant KM data only from KEYNOTE-204[3,4] would accurately reflect trial-based treatment costs. In this scenario (for the SCT-2L subgroup), ToT for pembrolizumab was estimated based on KM data from the ITT population in KEYNOTE 204. Given that KM data were not available for IGEV, ToT was set to equal PFS (26 weeks) for the comparator. For the SCT-3L+ and SCT+3L+ subgroups KM data were available for both pembrolizumab and BV and these were subsequently used to estimate treatment costs. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.11. Scenario 11: Alternative cut-points for modelling ToT (26 weeks)

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

The company estimated treatment costs in the base case by extrapolating ToT at 80 weeks (Section 4.2.8.2). The ERG understood the company's rationale of using as much KM data as possible before extrapolation; however, the company did not provide sensitivity analyses exploring the use of alternative cut points. The ERG considered the use of a 26-week cut point as the most appropriate time for modelling ToT, given that ToT should be largely coterminous with PFS. ITT data were used to undertake this analysis (as opposed to by subgroup) as this was what was provided by the company during clarification. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.12. Scenario 12: Alternate parametric fit (log normal) for ToT (pembrolizumab and comparator)

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

The company did not provide scenario analyses using alternative distributions for ToT (Section 4.2.8.2). Although the exponential distribution selected by the company exhibited the lowest AIC/BIC score, there was minimal difference between the scores for each distribution. This scenario analyses used the log-normal distribution as it resulted in the second lowest AIC/BIC scores in both treatment arms (SCT-2L: pembrolizumab and IGEV; SCT-3L+ and SCT+3L+: pembrolizumab and BV). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.13. Scenario 13: Subsequent treatments based on subgroup data from KEYNOTE-204

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	✓

In the base case analysis, subsequent treatment costs for each subgroup were based on the company's understanding of current UK clinical practice. However, the ERG noted several concerns surrounding the base case assumptions (Section 4.2.8.3). It is worth highlighting that subsequent treatment data for the SCT-2L and SCT-3L+ subgroups were also available from KEYNOTE-204[3,4] detailing the list of treatments and the associated uptake rates from the study. This scenario therefore used direct subgroup trial data to estimate subsequent treatment costs for these subgroups. Subsequent treatment data were not available for the SCT+3L+ subgroup, therefore treatments and uptake rates, for patients who failed pembrolizumab and BV, were derived from the ITT population (Table 13 in Section 4.2.8.3). Although this scenario is useful, the ERG outlined concerns surrounding the use of these data to estimate subsequent treatment costs (see Section 4.2.8.3). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.14. Scenario 14: Balzarotti et al. (2016) used to estimate OS (pembrolizumab and comparator)

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	×

As noted in Section 4.2.6.1, the ERG did not consider Gopal et al. (2015)[1] to be the most appropriate data source to derive OS estimates for the SCT-2L and SCT-3L+ subgroups.

Patients in Balzarotti et al. (2016)[2] (patients with HL who are R/R to firstline chemotherapy), appeared to better reflect these subgroups. This scenario analysis is unlikely to have a material impact on the ICER given that the same OS data are applied to both arms (SCT-2L: pembrolizumab and IGEV; SCT-3L+ and SCT+3L+: pembrolizumab and BV). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.15. Scenario 15: Balzarotti et al. (2016) for OS and alternative parametric fit applied

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	✓	SCT+3L+	×

In addition to Section 6.2.1.14, the ERG considered there was some uncertainty surrounding the impact of using an alternative parametric fit on the ICER (as the company did not provide sensitivity analysis using alternative parametric fits). This scenario analysis aims to explore OS uncertainty by using an alternative data source considered more generalisable to the SCT-2L and SCT-3L+ subgroup (Balzarotti et al., 2016[2]), as well as an alternative parametric fit (log-logistic). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.16. Scenario 16: Alternative parametric fit (Weibull) for PFS, applied to both pembrolizumab and IGEV treatment arms

Applicable to subgroup:	SCT-2L	✓	SCT-3L+	×	SCT+3L+	×

To extrapolate PFS in its base case, the company applied a log normal curve to both treatment arms (Section 4.2.6.2). Given that the company did not provide sensitivity analysis results using alternative fits, this scenario estimates the impact of using the next best curve fit on the ICER. The Weibull produced the lowest AIC/BIC scores and therefore was selected as the appropriate fit for this scenario. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.17. Scenario 17: Combined analysis: PFS (fully parametric) and OS (KEYNOTE-087)

Applicable to subgroup:	**SCT-2L**	✓ Weibull	**SCT-3L+**	✓ Gen Gam	**SCT+3L+**	✓ Gen Gam

The ERG considered that it may be useful to conduct a combined scenario analysis to explore the combined effect of alternative PFS and OS assumptions on the ICER. This scenario analysis models OS using an alternative data source (KEYNOTE 087[5]) and uses an alternative fully parametric fit (Weibull was used for SCT-2L and generalised gamma for both SCT-3L+ and SCT+3L+), for PFS for both pembrolizumab and comparator (SCT-2L: IGEV, and SCT-3L- and SCT+3L+ BV. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.18. Scenario 18: Subsequent treatments assumed to reflect UK practice

Applicable to subgroup:	**SCT-2L**	×	**SCT-3L+**	✓ BV > BEN	**SCT+3L+**	✓ Pembro > NIVO

Subsequent treatments included in the model are considered to have a large impact on the base case ICER, given the associated treatment acquisition costs.

As noted in Section 4.2.8.3, there were a number of concerns surrounding the company's base assumptions with respect to subsequent treatments in both the SCT-3L+ and SCT+3L+ subgroups.

SCT-3L+ subgroup: In the base case analysis the company assumed that patients who fail on BV go on to receive pembrolizumab (Section 4.2.8.3). However, the ERG noted that as pembrolizumab is within the Cancer Drugs Fund (CDF), it may therefore may not be routinely available. Based on clinician response to the ERG, bendamustine was suggested a plausible treatment option for these patients. Therefore, this scenario assumes that 100% of patients who fail on BV go on to receive bendamustine. It is anticipated that this scenario analysis will have a large upward impact on the ICER, as subsequent treatment costs for the comparator arm have decreased, relative to the base case.

SCT+3L+ subgroup: In the base case analysis the company assumed that patients who fail on pembrolizumab go on to receive BV, whilst 100% of patients who failed BV went on to receive

nivolumab. Based on a review of the treatment pathway for this subgroup, patients in both treatment arms should receive nivolumab as subsequent treatment (Section 4.2.8.3).

For this scenario analysis, subsequent treatment assumptions were as outlined in Table 24, which more appropriately reflect UK clinical practice. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

Table 24: ERG preferred subsequent treatments

	Subsequent treatment
SCT-3L+	
Pembrolizumab	100% receive BV
BV	100% receive bendamustine only
SCT+3L+	
Pembrolizumab	100% receive nivolumab
BV	100% receive nivolumab

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group

6.2.1.19. Scenario 19: Reduction in maximum number of cycles of BV

Applicable to subgroup:	SCT-2L	×	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.8.2, the company assumed a maximum treatment duration of 35 cycles (105 weeks) for both pembrolizumab and BV in both the SCT-3L and SCT+3L subgroups, which did not appear appropriate. Although 35 cycles were consistent with the two-year pembrolizumab stopping rule, based on the SmPC for BV, the maximum number of cycles should be given is 16. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.20. Scenario 20: Fully parametric approach to model PFS (using the generalised gamma curve)

Applicable to subgroup:	SCT-2L	×	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.6.2, the generalised gamma provided a superior statistical fit to the full SCT-3L+ subgroup data compared with the other parametric distributions (as assessed by AIC and BIC statistics), as well as the best fit to full SCT+3L+ subgroup data for the pembrolizumab arm. Therefore, this scenario analysis models PFS by applying full-fitted generalised gamma

distributions to each arm (i.e. with a break-point at Week 0). The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.21. Scenario 21: Alternative parametric fit (log-logistic) for Gopal et al. (2015) OS data for both pembrolizumab and BV

Applicable to subgroup:	SCT-2L	×	SCT-3L+	✓	SCT+3L+	✓

As noted in Section 4.2.6.1 the company did not provide sensitivity analysis using alternative distributions. For this scenario the ERG selected the log-logistic curve for use on the basis that it produces the next best fit, based on AIC/BIC statistics. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.1.22. Scenario 22: Model PFS using different data cut point (26 weeks)

Applicable to subgroup:	SCT-2L	×	SCT-3L+	✓	SCT+3L+	✓

In the base case analysis, the company extrapolated PFS using a 52-week cut point (see Section 4.2.6.2). For this scenario analysis, the log-normal distribution was fitted to the data at Week 26. Using a 26-week break-point means that more robust trial data is used to inform the parametric extrapolation, leading to less uncertain estimates. This is the ERG's preferred cut point. The incremental results and impact on the company base case are presented in Section 6.2.2 by subgroup.

6.2.2. Impact of scenario analyses on the ICER

The impact of each scenario on the ICER is provided for each of the subgroups: SCT-2L (Table 25), SCT-3L+ (Table 26), and SCT+3L+ (Table 27).

Table 25: Impact on the ICER of additional analyses undertaken by the ERG: SCT-2L

Subgroup	**Subgroup: SCT-2L**			
Scenario	Incr costs £	Incr QALYs	ICER £/QALY	+/-
ERG corrected company base case	[redacted]	[redacted]	53,099	-
Scenario 1: Utility value PD health state	[redacted]	[redacted]	94,284	[redacted]
Scenario 2: Equal PFS and PD utility value	[redacted]	[redacted]	799,995	[redacted]
Scenario 3: Waning of pembro PFS Tx effect	[redacted]	[redacted]	58,559	[redacted]
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	50,627	[redacted]
Scenario 5: No difference in SCT rates	[redacted]	[redacted]	64,332	[redacted]

Subgroup	Subgroup: SCT-2L			
Scenario 6: Dose intensity for pembro 100%	[redacted]	[redacted]	54,513	[redacted]
Scenario 7: Pembro 400 mg Q6W	[redacted]	[redacted]	53,742	[redacted]
Scenario 8: Time horizon 50 years	[redacted]	[redacted]	52,891	[redacted]
Scenario 9: KN-087 OS data (**pembro and IGEV**)	[redacted]	[redacted]	20,205	[redacted]
Scenario 10: ToT pembro based on KM data only	[redacted]	[redacted]	52,172	[redacted]
Scenario 11: 26-week cut-point for modelling ToT	[redacted]	[redacted]	103,052	[redacted]
Scenario 12: Log-normal fit for ToT (**pembro and IGEV**)	[redacted]	[redacted]	53,319	[redacted]
Scenario 13: Subsequent Tx based on SG data KN-204	[redacted]	[redacted]	57,148	[redacted]
Scenario 14: Balzarotti (2016) for OS (**pembro and IGEV**)	[redacted]	[redacted]	41,007	[redacted]
Scenario 15: Balzarotti (2016) for OS + log-logistic	[redacted]	[redacted]	44,996	[redacted]
Scenario 16: Weibull for PFS (**pembro & IGEV**)	[redacted]	[redacted]	53,745	[redacted]
Scenario 17: Combined PFS (**Weibull**) and OS (**KEYNOTE-087**) pembro and IGEV	[redacted]	[redacted]	20,799	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; incr, incremental; KN, KEYNOTE; OS, overall survival; PD, progressed disease; pembro, pembrolizumab; PFS, progression free survival; Q6W, every 6 weeks; QALYs, quality adjusted life years; SCT, stem cell transplant; SG, subgroup; ToT, time on treatment; Tx, treatment

Table 26: Impact on the ICER of additional analyses undertaken by the ERG: SCT-3L+

Subgroup	Subgroup: SCT-3L+			
Scenario	Incr costs £	Incr QALYs	ICER £/QALY	+/-
ERG corrected company base case	[redacted]	[redacted]	Dominant (-33,316)	-
Scenario 1: Utility value PD health state	[redacted]	[redacted]	Dominant (-52,833)	[redacted]
Scenario 2: Equal PFS and PD utility value	[redacted]	[redacted]	Dominant (-168,907)	[redacted]
Scenario 3: Waning of pembro PFS Tx effect	[redacted]	[redacted]	Dominant (-34,253)	[redacted]
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	Dominant (-40,840)	[redacted]
Scenario 5: No difference in SCT rates	[redacted]	[redacted]	Dominant (-36,184)	[redacted]
Scenario 6: Dose intensity for pembro 100%	[redacted]	[redacted]	Dominant (-32,154)	[redacted]
Scenario 7: Pembro 400 mg Q6W	[redacted]	[redacted]	Dominant (-33,314)	[redacted]
Scenario 8: Time horizon 50 years	[redacted]	[redacted]	Dominant (-33,234)	[redacted]
Scenario 9: KN-087 OS data (**pembro and BV**)	[redacted]	[redacted]	Dominant (-10,962)	[redacted]

Subgroup	Subgroup: SCT-3L+			
Scenario 10: ToT pembro based on KM data only	[redacted]	[redacted]	Dominant (-31,229)	[redacted]
Scenario 11: 26-week cut-point for modelling ToT	[redacted]	[redacted]	Dominant (-52,121)	[redacted]
Scenario 12: Log-normal fit for ToT (**pembro and BV**)	[redacted]	[redacted]	Dominant (-33,183)	[redacted]
Scenario 13: Subsequent Tx based on SG data KN-204	[redacted]	[redacted]	Dominant (-43,136)	[redacted]
Scenario 14: Balzarotti (2016) for OS (**pembro and BV**)	[redacted]	[redacted]	Dominant (-24,450)	[redacted]
Scenario 15: Balzarotti (2016) for OS + log-logistic	[redacted]	[redacted]	Dominant (-26,254)	[redacted]
Scenario 16: Weibull for PFS (**pembro and BV**)	[redacted]	[redacted]	Dominant (-33,316)	[redacted]
Scenario 17: Combined PFS (**generalised gamma**) and OS (**KEYNOTE-087**) pembro and BV	[redacted]	[redacted]	Dominant (-12,622)	[redacted]
Scenario 18: Subsequent treatments assumed to reflect UK practice (**100% bendamustine** on BV failure)	[redacted]	[redacted]	15,703	[redacted]
Scenario 19: Reduction in maximum number of BV cycles	[redacted]	[redacted]	Dominant (-17,935)	[redacted]
Scenario 20: Fully parametric approach to model PFS (generalised gamma curve)	[redacted]	[redacted]	Dominant (-35,005)	[redacted]
Scenario 21: Alternative parametric fit (**log logistic**) for Gopal et al. (2015) OS data (pembro and BV)	[redacted]	[redacted]	Dominant (-33,110)	[redacted]
Scenario 22: 26-week data cut point for PFS	[redacted]	[redacted]	Dominant (-36,396)	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; incr, incremental; KN, KEYNOTE; OS, overall survival; PD, progressed disease; pembro, pembrolizumab; PFS, progression free survival; Q6W, every 6 weeks; QALYs, quality adjusted life years; SCT, stem cell transplant; SG, subgroup; ToT, time on treatment; Tx, treatment

Table 27: Impact on the ICER of additional analyses undertaken by the ERG: SCT+3L+

Subgroup	Subgroup: SCT+3L+			
Scenario	Incr costs £	Incr QALYs	ICER £/QALY	+/-
ERG corrected company base case	[redacted]	[redacted]	Dominant (-73,896)	-
Scenario 1: Utility value PD health state	[redacted]	[redacted]	Dominant (-107,883)	[redacted]
Scenario 2: Equal PFS and PD utility value	[redacted]	[redacted]	Dominant (-458,591)	[redacted]
Scenario 3: Waning of pembro PFS Tx effect	[redacted]	[redacted]	Dominant (-75,473)	[redacted]
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	Dominant (-79,965)	[redacted]

Subgroup	Subgroup: SCT+3L+			
Scenario 5: No difference in SCT rates	[redacted]	[redacted]	Dominant (-78,183)	[redacted]
Scenario 6: Dose intensity for pembro 100%	[redacted]	[redacted]	Dominant (-72,152)	[redacted]
Scenario 7: Pembro 400 mg Q6W	[redacted]	[redacted]	Dominant (-74,342)	[redacted]
Scenario 8: Time horizon 50 years	[redacted]	[redacted]	Dominant (-73,726)	[redacted]
Scenario 9: KN-087 OS data (**pembro and BV**)	[redacted]	[redacted]	Dominant (-29,418)	[redacted]
Scenario 10: ToT pembro based on KM data only	[redacted]	[redacted]	Dominant (-73,374)	[redacted]
Scenario 11: 26-week cut-point for modelling ToT	[redacted]	[redacted]	Dominant (-139,123)	[redacted]
Scenario 12: Log-normal fit for ToT (**pembro and BV**)	[redacted]	[redacted]	Dominant (-73,967)	[redacted]
Scenario 13: Subsequent Tx based on SG data KN-204	[redacted]	[redacted]	Dominant (-28,585)	[redacted]
Scenario 17: Combined PFS (**generalised gamma**) and OS (**KEYNOTE-087**) pembro and BV	[redacted]	[redacted]	Dominant (-30,704)	[redacted]
Scenario 18: Subsequent treatments assumed to reflect UK practice (**100% nivolumab** on pembro failure)	[redacted]	[redacted]	Dominant (-45,625)	[redacted]
Scenario 19: Reduction in maximum number of BV cycles	[redacted]	[redacted]	Dominant (-65,013)	[redacted]
Scenario 20: Fully parametric approach to model PFS (generalised gamma curve)	[redacted]	[redacted]	Dominant (-54,360)	[redacted]
Scenario 21: Alternative parametric fit (**log-logistic**) for Gopal et al. (2015) OS data (pembro and BV)	[redacted]	[redacted]	Dominant (-74,240)	[redacted]
Scenario 22: 26-week data cut point for PFS	[redacted]	[redacted]	Dominant (-57,940)	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; incr, incremental; KN, KEYNOTE; OS, overall survival; PD, progressed disease; pembro, pembrolizumab; PFS, progression free survival; Q6W, every 6 weeks; QALYs, quality adjusted life years; SCT, stem cell transplant; SG, subgroup; ToT, time on treatment; Tx, treatment

6.2.3. Adjustment to the probabilistic sensitivity analysis

In the company's analysis, a single set of distribution parameters informs the OS curves in both treatment arms and, as a result, these curves are varied in exactly the same way in the probabilistic sensitivity analysis (PSA).

The ERG noted that this may not adequately reflect uncertainty surrounding the OS parameters: this uncertainty would be captured better by using two sets of OS parameters, one for each arm.

These sets contained identical values in the deterministic analysis, but were varied separately in the ERG probabilistic analysis. The same Cholesky matrix was used for each set to account for the correlation among the parameters in that set, but the matrix was multiplied by a different random vector for each set, the values of which were drawn from an inverse Normal distribution.

The choice of OS distribution based on the data from Balzarotti et al. (2016)[2] needed to be specified for the PSA in the ERG probabilistic analysis for the SCT-2L and SCT-3L+ subgroups.

The PSA sample mean for the PFS hazard ratio (HR) obtained from the MAIC was also missing from the company's model. In the ERG and corrected company probabilistic analysis for the SCT-2L subgroup, the HR was varied using a log-normal distribution, with a standard error based on the 95% confidence interval obtained from the MAIC.

6.3. ERG's preferred assumptions

The ERG's preferred base case analysis for each subgroup comprised alternative assumptions and amended model errors and settings

Table 28: ERG's preferred model assumptions (SCT-2L)

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case	[redacted]	[redacted]	£53,581
ERG corrected company base case	[redacted]	[redacted]	£53,099
Scenario 14: Balzarotti et al (2016) used as the data source for estimating OS for both pembrolizumab and chemotherapy (IGEV)	[redacted]	[redacted]	£41,007
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms	[redacted]	[redacted]	£94,319
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	£89,930
Scenario 5: No difference in SCT rates between treatment arms (apply pembrolizumab allo-SCT and auto SCT rate to both arms)	[redacted]	[redacted]	£109,876
Scenario 6: Dose intensity for pembrolizumab assumed to be 100%	[redacted]	[redacted]	£112,387
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	£112,284
Scenario 11: 26-week data cut point for ToT	[redacted]	[redacted]	£202,428

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Table 29: Comparison of company and ERG results (SCT-2L)

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
ERG corrected company base case (deterministic)							
Pembro	[redacted]	[redacted]	[redacted]	-	-	-	-
IGEV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£53,099
ERG base case (deterministic)							
Pembro	[redacted]	[redacted]	[redacted]	-	-	-	-
IGEV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£202,428
ERG corrected company base case (probabilistic)							
Pembro	[redacted]	[redacted]	[redacted]	-	-	-	
IGEV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£56,446
ERG base case (probabilistic)							
Pembro	[redacted]	[redacted]	[redacted]	-	-	-	-
IGEV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£176,859

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life year; Pembro, pembrolizumab; QALY, quality adjusted life year

Note: It was not possible to obtain LY results from the cost-effectiveness model

Table 30: ERG's preferred model assumptions (SCT-3L+)

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case[a]	[redacted]	[redacted]	**Dominant (-£33,316)**
Scenario 14: Balzarotti et al (2016) used as the data source for estimating OS for both pembrolizumab and chemotherapy	[redacted]	[redacted]	Dominant (-£24,450)
Scenario 22: Semi parametric approach to modelling PFS (cut point for PFS set at 26 weeks)	[redacted]	[redacted]	Dominant (-£27,163)
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms	[redacted]	[redacted]	Dominant (-£61,670)
Scenario 18: Subsequent treatment assumed to reflect UK practice: 100% of patients who fail pembrolizumab go on to receive BV AND 100% of patients who fail on BV go on to receive bendamustine alone	[redacted]	[redacted]	£24,265
Scenario 19: Maximum ToT for brentuximab set to 16 cycles (not 35 as per base case)	[redacted]	[redacted]	£52,006
Scenario 11: Cut-off for ToT to reflect PFS data cut point (26 weeks)	[redacted]	[redacted]	£79,232

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	£67,399
Scenario 5: No difference in SCT rates between treatment arms (pembrolizumab allo-SCT and auto-SCT rate to both arms)	[redacted]	[redacted]	£62,226
Scenario 6: Dose intensity for pembrolizumab 100%	[redacted]	[redacted]	£65,018
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	£64,124

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Note:

a ERG corrected company base case not applicable for this subgroup (see Section 5.3)

Table 31: Comparison of company and ERG results (SCT-3L+)

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company base case (deterministic)[a]							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£33,316)
ERG base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£64,124
Company base case (probabilistic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£31,773)
ERG base case (probabilistic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£58,738

Abbreviations: BV, brentuximab vedotin; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year

Note:

It was not possible to obtain LY results from the cost-effectiveness model

a ERG corrected company base case not applicable for this subgroup (see Section 5.3)

Table 32: ERG's preferred model assumptions (SCT+3L+)

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Company base-case	[redacted]	[redacted]	**Dominant (-£73,896)**
Scenario 22: Semi parametric approach to modelling PFS (cut point for PFS set at 26 weeks)	[redacted]	[redacted]	Dominant (-£57,940)
Scenario 1: Utility value for PD health state set to [redacted] for both treatment arms	[redacted]	[redacted]	Dominant (-£79,339)
Scenario 19: Maximum ToT for brentuximab set to 16 cycles (not 35 as per base case)	[redacted]	[redacted]	Dominant (-£68,202)
Scenario 11: Cut-off for ToT to reflect PFS data cut point (26 weeks)	[redacted]	[redacted]	Dominant (-£49,001)
Scenario 4: Higher resource use in the PD health state	[redacted]	[redacted]	Dominant (-£61,514)

Preferred assumption	Incr. Costs	Incr. QALYs	Cumulative ICER £/QALY
Scenario 5: No difference in SCT rates between treatment arms (pembrolizumab allo-SCT and auto-SCT rate to both arms)	[redacted]	[redacted]	Dominant (-£66,889)
Scenario 6: Dose intensity for pembrolizumab 100%	[redacted]	[redacted]	Dominant (-£64,127)
Scenario 8: Time horizon increased to 50 years	[redacted]	[redacted]	Dominant (-£63,904)
Scenario 18: Subsequent treatment assumed to reflect UK practice: 100% of patients who fail pembrolizumab go on to receive nivolumab AND 100% of patients who fail on BV go on to receive nivolumab	[redacted]	[redacted]	Dominant (-£33,849)

Abbreviations: BV, brentuximab vedotin; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PFS, progression free survival; QALY, quality adjusted life year; SCT, stem cell transplant

Note: a ERG corrected company base case not applicable for this subgroup (see Section 5.3)

Table 33: Comparison of company and ERG results (SCT+3L+)

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company base case (deterministic)[a]							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£73,896)
ERG base case (deterministic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£33,849)
Company base case (probabilistic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£66,098)
ERG base case (probabilistic)							
Pembrolizumab	[redacted]	[redacted]	[redacted]	-	-	-	-
BV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Dominant (-£34,156)

Abbreviations: BV, brentuximab vedotin; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year

Note:

It was not possible to obtain LY results from the cost-effectiveness model

a ERG corrected company base case not applicable for this subgroup (see Section 5.3)

6.4. Conclusions of the cost-effectiveness section

6.4.1. SCT-2L

The company's base case results (ERG corrected for errors) indicated that pembrolizumab resulted in an ICER of £53,099 when compared to salvage chemotherapy (IGEV). The ERG acknowledged that this result was subject to uncertainty due to concerns surrounding the use of MAIC data in the economic analysis, which was used to estimate the clinical effectiveness of pembrolizumab. As such results should be interpreted with caution.

Using the ERG's preferred assumptions, the ICER for pembrolizumab increased to £202,428 based on an incremental cost of [redacted] and an incremental QALY gain of [redacted]. Based on this analysis, pembrolizumab does not appear to be a cost-effective treatment option for patients with R/RcHL who did not have at least two prior therapies when autologous stem cell transplant is not a treatment option, when compared to salvage chemotherapy. The ERG conducted a large number of scenario analyses to test uncertainty surrounding key modelled parameters. As outlined in Table 28, the ICER was particularly sensitive to several ERG preferred assumptions including the use of alternative utility and ToT assumptions.

6.4.2. SCT-3L+ and SCT+3L+

For the SCT-3L+, the company's base case results (ERG corrected for errors) indicated that pembrolizumab was the dominant treatment when compared to BV. Using the ERG's preferred assumptions, pembrolizumab resulted in an ICER of £64,124 based on an incremental cost of [redacted] and an incremental QALY gain of [redacted]. The ERG conducted scenario analyses to test uncertainty surrounding key modelled parameters. The ICER was particularly sensitive to several preferred ERG assumptions including alternative subsequent treatments, reduced maximum ToT for BV and the use of an earlier cut point for ToT (26 weeks).

For the SCT+3L+ subgroup, the company's base case results (ERG corrected for errors) indicated that pembrolizumab remained dominant when compared to BV. Using the ERG's preferred assumptions, pembrolizumab remained dominant resulting in incremental savings of [redacted] and an incremental QALY gain of [redacted]. Incremental results were most sensitive to alternative ERG preferred assumptions including utility and subsequent treatments as noted in section 6.3.

7. END OF LIFE

Pembrolizumab does not meet NICE's end of life criteria outlined below.

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months; and
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional three months, compared to current NHS treatment.

References

1. Gopal AK, Chen R, Smith SE, Ansell SM, Rosenblatt JD, Savage KJ, et al. Durable remissions in a pivotal phase 2 study of brentuximab vedotin in relapsed or refractory Hodgkin lymphoma. Blood. 2015;125(8):1236-43.
2. Balzarotti M, Brusamolino E, Angelucci E, Carella AM, Vitolo U, Russo E, et al. B-IGEV (bortezomib plus IGEV) versus IGEV before high-dose chemotherapy followed by autologous stem cell transplantation in relapsed or refractory Hodgkin lymphoma: a randomized, phase II trial of the Fondazione Italiana Linfomi (FIL). Leukemia & Lymphoma. 2016;57(10):2375-81.
3. Merck Sharp Dohme. KEYNOTE-204 CSR. Data on file. 2020.
4. Kuruvilla J, Ramchandren R, Santoro A, Paszkiewicz-Kozik E, Gasiorowski R, Johnson N, et al. KEYNOTE-204: Randomized, open-label, phase III study of pembrolizumab (pembro) versus brentuximab vedotin (BV) in relapsed or refractory classic Hodgkin lymphoma (R/R cHL). Journal of Clinical Oncology. 2020;38(15 Suppl):8005.
5. Chen R, Zinzani PL, Fanale MA, Armand P, Johnson NA, Brice P, et al. Phase II study of the efficacy and safety of pembrolizumab for relapsed/refractory classic Hodgkin lymphoma. Journal of Clinical Oncology. 2017;35(19):2125-32.
6. Cancer Research UK. Hodgkin lymphoma, 2020. Available from: https://www.cancerresearchuk.org/about-cancer/hodgkin-lymphoma/types.
7. Martin P, Leonard JP. Hodgkin lymphoma: Merck Manuals Professional Version; 2020. Available from: https://www.merckmanuals.com/professional/hematology-and-oncology/lymphomas/hodgkin-lymphoma.
8. Cancer Research UK. Hodgkin lymphoma survival statistics, 2020. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma/survival#heading-Zero
9. Engelhardt BG, Holland DW, Brandt SJ, Chinratanalab W, Goodman SA, Greer JP, et al. High-dose chemotherapy followed by autologous stem cell transplantation for relapsed or refractory Hodgkin lymphoma: prognostic features and outcomes. Leukemia & Lymphoma. 2007;48(9):1728-35.
10. Moskowitz AJ, Perales M-A, Kewalramani T, Yahalom J, Castro-Malaspina H, Zhang Z, et al. Outcomes for patients who fail high dose chemoradiotherapy and autologous stem cell rescue for relapsed and primary refractory Hodgkin lymphoma. British Journal of Haematology. 2009;146(2):158-63.
11. Cancer Research UK. Hodgkin lymphoma incidence statistics, 2020. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma/incidence#heading-One.
12. National Institute for Health and Care Excellence (NICE). Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma [TA524], 2018. Available from: https://www.nice.org.uk/guidance/ta524.
13. National Institute for Health and Care Excellence (NICE). Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma [TA462], 2017. Available from: https://www.nice.org.uk/guidance/ta462.
14. National Institute for Health and Care Excellence (NICE). Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [TA540], 2018. Available from: https://www.nice.org.uk/guidance/ta540.
15. National Institute for Health and Care Excellence (NICE). Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after autologous stem cell transplant, or at least one prior therapy. Final scope, 2020. Available from: https://www.nice.org.uk/guidance/gid-ta10485/documents/final-scope-2.

16. Armand P, Shipp MA, Ribrag V, Michot JM, Zinzani PL, Kuruvilla J, et al. Pembrolizumab in patients with classical Hodgkin lymphoma after brentuximab vedotin failure: long-term efficacy from the phase 1b KEYNOTE-013 study. Blood Conference: 58th Annual Meeting of the American Society of Hematology. 2016;128(22):1108.
17. Geoerger B, Kang HJ, Yalon-Oren M, Marshall LV, Vezina C, Pappo A, et al. Pembrolizumab in paediatric patients with advanced melanoma or a PD-L1-positive, advanced, relapsed, or refractory solid tumour or lymphoma (KEYNOTE-051): interim analysis of an open-label, single-arm, phase 1-2 trial. Lancet Oncology. 2020;21(1):121-33.
18. Wells G, Shea B, O'connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Ottawa Hospital Research Institute; 2011. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
19. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. Journal of Clinical Oncology. 2007;25(5):579-86.
20. Miettinen O, Nurminen M. Comparative analysis of two rates. Statistics in Medicine. 1985;4(2):213-26.
21. EORTC. EORTC-QLQ-C30 questionnaire (version 3.0) 1995. Available from: https://www.eortc.org/app/uploads/sites/2/2018/08/Specimen-QLQ-C30-English.pdf.
22. EuroQol. EQ-5D-3L 2020. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/.
23. Eyre TA, Phillips EH, Linton KM, Arumainathan A, Kassam S, Gibb A, et al. Results of a multicentre UK-wide retrospective study evaluating the efficacy of brentuximab vedotin in relapsed, refractory classical Hodgkin lymphoma in the transplant naive setting. British Journal of Haematology. 2017;179(3):471-9.
24. Precision HEOR. Comparative efficacy of pembrolizumab and competing interventions for relapsed or refractory cHL: Feasibility assessment – UK context. 2020.
25. Parker C, Woods B, Eaton J, Ma E, Selby R, Benson E, et al. Brentuximab vedotin in relapsed/refractory Hodgkin lymphoma post-autologous stem cell transplant: a cost-effectiveness analysis in Scotland. Journal of Medical Economics. 2017;20(1):8-18.
26. Dolan P. Modeling valuations for EuroQol health states. Medical Care. 1997;35(11):1095-108.
27. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology. 2012;12(1):9.
28. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data, 2011: Decision Support Unit, ScHARR, University of Sheffield. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.
29. Hansen BE. The new econometrics of structural change: dating breaks in U.S. labour productivity. Journal of Economic Perspectives. 2001;15(4):117-28.
30. Burnham KP, Anderson DR. Model selection and multimodel inference. A practical information-theoretic approach. New York: Springer-Verlag; 2002.
31. Jackson CH, Sharples LD, Thompson SG. Survival models in health economic evaluations: balancing fit and parsimony to improve prediction. International Journal of Biostatistics. 2010;6(1):34.
32. National Institute for Health and Care Excellence (NICE). Nivolumab for advanced squamous non-small-cell lung cancer after chemotherapy [TA655], 2020. Available from: https://www.nice.org.uk/guidance/ta655.
33. Santoro A, Magagnoli M, Spina M, Pinotti G, Siracusano L, Michieli M, et al. Ifosfamide, gemcitabine, and vinorelbine: a new induction regimen for refractory and relapsed Hodgkin's lymphoma. Haematologica. 2007;92(1):35-41.

34. Swinburn P, Shingler S, Acaster S, Lloyd A, Bonthapally V. Health utilities in relation to treatment response and adverse events in relapsed/refractory Hodgkin lymphoma and systemic anaplastic large cell lymphoma. Leukemia & Lymphoma. 2015;56(6):1839-45.
35. Armand P, Engert A, Younes A, Fanale M, Santoro A, Zinzani PL, et al. Nivolumab for relapsed/refractory classic hodgkin lymphoma after failure of autologous hematopoietic cell transplantation: extended follow-up of the multicohort single-arm phase II Checkmate 205 trial. Journal of Clinical Oncology. 2018;36(14):1428-39.
36. Scottish Medicines Consortium. Nivolumab 10mg/mL concentrate for solution for infusion (Opdivo). SMC No (1240/17), 2017. Available from: https://www.scottishmedicines.org.uk/media/2051/nivolumab_opdivo_chl_final_june_2017_for_website.pdf.
37. EMC. SMPC: KEYTRUDA 50 mg powder for concentrate for solution for infusion, 2020. Available from: https://www.medicines.org.uk/emc/product/6947/smpc.
38. National Institute for Health and Care Excellence (NICE). Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy [TA428], 2017. Available from: https://www.nice.org.uk/guidance/ta428.
39. National Institute for Health and Care Excellence (NICE). Brentuximab vedotin for treating CD30-positive Hodgkin's lymphoma: Committee Papers [TA446], 2016. Available from: https://webarchive.nationalarchives.gov.uk/20180501231126/https://www.nice.org.uk/guidance/ta446/documents/committee-papers.
40. NHS England. National Cost Collection: National Schedule of NHS costs - Year 2018-19 - NHS trust and NHS foundation trusts 2018. Available from: https://www.england.nhs.uk/wp-content/uploads/2020/08/2_-_National_schedule_of_NHS_costs_V2.xlsx.
41. Radford J, McKay P, Malladi R, Johnson R, Bloor A, Percival F, et al. Treatment pathways and resource use associated with recurrent Hodgkin lymphoma after autologous stem cell transplantation. Bone Marrow Transplantation. 2017;52(3):452-4.
42. Brown T, Pilkington G, Bagust A, Boland A, Oyee J, Smith CT, et al. Clinical effectiveness and cost-effectiveness of first-line chemotherapy for adult patients with locally advanced or metastatic non-small cell lung cancer: a systematic review and economic evaluation. Health Technology Assessment. 2013;17(31):1-278.