NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]

Contents:

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

Pre-meeting briefing

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

1

Abbreviations

AE	Adverse event	**ICER**	Incremental cost-effectiveness ratio
ASaT	All Subjects as Treated	**MAIC**	Matched Adjusted Indirect treatment comparison
AutoSCT	Autologous Stem Cell Transplant	**OS**	Overall survival
AlloSCT	Allogeneic Stem Cell Transplant	**PD**	Progressive Disease
BV	Brentuximab Vedotin	**PFS**	Progression free survival
BICR	Blinded Independent Central Radiologists	**PR**	Partial response
BSC	Best supportive care	**RRcHL**	Relapsed or refractory classical Hodgkin lymphoma
CAA	Commercial access agreement	**SD**	Stable Disease
cHL	Classical Hodgkin lymphoma	**SOC**	Standard of Care
CR	Complete response		

2

Key issues 1: Clinical management and effectiveness

- How long would pembrolizumab treatment be continued for in clinical practice?
  - SmPC: "Patients should be treated…until disease progression or unacceptable toxicity."
  - Time to discontinuation assumed to be 24 months after starting treatment in base case model for people who don't have alloSCT, inline with KEYNOTE-087 protocol
- Is TA462 (Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma) relevant for cohort 1 in this appraisal?
- How effective is pembrolizumab for treating people with relapsed or refractory classical Hodgkin lymphoma compared with current practice?
  - Data are based on a non-comparative, single arm trial
  - Follow-up of participants from the trial is ongoing and data is potentially immature

3

Key issues 2: Clinical effectiveness

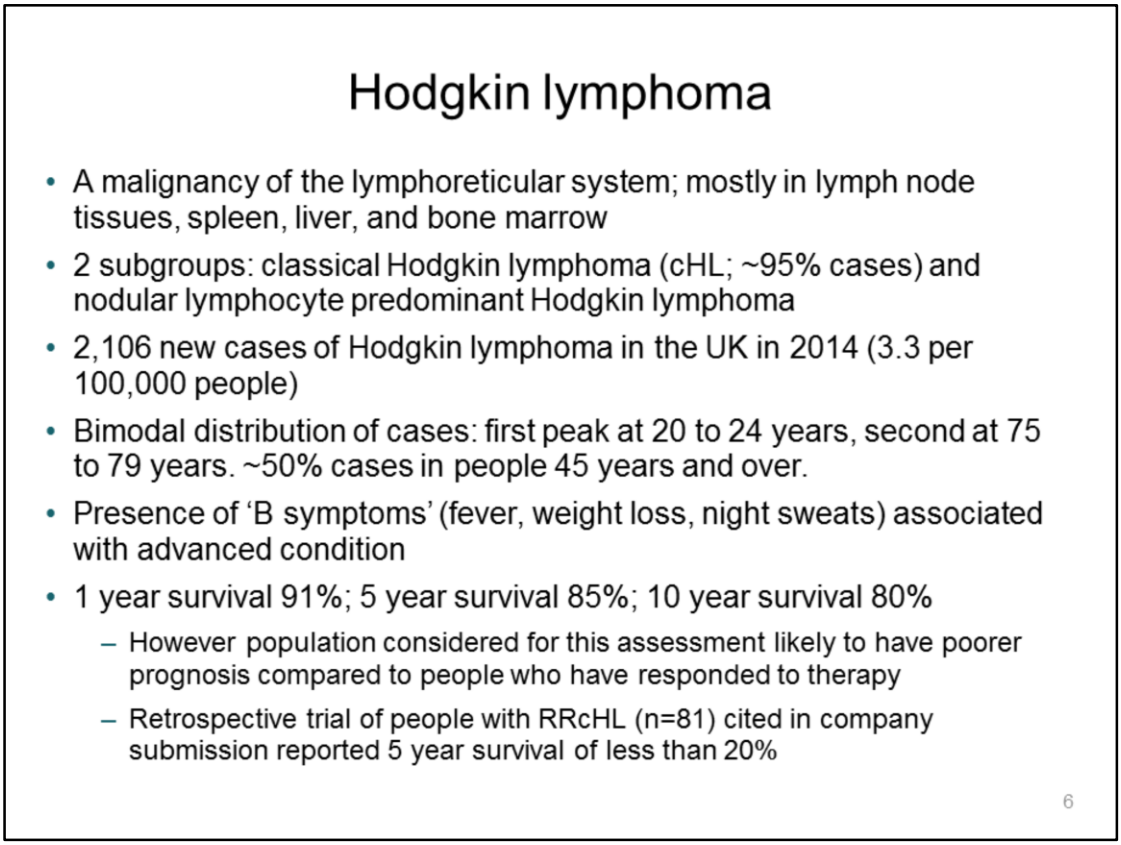
- How robust are the indirect comparisons of pembrolizumab and SOC?
  - Does the population in the comparator study (Cheah et al. 2016) adequately represent the UK clinical population?
- Is it more appropriate to use a naïve indirect comparison or MAIC to compare KEYNOTE-087 and Cheah et al. (2016) data?
- How well does the population in Cheah et al. (2016) match cohort 2 from KEYNOTE-087?

4

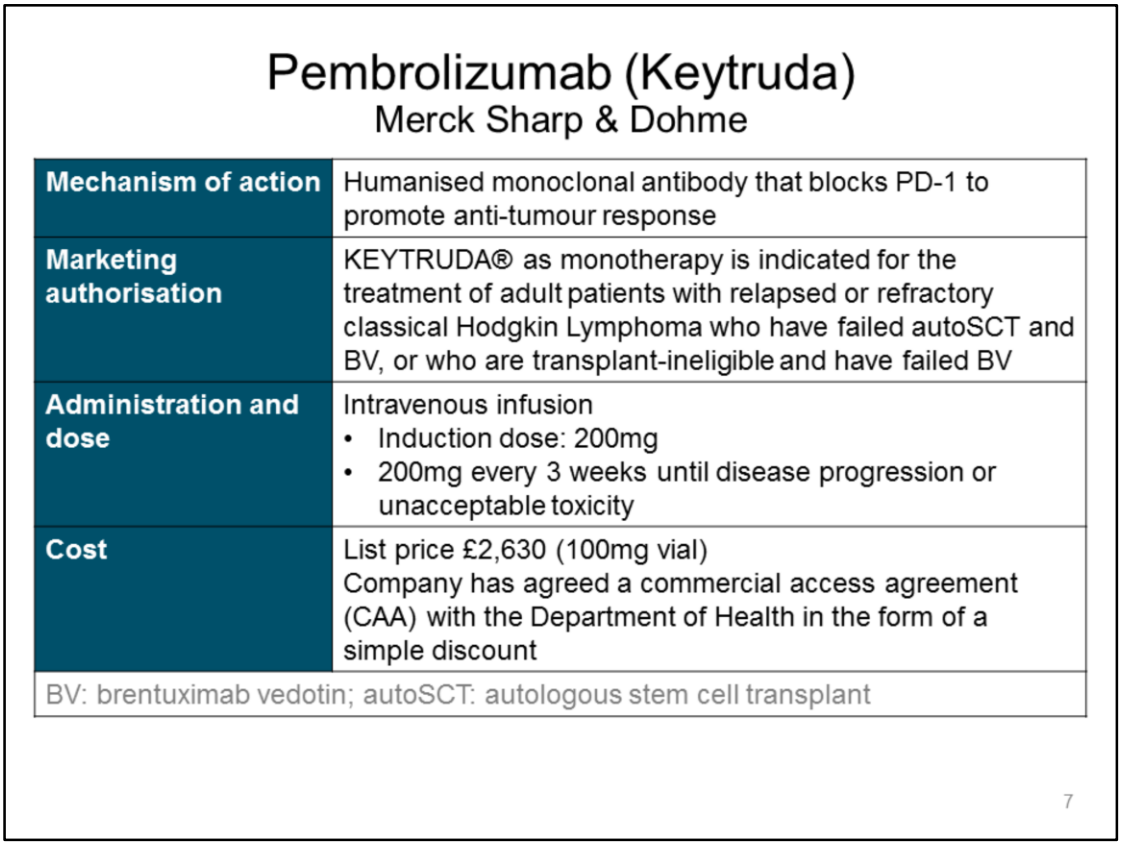
Key issues 3: Cost effectiveness

- Is the structural assumption that all alloSCTs would occur 12 weeks after starting treatment appropriate?
- Is it appropriate that best supportive care (BSC) was not considered as a comparator in the base-case analysis?
- Is the assumption that no patients with progressive disease would have alloSCT appropriate?
- Is the calculated utility for progressive disease more appropriate for use than the utility score for this state from KEYNOTE-087?
- Does pembrolizumab meet the criteria for a life-extending treatment at the end of life?
- Does pembrolizumab represent an innovative treatment?

5



Source: Company submission, section 3.1 (pages 35 and 36), section 3.2 (pages 36 to 37); ERG report, section 2.1 (page 20)



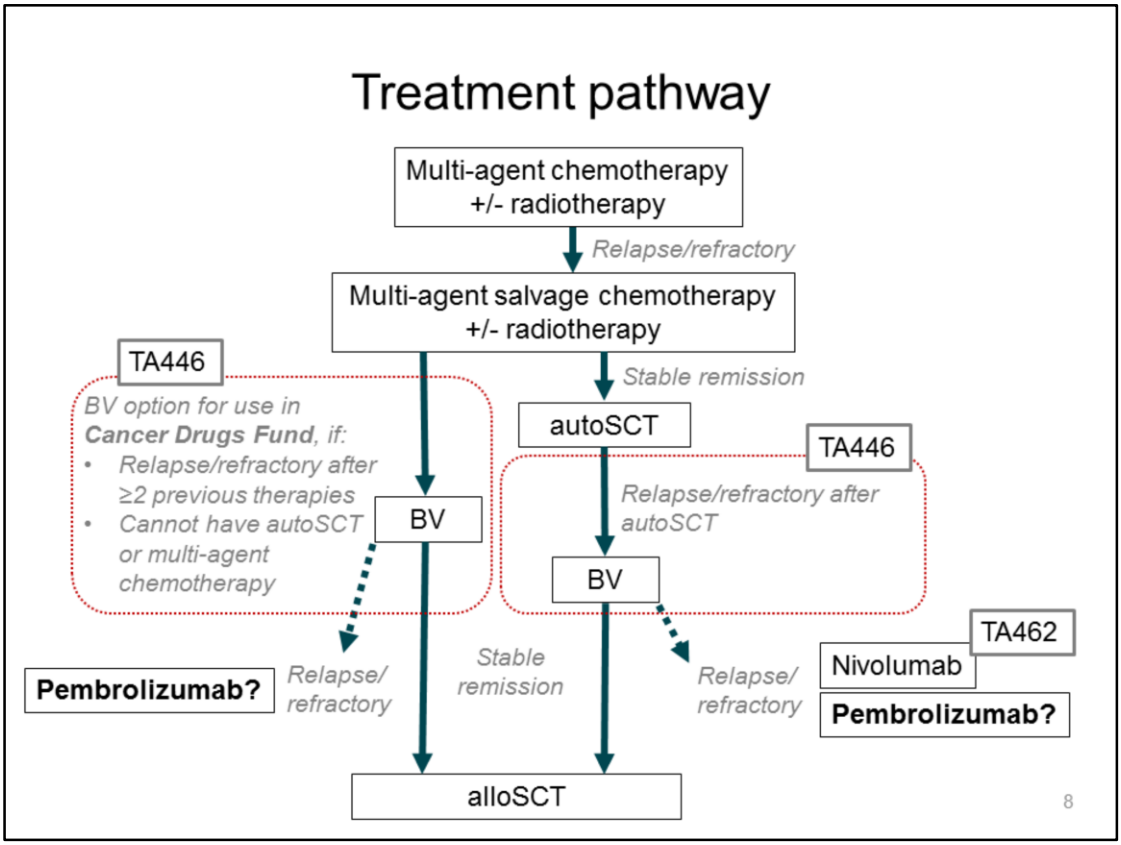
Pembrolizumab (Keytruda)
Merck Sharp & Dohme

Mechanism of action	Humanised monoclonal antibody that blocks PD-1 to promote anti-tumour response
Marketing authorisation	KEYTRUDA® as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and BV, or who are transplant-ineligible and have failed BV
Administration and dose	Intravenous infusion • Induction dose: 200mg • 200mg every 3 weeks until disease progression or unacceptable toxicity
Cost	List price £2,630 (100mg vial) Company has agreed a commercial access agreement (CAA) with the Department of Health in the form of a simple discount

BV: brentuximab vedotin; autoSCT: autologous stem cell transplant

7

Source: Company submission: section 2.1 (page 28), section 2.2. (page 29), section 2.3 (page 30)



Source: Adapted from Royal College of pathologist's submission; Company submission, section 3.3 (page 38), section 3.5 (page 40)

- No NICE clinical guidance on treatment of Hodgkin lymphoma
- British Committee for Standards (BCSH) in Haematology guidelines suggested as relevant to UK practice
- 1st line therapy can include:
    - ABVD chemotherapy regimen with radiotherapy
    - BEACOPP chemotherapy regimen
- If no long term remission, 'salvage therapy' may include chemotherapy and/or radiotherapy to enable autoSCT.
- Some patients ineligible for autoSCT; typically because of lack of clinical response or factors such as age or comorbidity.
- Further detail on recommendations from the BCSH guidelines for the treatment of classical Hodgkin Lymphoma can be found in the Company's submission, section 3.5, table 5 (page 40).

NICE TA446: Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma
Recommendations:

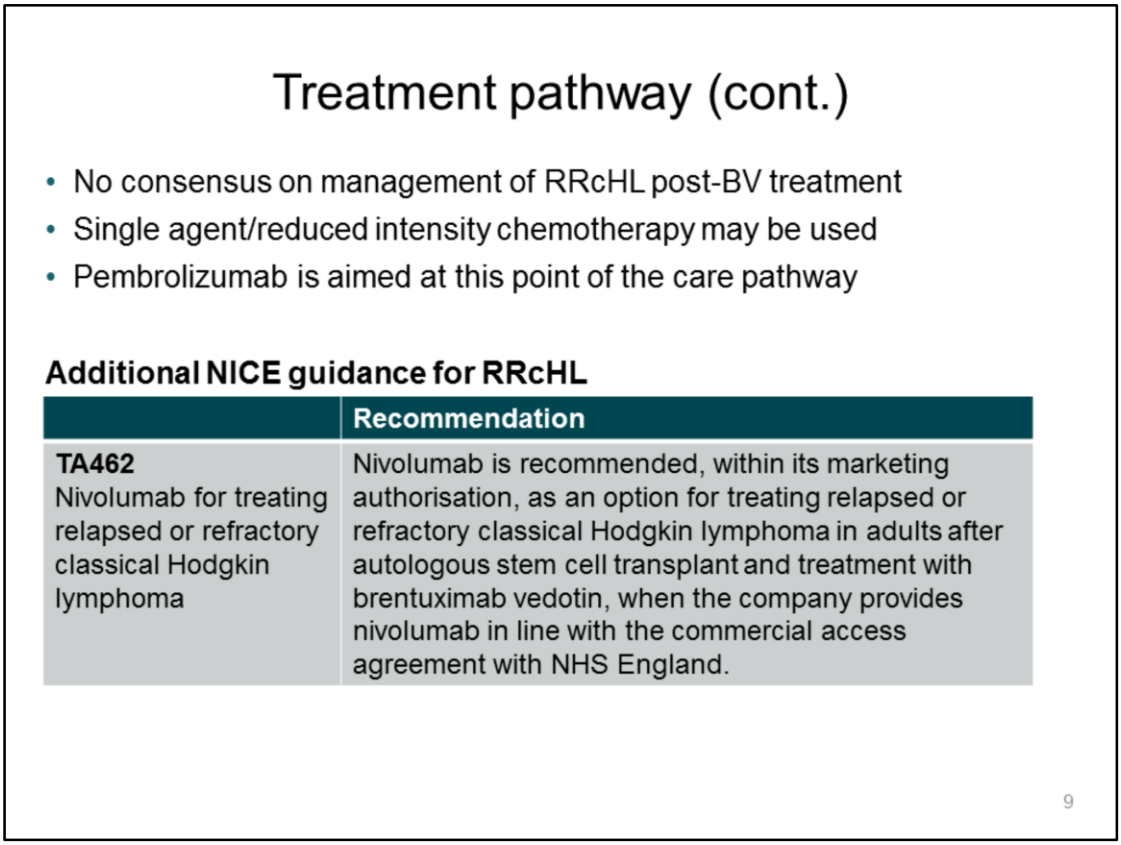
Brentuximab vedotin is recommended as an option for treating CD30-positive Hodgkin lymphoma in adults, only if:

- they have relapsed or refractory disease after autologous stem cell transplant and
- the company provides brentuximab vedotin at the price agreed with NHS England in the commercial access agreement.

1.2 Brentuximab vedotin is recommended for use within the Cancer Drugs Fund as an option for treating CD30-positive Hodgkin lymphoma in adults, only if:

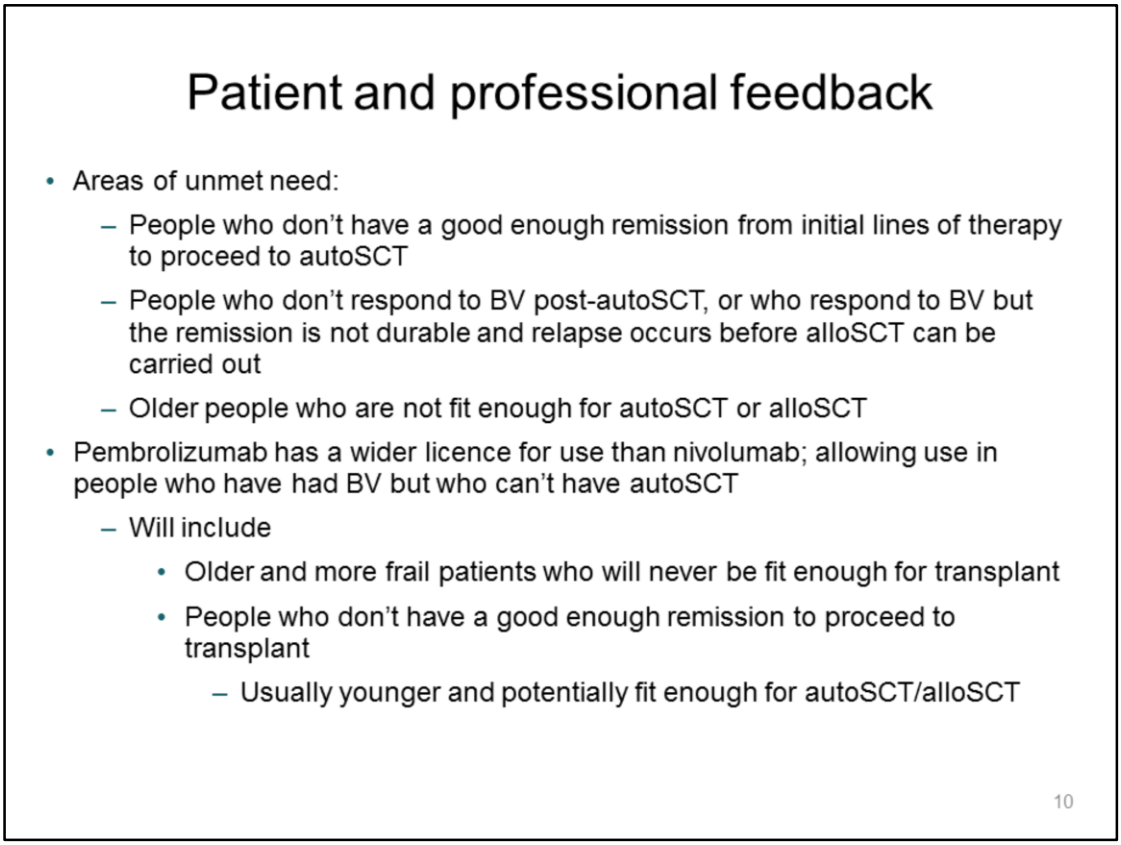
- they have relapsed or refractory disease after at least 2 previous therapies and
- they cannot have autologous stem cell transplant or multi-agent chemotherapy and
- the conditions of the managed access agreement are followed.

1.3 These recommendations are not intended to affect treatment with brentuximab vedotin that was started in the NHS before this guidance was published. Adults having treatment outside these recommendations may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop.

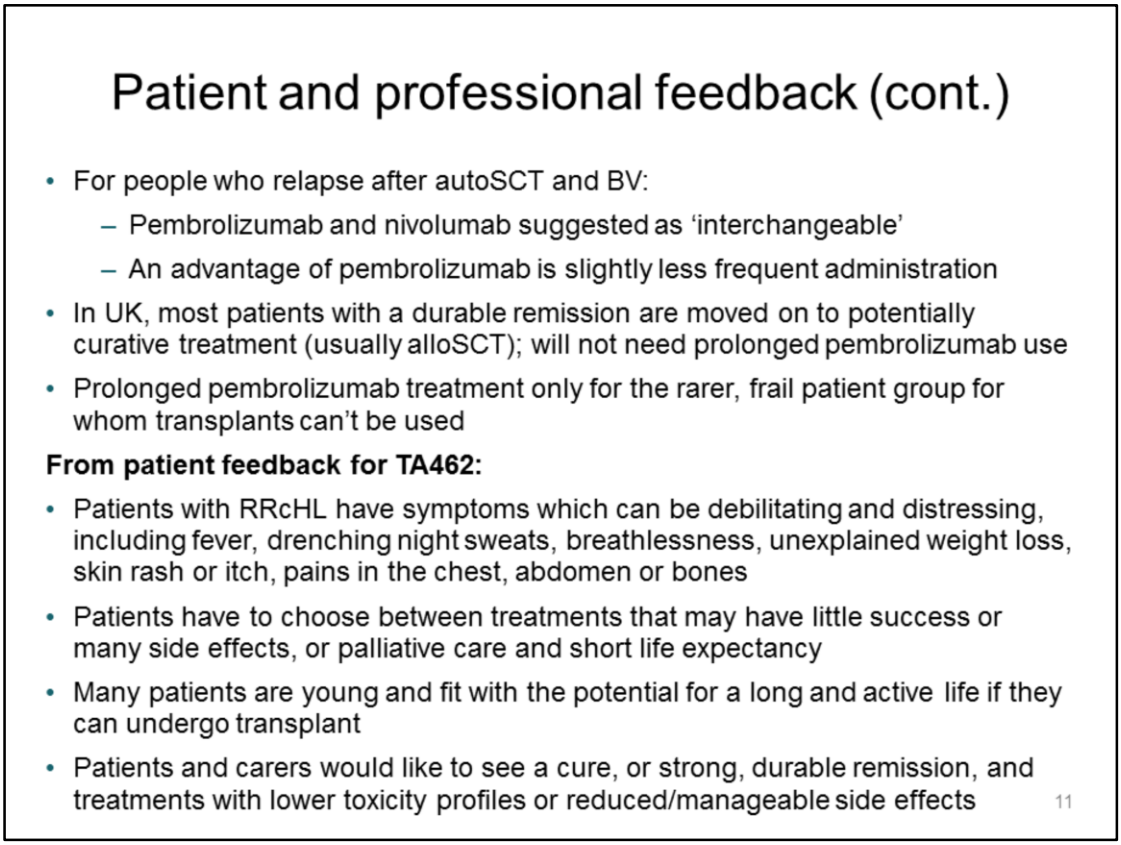


Source: Company submission, section 3.3 (page 38), section 3.5 (page 40); ERG report, section 2.2 (page 21)

Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma (2017) NICE technology appraisal guidance 462, published 26 July 2017



Source: Submission from the Royal College of Pathologists



Source: Submission from the Royal College of Pathologists; patient feedback for TA462

Decision problem

	NICE scope	Company submission	ERG's comments
Population	People with RRcHL who have received: • autoSCT and BV • BV when autoSCT is not a treatment option	As per NICE scope	-
Comparators	Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin Best supportive care (BSC)	Standard of care as per Cheah et al. (2016). BSC assessed as a subsequent therapy in base case and as a comparator in a scenario analysis	Cheah et al. includes multiple comparators – some of which are within scope, others are not. Broadly matches comparator in NICE scope This study was used to provide comparator data in TA462 ERG not aware of a more appropriate data source for SOC comparator
Outcomes	• Overall survival • Progression-free survival • Response rates • Adverse effects of treatment • Health-related quality of life	As per NICE scope; except no long term overall survival data	Mostly in-line with final scope. However survival data is immature and only 2 outcomes (PFS and ORR) have been included in indirect comparisons

12

Source: Company submission, section 1.1 (page 19); section 4.1.1; ERG report, section 3 (pages 22 to 26)

Standard of care as per Cheah et al. (2016) includes:

- Investigational agent(s)
- Gemcitabine
- Bendamustine
- Other alkylatory
- BV retreatment
- Platinum based
- autoSCT
- Other

Use of best supportive care (BSC) considered by company to be minimal at this stage of the treatment pathway (eligible patients will receive therapy if feasible). Therefore BSC applied by company in base-case model as subsequent therapy. Company submission, section 5.2.4 (page 148)

Company's clinical evidence
KEYNOTE-087

	KEYNOTE-087	
Design	Phase II single arm, open label trial	
Population	Adults with RRcHL after:	
	Cohort 1 (n=69; 4 from UK):	autoSCT and BV (post-autoSCT)
	Cohort 2 (n=81; 10 from UK)	Salvage chemotherapy and BV (no autoSCT)
	Cohort 3 (n= 60)	autoSCT (didn't receive BV post-autoSCT) – ***Not in assessment***
Setting	51 study sites: 26 Europe (3 in UK), 11 US, 7 Japan, 4 Israel, 2 Australia, 1 Canada.	
Intervention	Pembrolizumab 200mg as a 30 minute infusion every 3 weeks in an outpatient setting. On treatment for up to 2 years, or until unacceptable toxicity or progression.	
Outcomes	Primary: Overall response rate (ORR) / Safety and tolerability Secondary includes: ORR (investigator assessment), progression-free survival, duration of response and overall survival	

13

Source: Company submission: section 4.3.1 (page 50), section 4.7.1 (page 69)

On-going study. Efficacy data from most recent cut-off (21 March 2017).

No RCTs relevant to the decision problem were identified. Evidence for pembrolizumab is based on the ongoing KEYNOTE-087 trial. Evidence for SOC is based on a retrospective observational study (reported in Cheah et al. 2016; see later slides).

Rationale for use of single-arm, non-comparative trial: absence of established clinical practice in this later line setting and limited number of eligible participants. From company submission, section 4.3.1 (page 50).

Company submission focuses on cohorts 1 and 2 as per EMA licence requirements for pembrolizumab. From company submission, section 4.3.1 (page 50).

Cohort 1: People with RRcHL who have failed to achieve a response or progressed after autologous stem cell transplant (auto-SCT) and have relapsed after treatment with, or failed to respond to, brentuximab vedotin post auto-SCT.
Cohort 2: People with RRcHL who were unable to achieve a complete response (CR) or

partial response (PR) to salvage chemotherapy and did not receive auto-SCT, but have relapsed after treatment with, or failed to respond to, brentuximab vedotin.
Cohort 3: People with RRcHL who have failed to respond to, or progressed after, auto-SCT and have not received brentuximab vedotin post auto-SCT. These patients may or may not have received brentuximab vedotin as part of primary or salvage treatment.

The analysis of primary efficacy endpoints were based on the All Subjects as Treated (ASaT) population, i.e., patients will be included if they receive at least one dose of study medication. From company submission, section 4.4 (page 61).

Overall Response Rate (ORR): The proportion of patients in the analysis population who have complete remission (CR) or partial remission (PR) using IWG criteria (Cheson 2007) at any time during the study. Response for the primary analysis was determined by blinded, independent central review (BICR). Company submission (page 55).

Further exploratory end-points include changes in health-related quality-of-life assessments from baseline using the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life (QoL) Questionnaire C30 (QLQ-C30) and European Quality of Life Five Dimensions Questionnaire (EuroQoL EQ-5D). Full details can be found in the company submission, section 4.3.1 (pages 55 and 56).

CONFIDENTIAL

Company's clinical evidence

KEYNOTE-087: Response rates

	Response at week 12		Best overall response (at March 2017)	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2
CR (n)	[redacted]	[redacted]	27.5% (19)	24.7% (20)
PR (n)	[redacted]	[redacted]	47.8% (33)	42.0% (34)
OR [CR+PR] (n)	[redacted]	[redacted]	75.4% (52)	66.7% (54)
SD (n)	[redacted]	[redacted]	[redacted]	[redacted]
PD (n)	[redacted]	[redacted]	[redacted]	[redacted]
No assessment	[redacted]	[redacted]	[redacted]	[redacted]
Median time to response (range)			[redacted]	[redacted]
Median response duration			[redacted]	[redacted]

CR: Complete remission, PR: Partial remission, OR: Objective response, SD: Stable disease; PD: Progressive disease

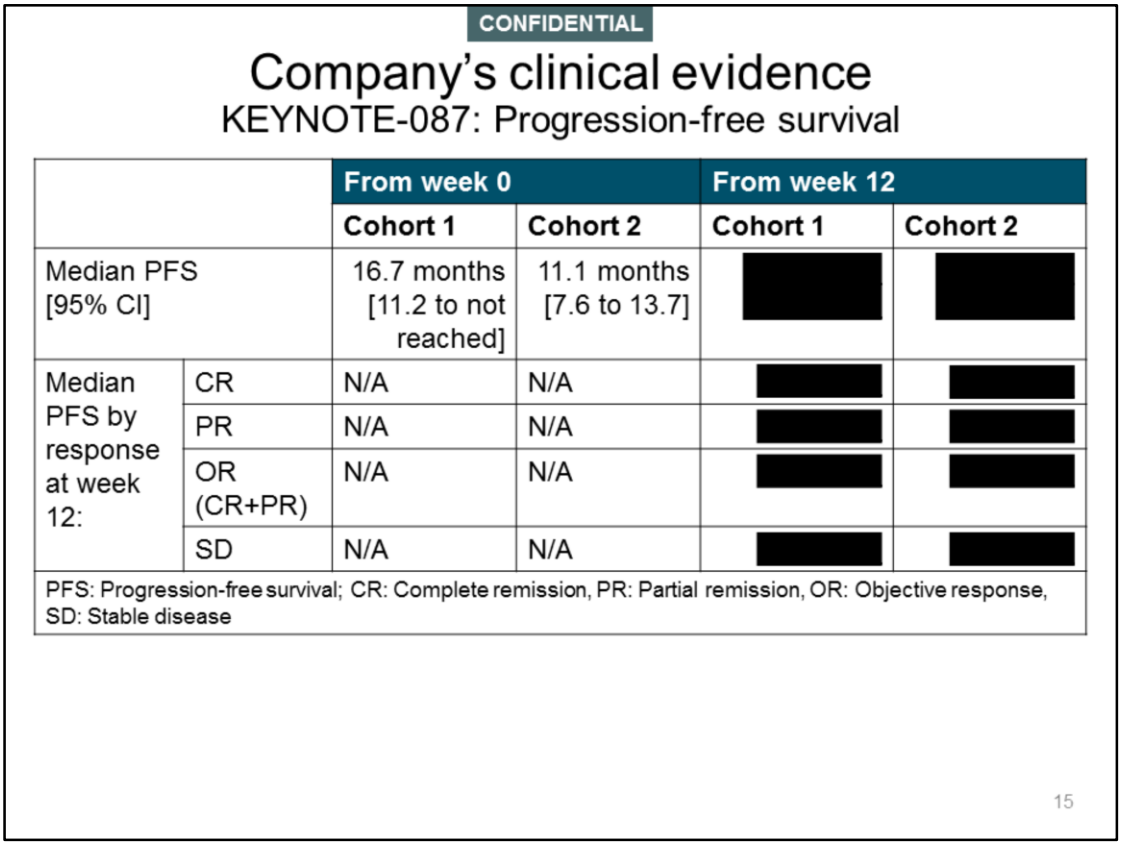
14

Source: Company submission, section 4.7, table 14 (page 71), table 15 (page 72); section 4.8.2, table 19 (page 79)

Data from week 12 used in cost-effectiveness model (see later slides).

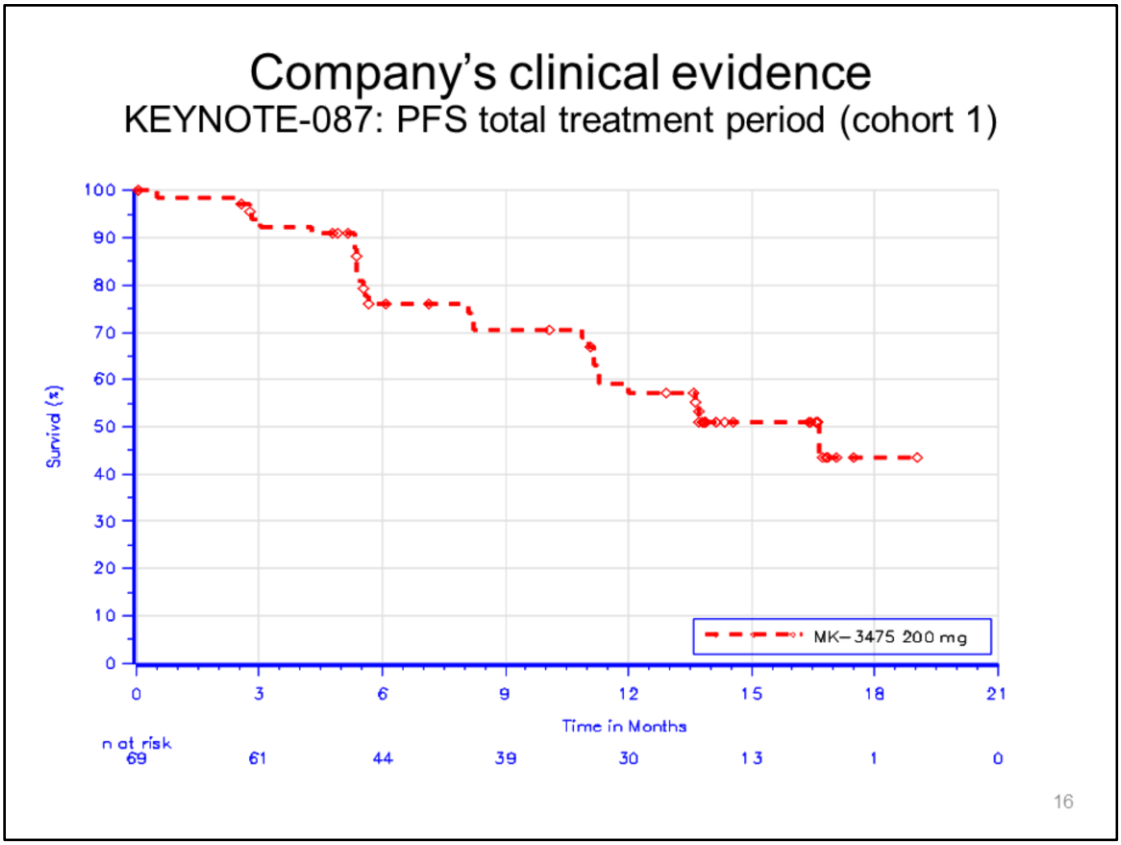
Overall Response Rate (ORR): The proportion of patients in the analysis population who have complete remission (CR) or partial remission (PR) using IWG criteria at any time during the study. Response for the primary analysis was determined by blinded, independent central review (BICR). From company submission, section 5.3.1 (page 55)

Duration of response: For the subgroup of patients who achieve CR or PR, the time from start of the first documentation of objective tumour response (CR or PR) to the first documentation of tumour progression or death due to any cause, whichever comes first.

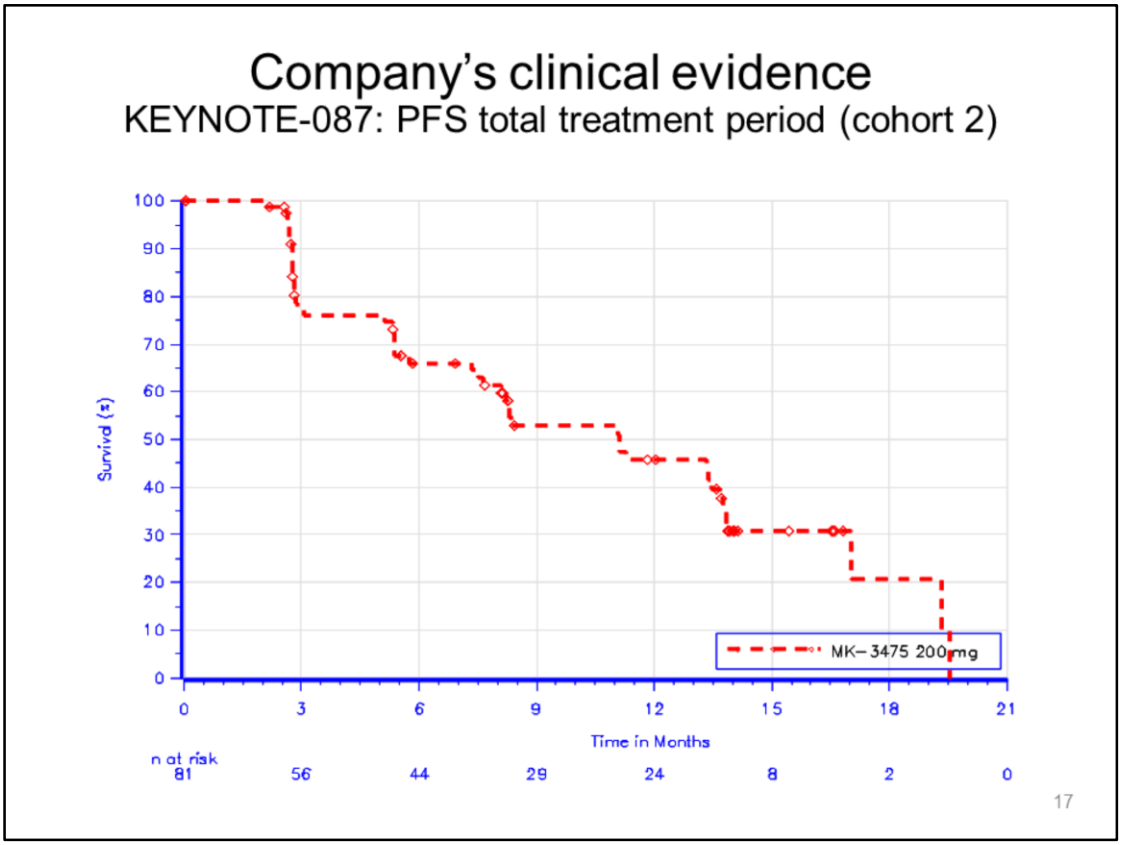


Source: Company submission, section 4.7, table 17 (page 74), section 4.8.2, table 21 (page 80) and table 22 (page 82)

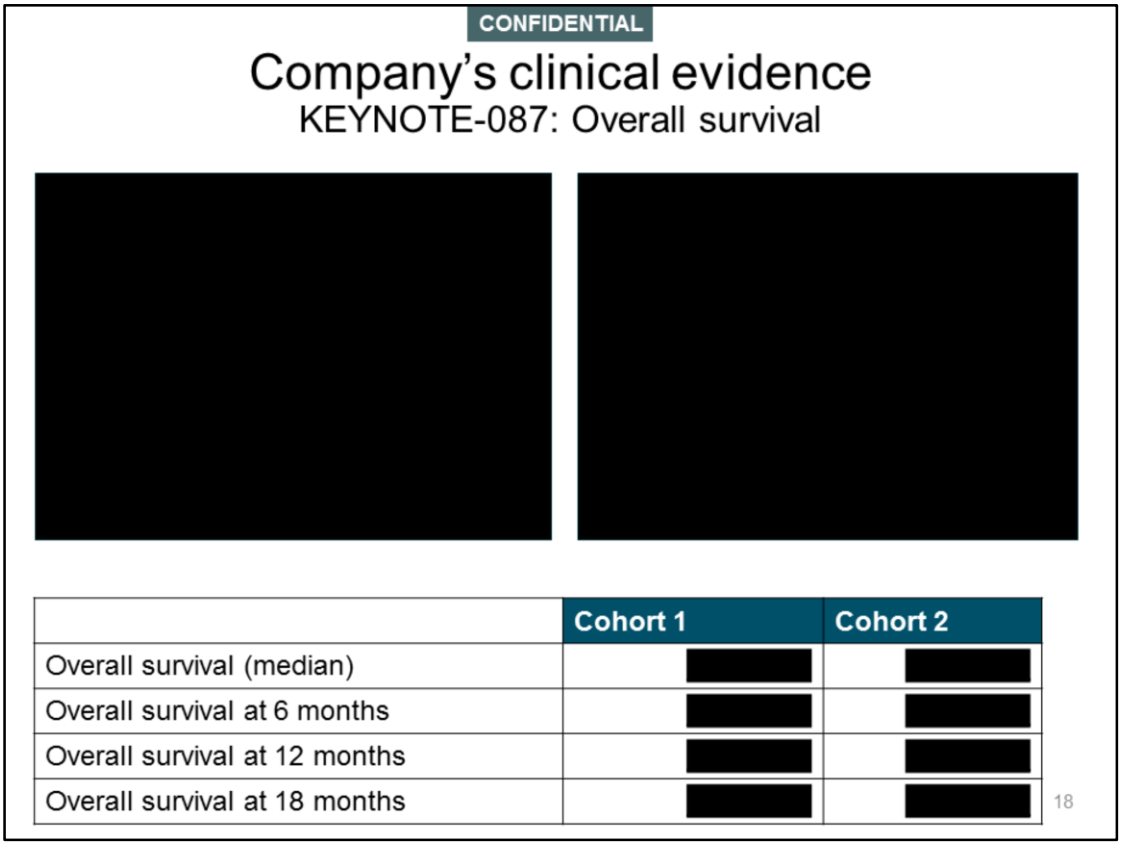
PFS is computed from a patients start point towards the first documented progression of disease according to IWG criteria or death due to any cause, whichever occurs first, expressed in days. Patients without an event (progression or death) at the time of last tumour assessment are considered right censored at the last disease assessment date. Responses are based on BICR assessment using IWG criteria (Cheson 2007). From company submission, section 4.8.1 (page 76).



Source: Company submission, figure 8 (page 85)

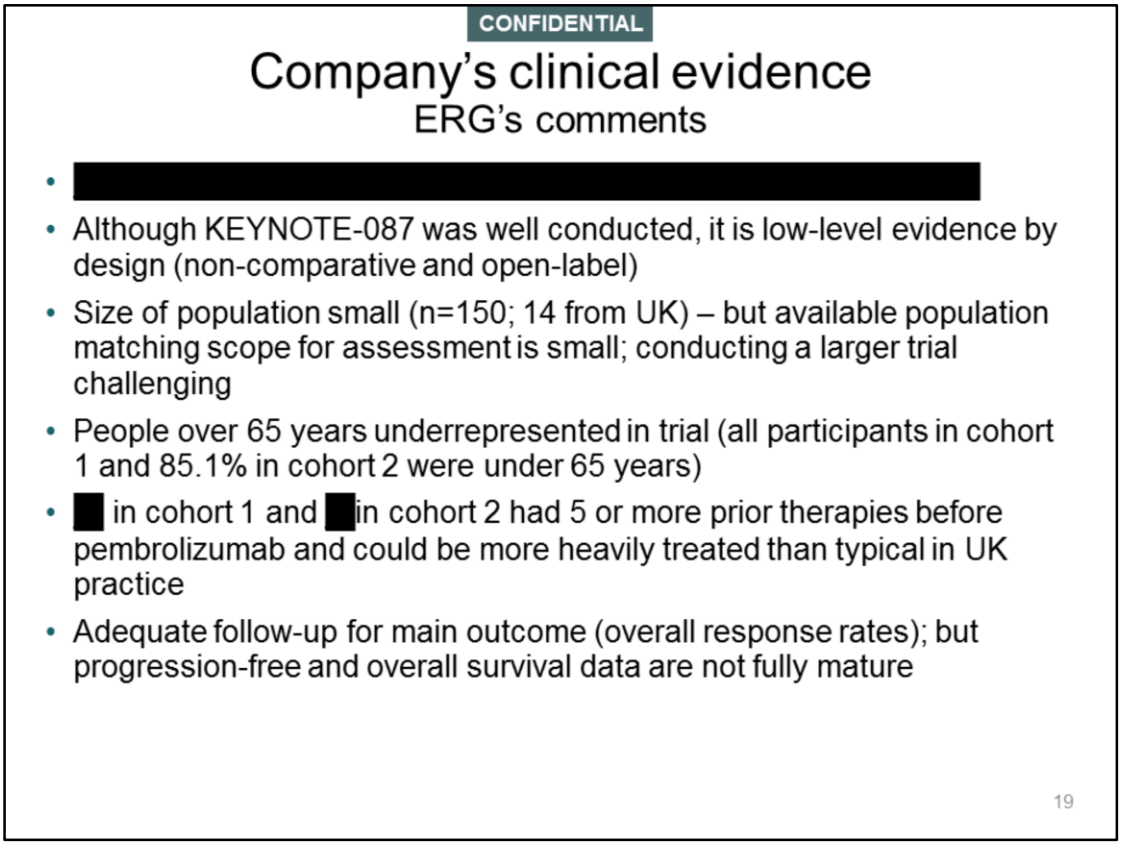


Source: Company submission, figure 9 (page 85)

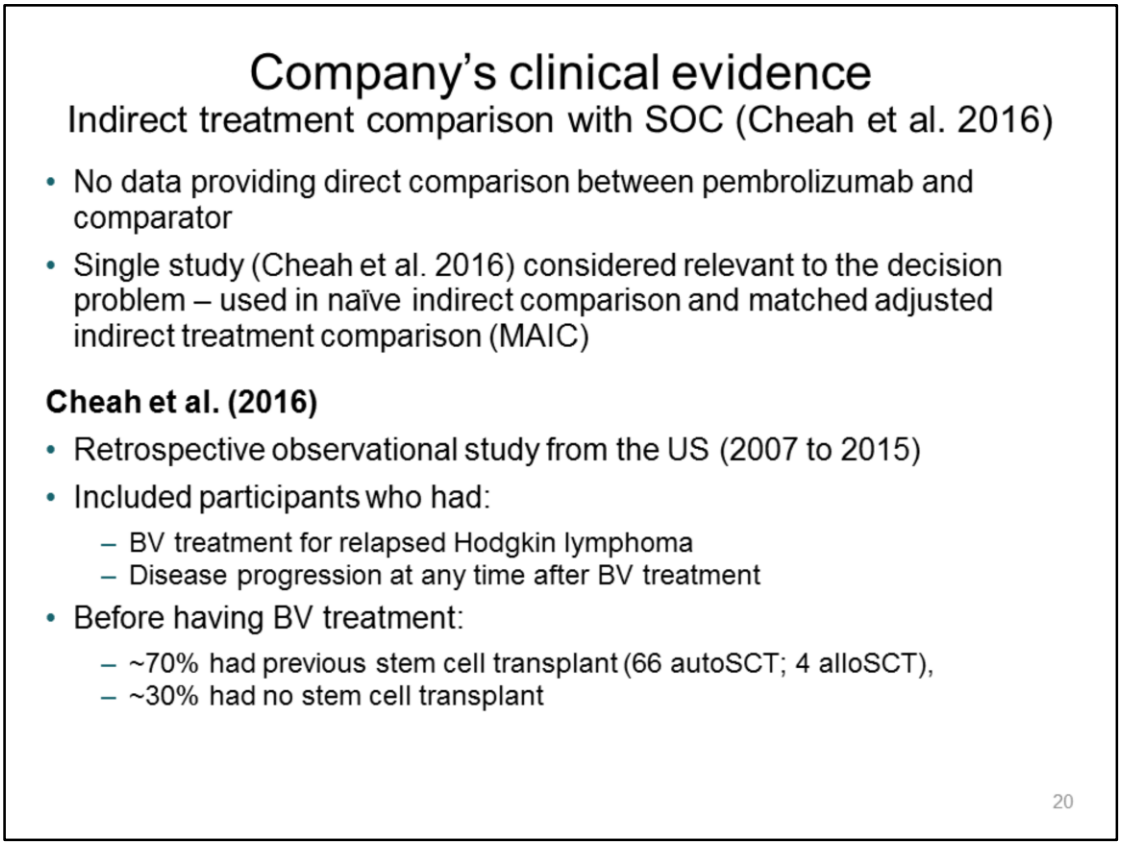


Source: Company submission: figures 10 and 11 (page 87); tables 25 and 26 (page 86)

Overall survival (OS) was defined as time from first dose intake to death due to any cause, expressed in days. Patients without documented death are considered right censored at the day of last contact. Patients who had a survival update after the data cut-off date of March 2017 are censored at the cut-off date. From company submission, section 4.8.1 (page 77).



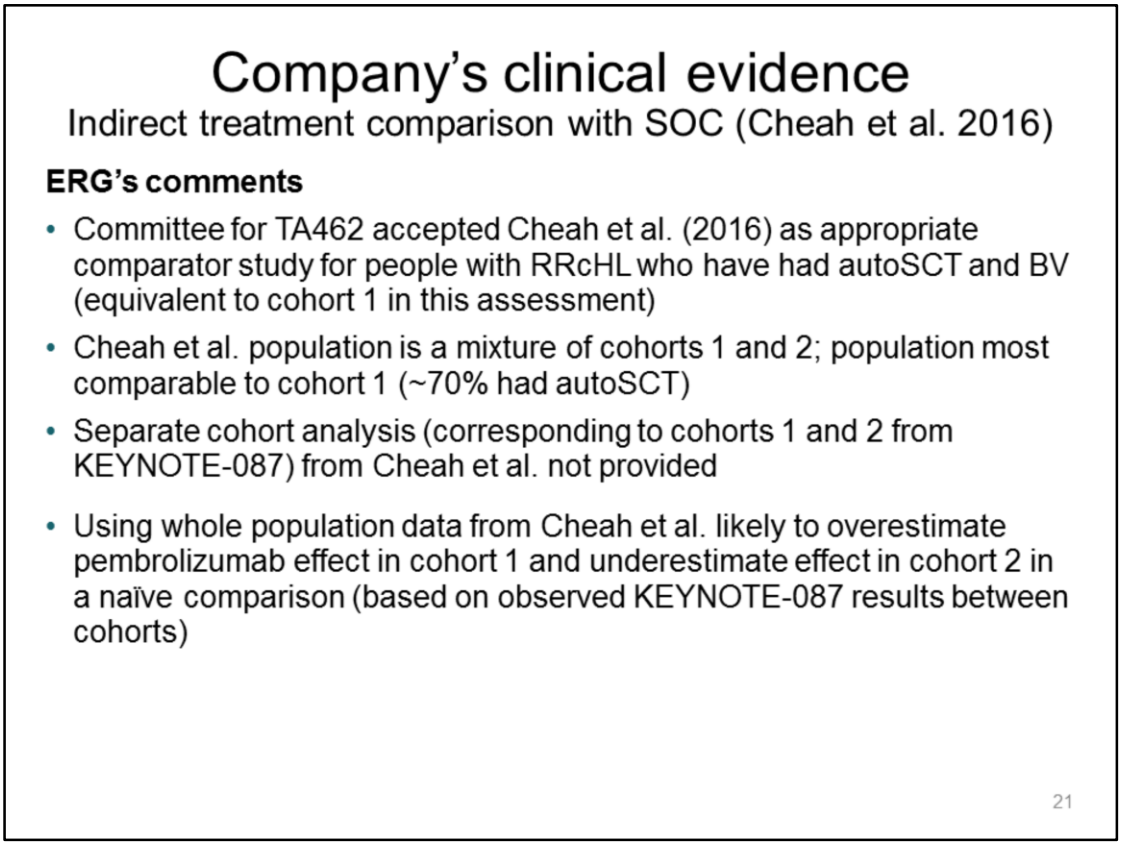
Source: ERG report, section 4.2 (pages 31 to 38); section 4.2.2.5 (page 39); section 4.2.2.6 (page 41),



Source: Company submission, section 4.10.4 (page 89), section 4.10.8 (page 91); ERG report, section 4.3 (pages 46 to 49)

Further details of Cheah et al. (2016) can be found in the company submission, section 4.10.4 (page 89). Further details on methods used for the indirect comparisons can be found in the company submission, section 4.10.12 (pages 92 and 93) and in the ERG report, section 4.4.1 (page 50 onwards)

The MAIC used weighting to match individual patient data from KEYNOTE-087 to summary data from Cheah et al. Initial matching used all variables available in both KEYNOTE-087 and Cheah et al. These were: ECOG >0 (%), B symptoms (%), Age >45 (%), Albumin <40 g/l (%), Haemoglobin <105 g/l (%), Lymphocytes <0.6 x 109 (%), White blood cells >15 x 109 (%), Max Tumour Diameter >4 cm (%), Any extranodal site (%), Female (%), and Prior lines (mean/median). Variables were only excluded from the matching if there were problems with model convergence. Most analyses only excluded one variable 'median prior lines', but the analysis of ORR for cohort 1 in the 12-week scenario only included four variables in the matching model. From ERG report, section 4.4.1 (page 52)



Source: ERG report, section 4.3 (pages 46 and 47)

CONFIDENTIAL

Company's clinical evidence

Indirect comparison: Progression-free survival

Cohort	Comparison	Hazard ratio (95% CI) Pembrolizumab (KEYNOTE-087) versus SOC (Cheah)	
		From study initiation to week 12	From study initiation to most recent observation
1	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]

CI: Confidence interval; MAIC: Matched Adjusted Indirect treatment comparison; SOC: Standard of care

- Hazard ratio for cohort 1 more favourable to pembrolizumab in the MAIC

ERG comments

- Almost all PFS results show significant benefit for pembrolizumab versus SOC
- One exception: naïve comparison in cohort 1 at week 12 – non-significant difference favouring pembrolizumab

22

Source: Company submission, tables 27 and 28 (page 96); ERG report, section 4.4.2 (page 47)

Methods used for the naïve indirect comparison and matched adjusted indirect treatment comparison (MAIC) for progression-free survival analysis can be found in the company submission, section 4.10.12 (pages 92 and 93).

CONFIDENTIAL

Company's clinical evidence

Indirect comparison: Objective response rate (ORR)

Cohort	Comparison	Odds ratio (95% CI) Pembrolizumab (KEYNOTE-087) versus SOC (Cheah)	
		Response at week 12 (KEYNOTE-087) versus best overall response (Cheah et al.)	Best overall response
1	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]

CI: Confidence interval; MAIC: Matched Adjusted Indirect treatment comparison; SOC: Standard of care

- MAIC increases odds ratio (relative to naïve comparison)

ERG comment

- All results for ORR significantly favour pembrolizumab over SOC

23

Source: Company submission, tables 29 and 30 (page 97); ERG report, section 4.4.2 (page 54)

Response at week 12: Compares response at 12 weeks in KEYNOTE-087 (Pembrolizumab) and best overall response in Cheah et al. (SOC).

Odds ratios for ORR were also provided using Cheah et al. with data from patients who received investigational agents (n=28) removed. See ERG report, section 4.4.2, table 4.15 (page 54)

CONFIDENTIAL

Company's clinical evidence

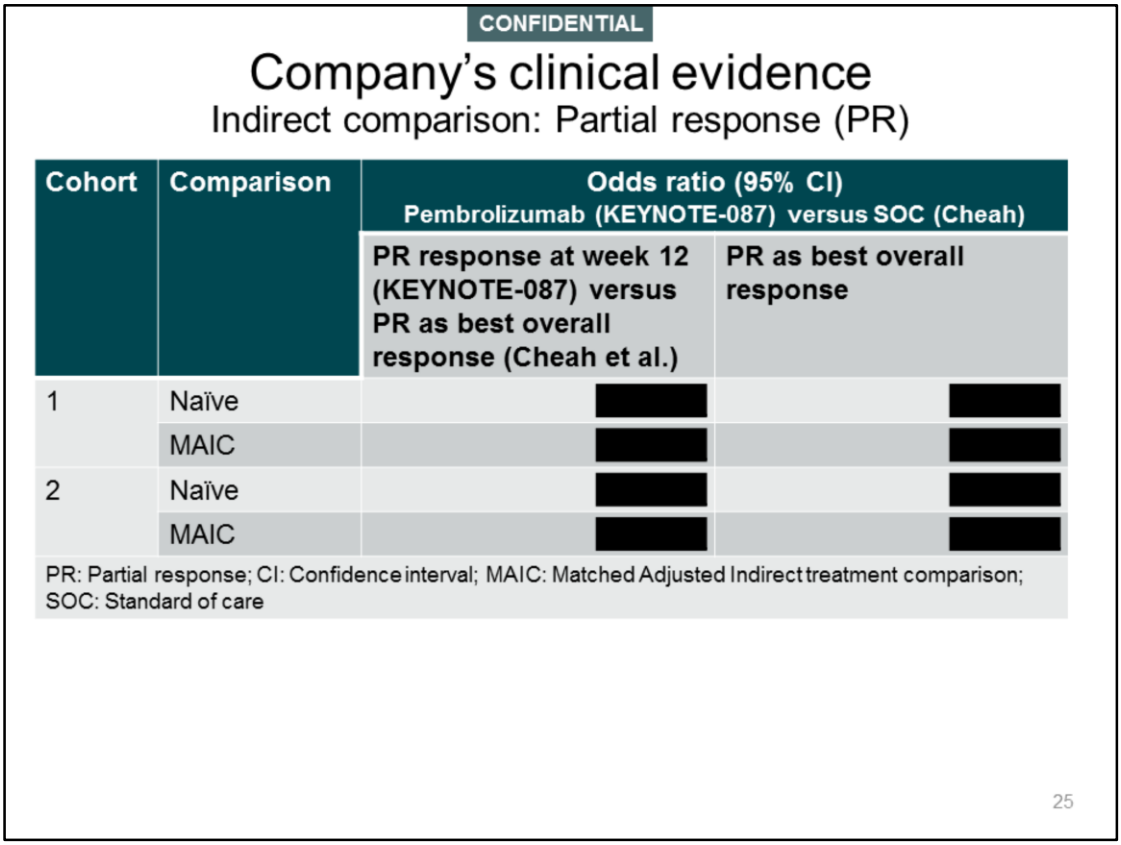
Indirect comparison: Complete response (CR)

Cohort	Comparison	Odds ratio (95% CI) Pembrolizumab (KEYNOTE-087) versus SOC (Cheah)	
		CR response at week 12 (KEYNOTE-087) versus CR as best overall response (Cheah et al.)	CR as best overall response
1	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]

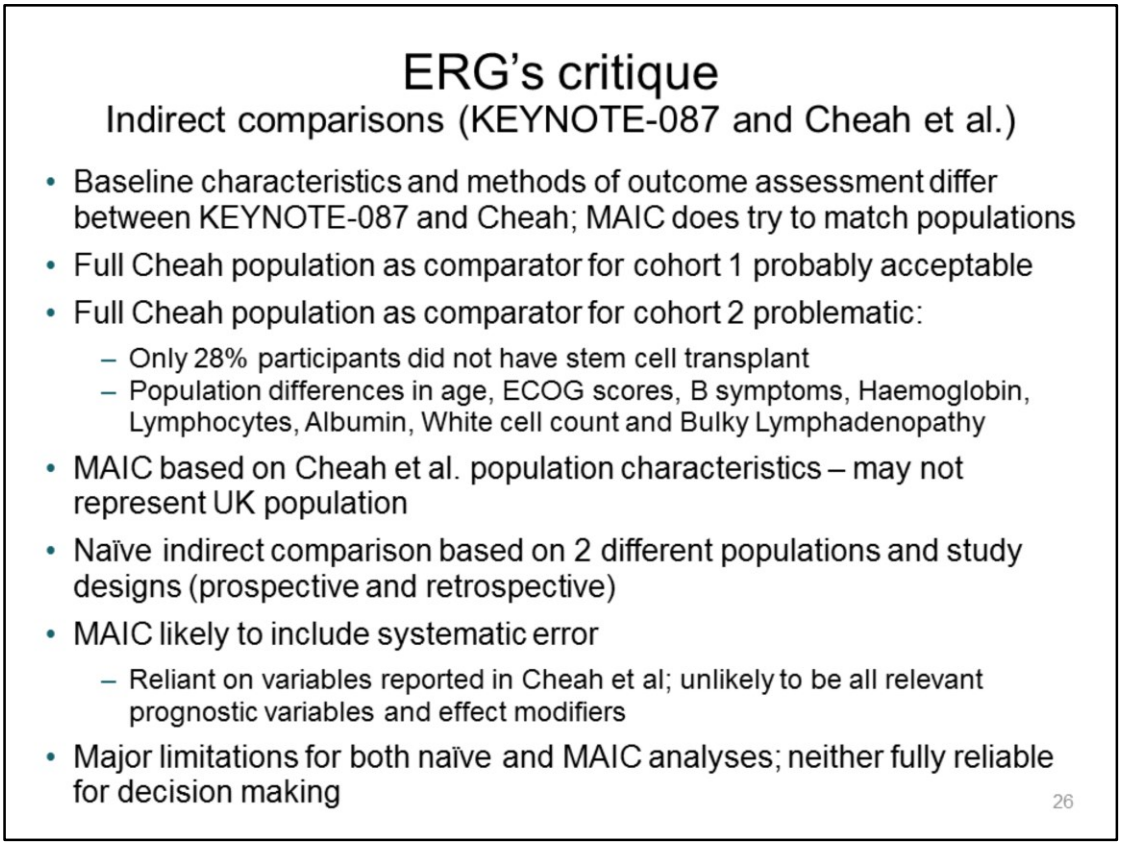
CR: Complete response; CI: Confidence interval; MAIC: Matched Adjusted Indirect treatment comparison; SOC: Standard of care

24

Source: Company submission, tables 31 and 32 (pages 98 and 99).



Source: Company submission, tables 33 and 34 (page 100)



Source: ERG report, section 4.4.1 (pages 50 to 53)

In the economic model, the naïve indirect comparison results are used in the base-case analysis, and the MAIC results are used in a scenario analysis

CONFIDENTIAL

Company's clinical evidence

Adverse events: KEYNOTE-087

	Cohort 1 (n=69)	Cohort 2 (n=81)
1 or more adverse events (n)	[redacted]	[redacted]
Drug related adverse event* (n)	[redacted]	[redacted]
Toxicity grade 3-5 adverse event (n)	[redacted]	[redacted]
Toxicity grade 3-5 drug-related adverse events (n)	[redacted]	[redacted]
Non-serious adverse events (n)	[redacted]	[redacted]
Serious adverse events (n)	[redacted]	[redacted]
Serious drug-related adverse events (n)	[redacted]	[redacted]
Discontinued due to an adverse event (n)	[redacted]	[redacted]
Discontinued due to drug related adverse event (n)	[redacted]	[redacted]
Discontinued due to a serious drug-related adverse event (n)	[redacted]	[redacted]

* Determined by investigator to be related to the drug

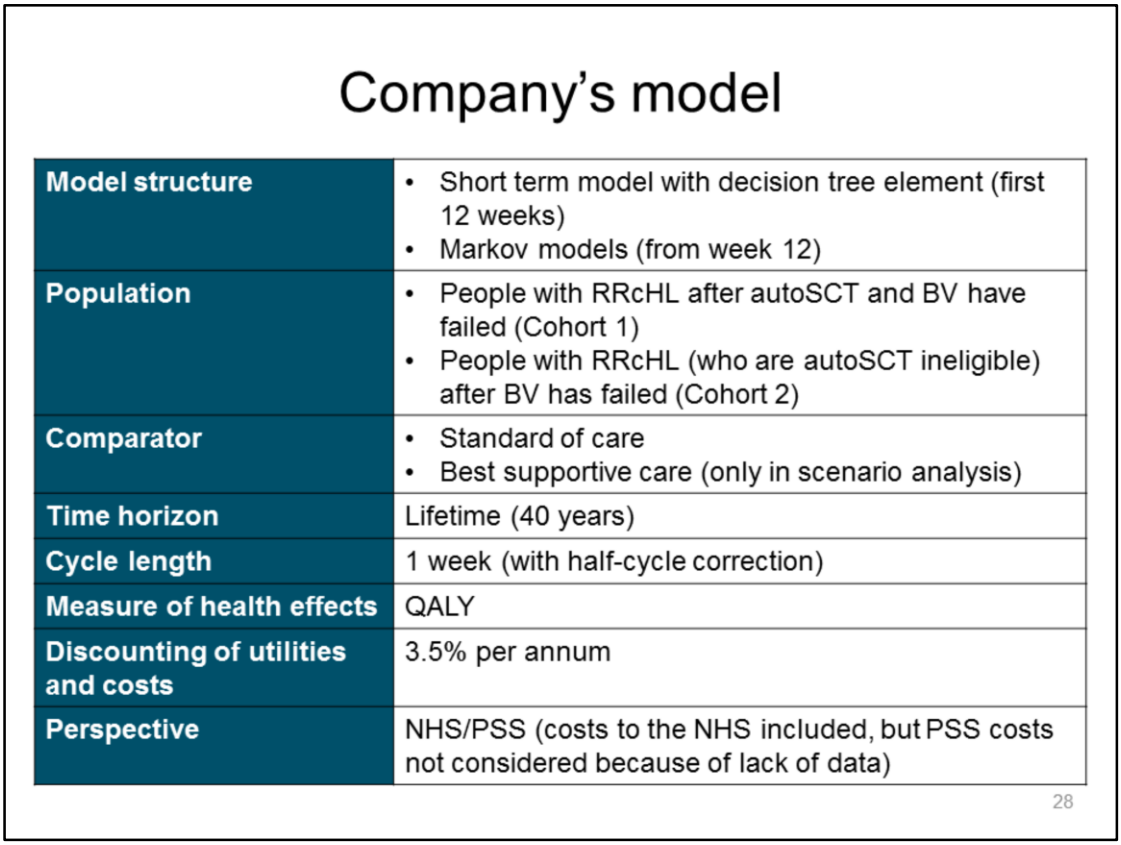
27

Source: Adapted from company submission, section 4.12.1, table 39 (pages 107 and 108)

Data from KEYNOTE-087 with a cut-off date of 25 September 2016

ASaT population: all enrolled patients who received at least one dose of study treatment.

Further detail on adverse reactions from KEYNOTE-087 can be found in the company submission, section 4.12 (starts page 105)



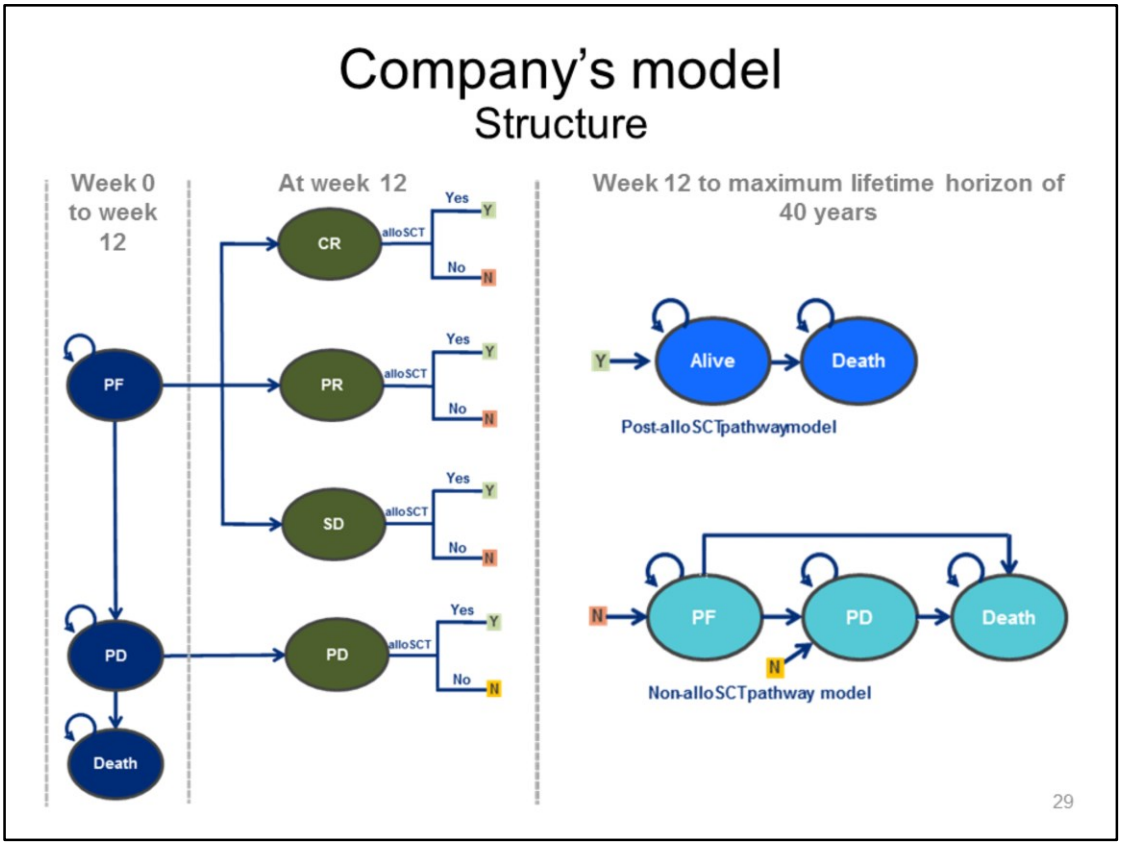
Company's model

Model structure	• Short term model with decision tree element (first 12 weeks) • Markov models (from week 12)
Population	• People with RRcHL after autoSCT and BV have failed (Cohort 1) • People with RRcHL (who are autoSCT ineligible) after BV has failed (Cohort 2)
Comparator	• Standard of care • Best supportive care (only in scenario analysis)
Time horizon	Lifetime (40 years)
Cycle length	1 week (with half-cycle correction)
Measure of health effects	QALY
Discounting of utilities and costs	3.5% per annum
Perspective	NHS/PSS (costs to the NHS included, but PSS costs not considered because of lack of data)

28

Source: Company submission, section 5.2.1 (page 137), section 5.2.2. (page 138), section 5.2.3 (page 145), section 5.2.4 (page 148).

The use of best supportive care (BSC) suggested to be minimal at this stage of the care pathway (based on BCSH guidelines and clinician opinion) because eligible patients likely to receive treatment where possible. From company submission, section 5.2.4. (page 148)



Source: Adapted from company submission, section 5.2.2, figure 13 (page 140); section 5.2.2 (pages 142 and 143)

Short-term decision tree model (weeks 0 to 12)

- Patients enter the model as progression free (PF) and receive treatment (pembrolizumab or SOC)
- Over first 12 weeks:
    - Patients can remain PF, progress (PD) or die
- At 12 weeks, patients who are PF are partitioned by response:
    - Complete response (CR)
    - Partial response (PR)
    - Stable disease (SD)
- At week 12, patients can have alloSCT or continue on treatment (pembrolizumab or SOC)
- Probability of having alloSCT depends on response status of patient (CR,PR,SD or PD)
    - Assumed that no people with progressed disease have alloSCT

Longer term Markov model (week 12 to death)

- After the decision tree, patients enter one of 2 independent Markov models depending

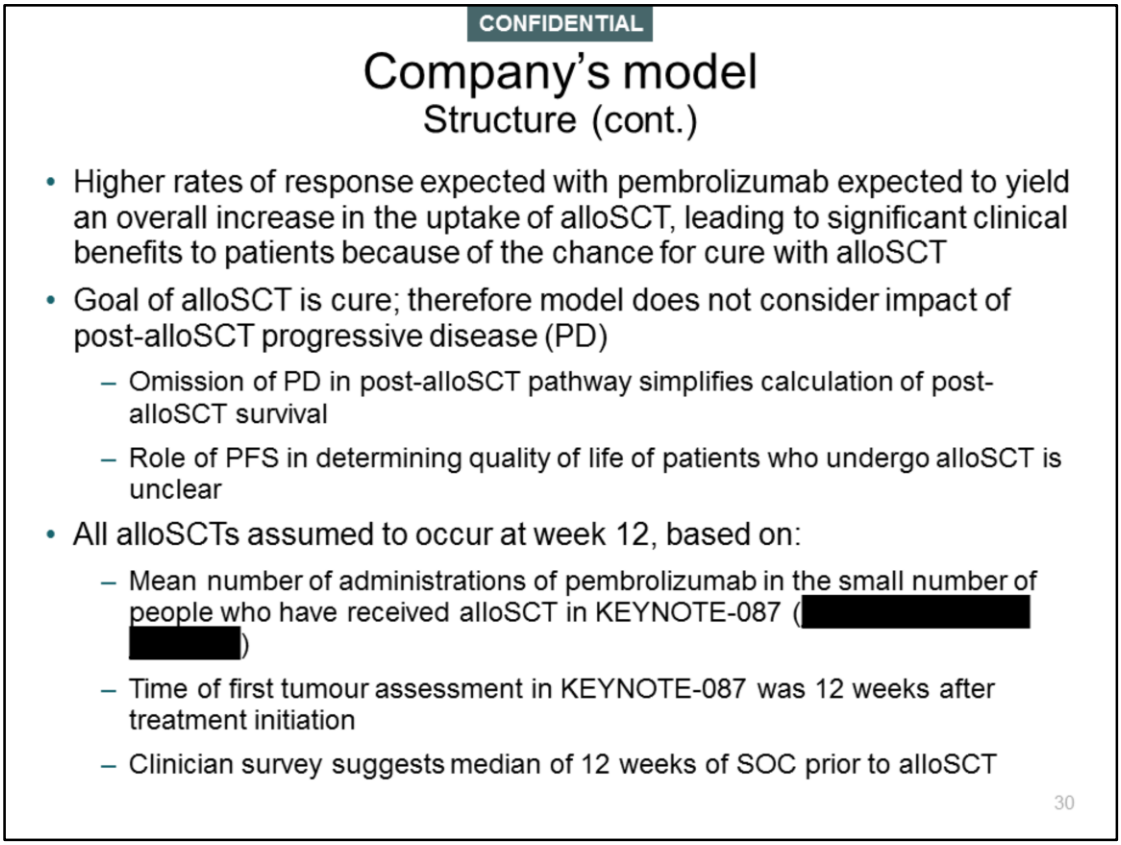
on whether they have alloSCT or not

Non-alloSCT pathway:

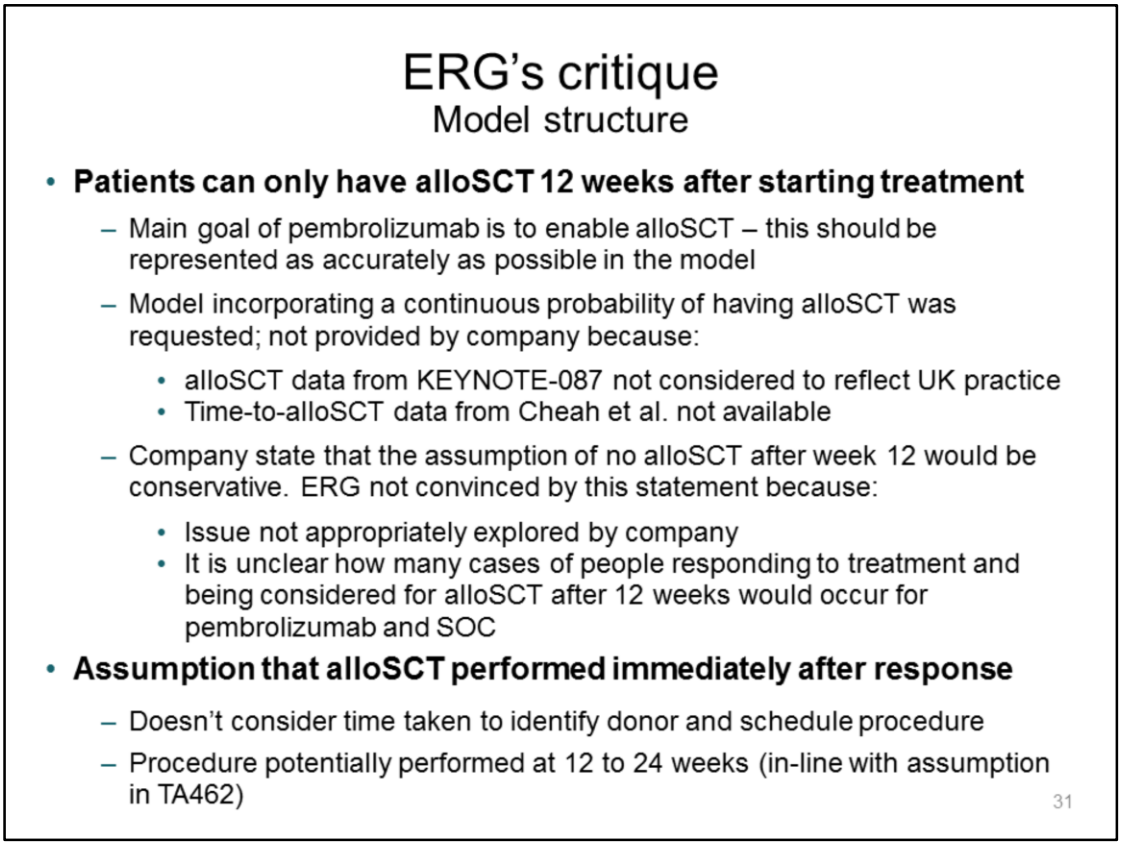
- Patients allocated to PF state (if CR, PR or SD at week 12) or PD (if PD at week 12)
- Patients in PF state continue with treatment (pembrolizumab or SOC) until toxicity, PD or death

Post-alloSCT pathway

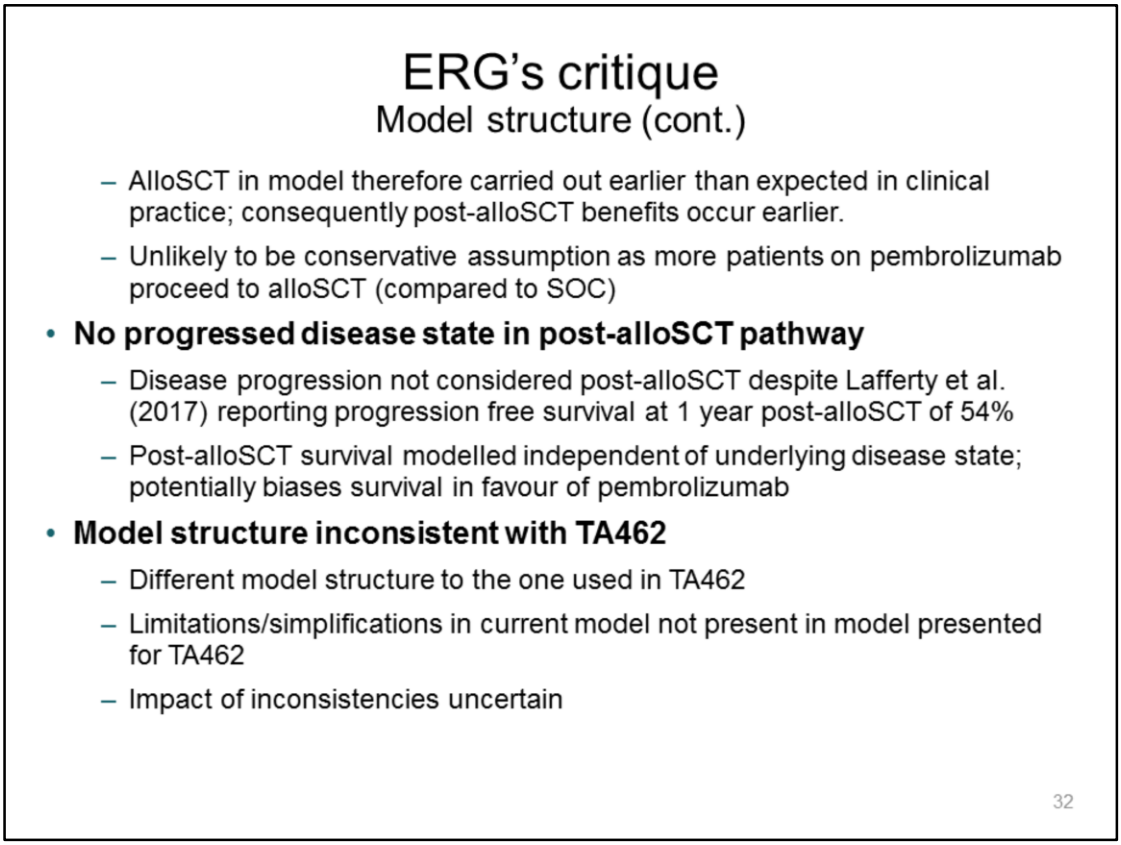
- Patients enter 'alive' state at week 12 and discontinue previous treatment
- Outcomes assumed to be the same regardless of previous treatment (pembrolizumab or SOC)
- Health state utility in 'alive' state varies between first 100 days and post-100 days to account for effect of treatment and recovery
- No PD state in the post-alloSCT pathway



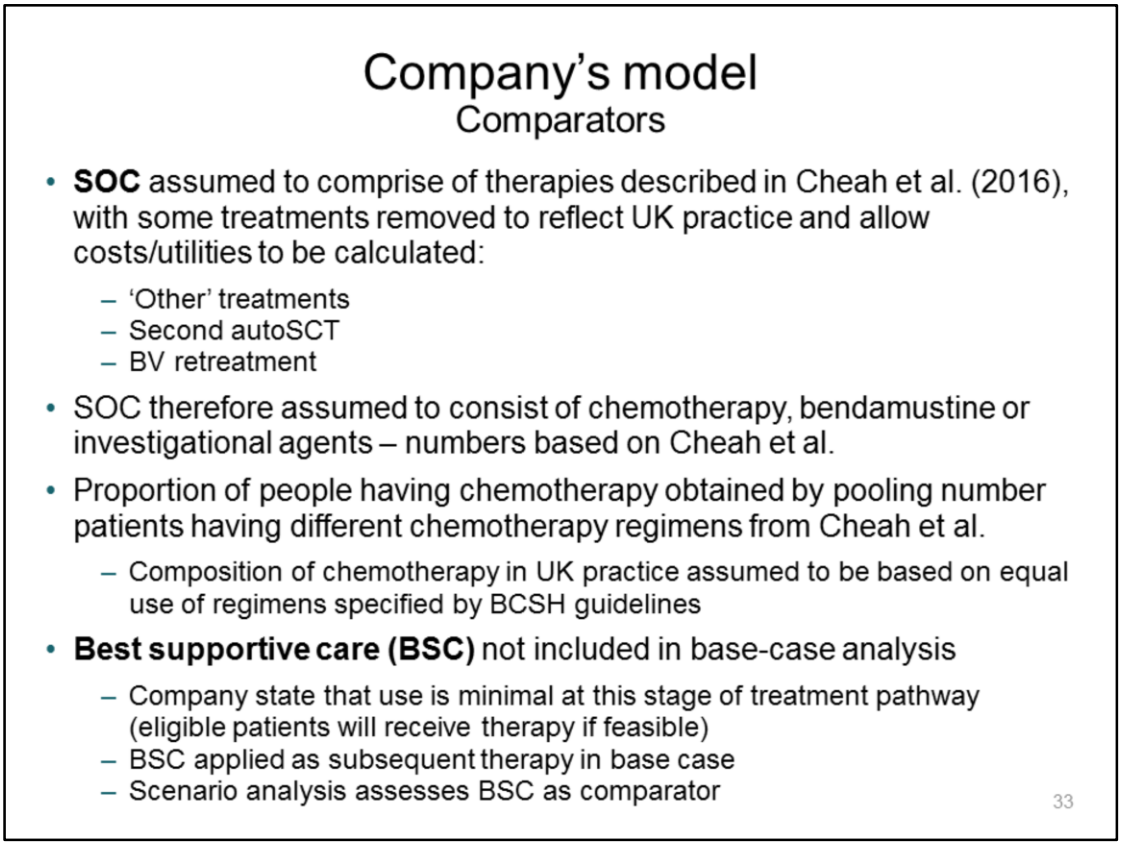
Source: Company submission, section 5.2.2 (pages 142 and 143); section 5.6.3 (page 216)



Source: ERG report, section 5.2.2 (pages 64 and 65)

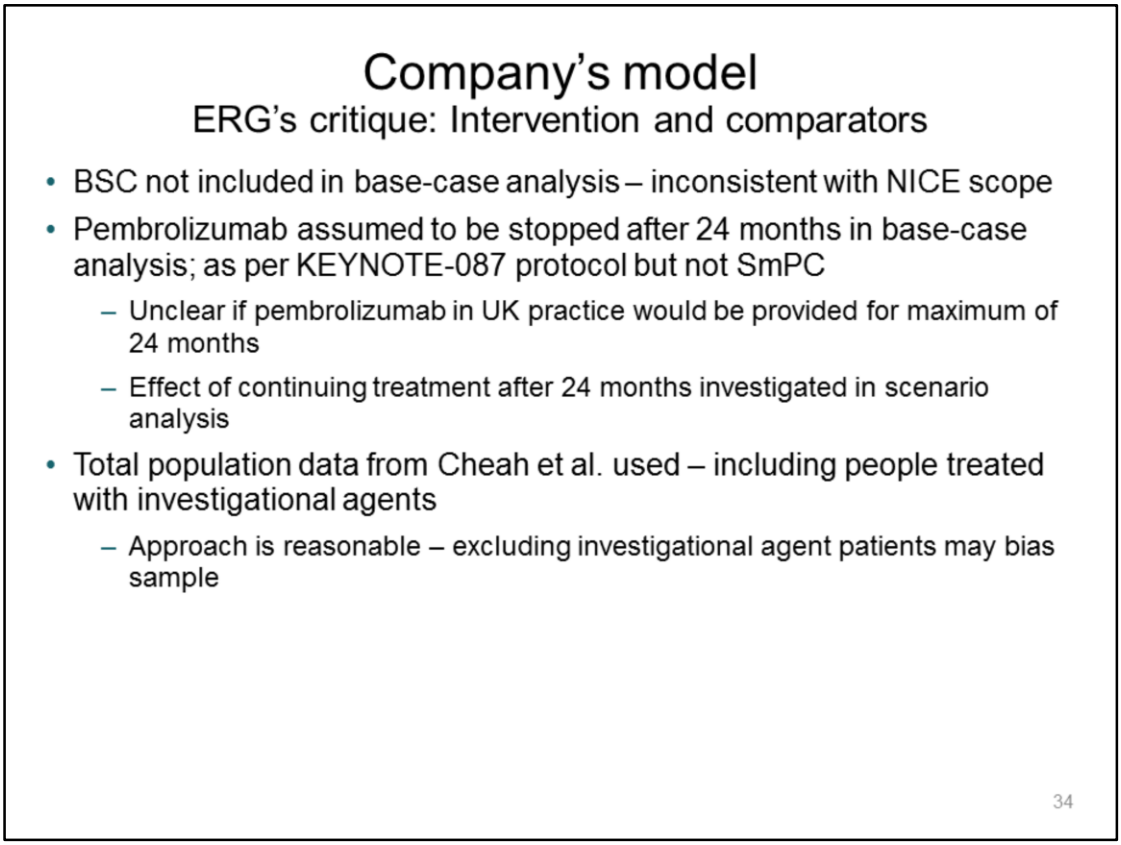


Source: ERG report, section 5.2.2 (pages 64 and 65)

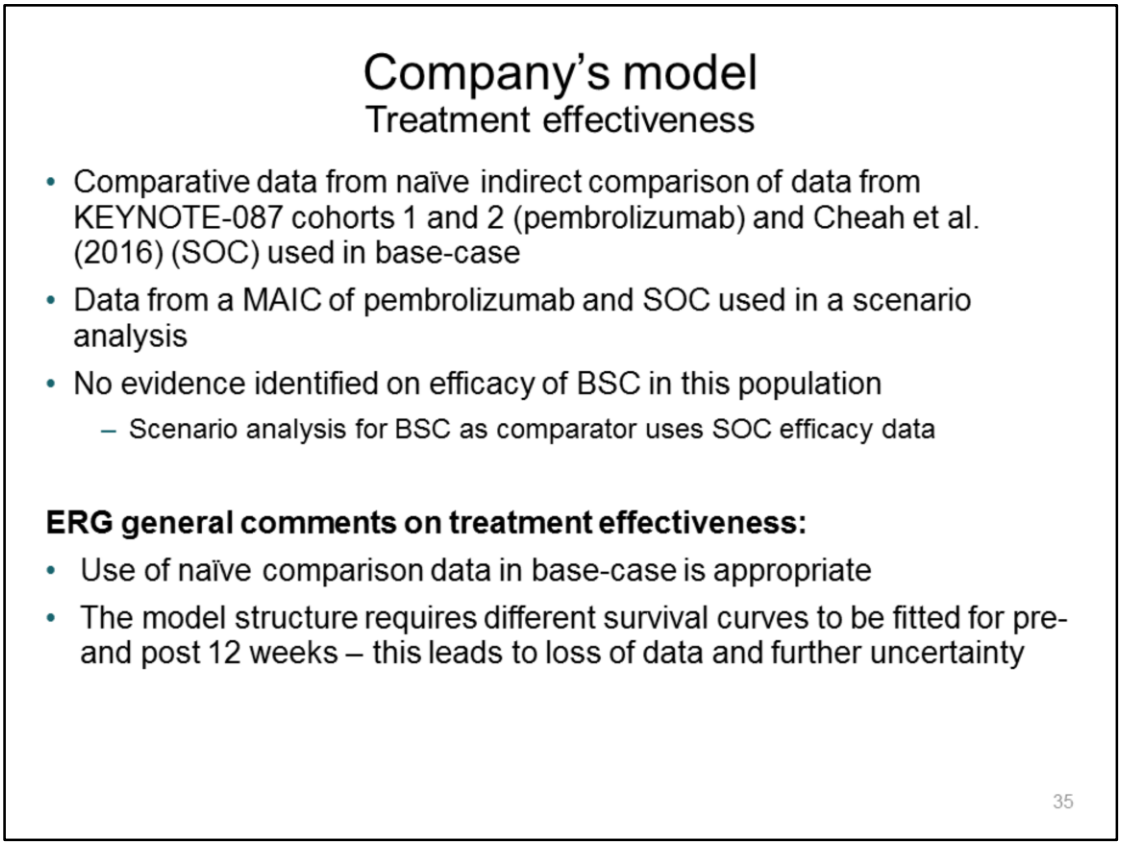


Source: Company submission, section 5.2.4 (pages 146 and 147)

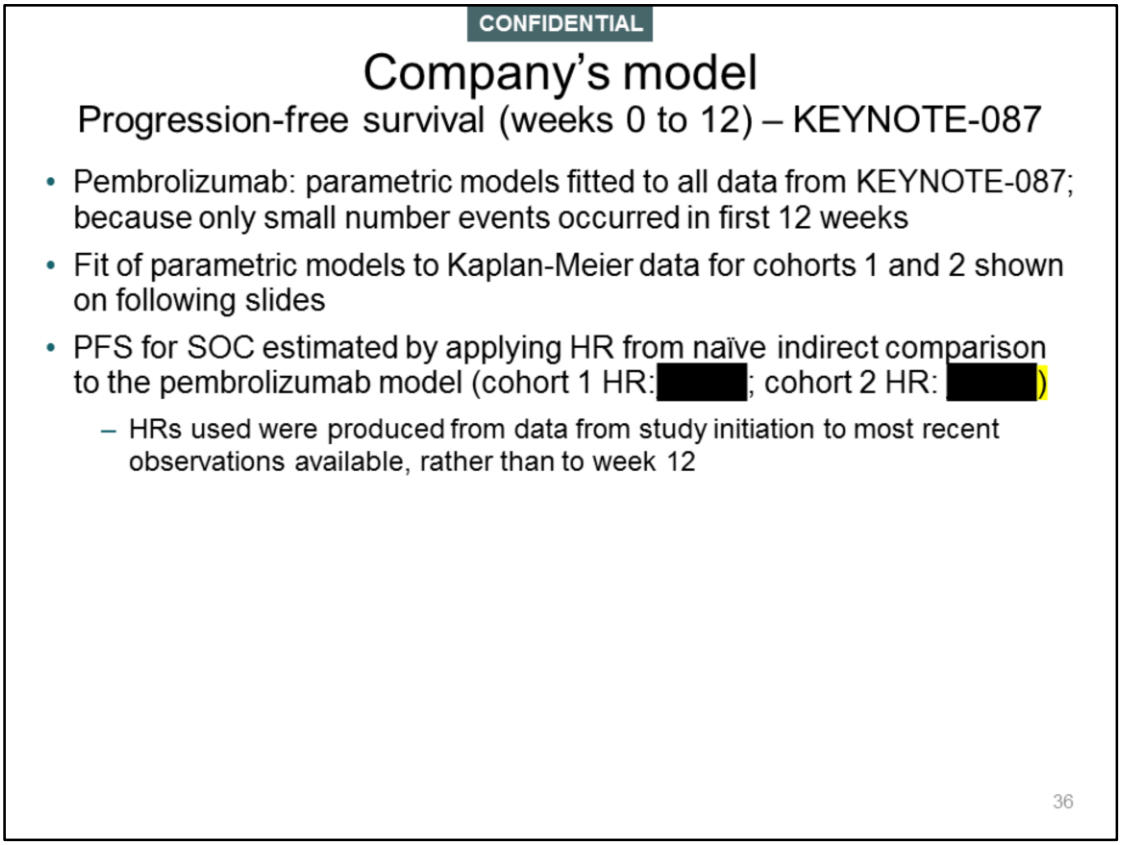
Standard of care chemotherapy was assumed to be equal usage of all regimens specified for the treatment of relapsed or refractory HL within BCSH guidelines (ASHAP, DHAOx, DHAP, ESHAP, GDP , GEM-P, GVD, ICE, IGEV, IVE, IVOx, MINE). From company submission, section 5.5.5.1 (page 202)



Source: ERG report, section 5.2.4 (pages 67 and 68)



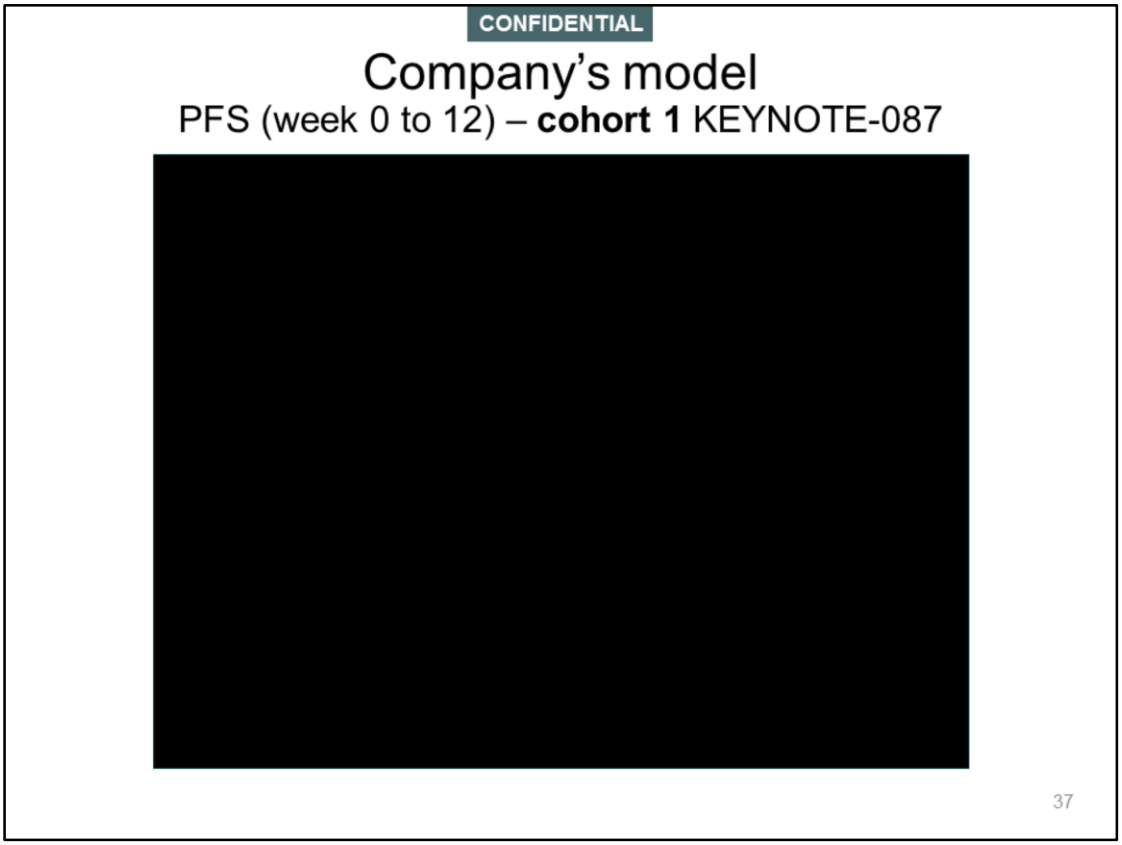
Source: ERG report, section 5.2.6 (pages 68 to 70)



Source: Company submission, section 5.3.1 (page 153)

Naïve direct comparison hazard ratios (HRs) previously described (slide 22)

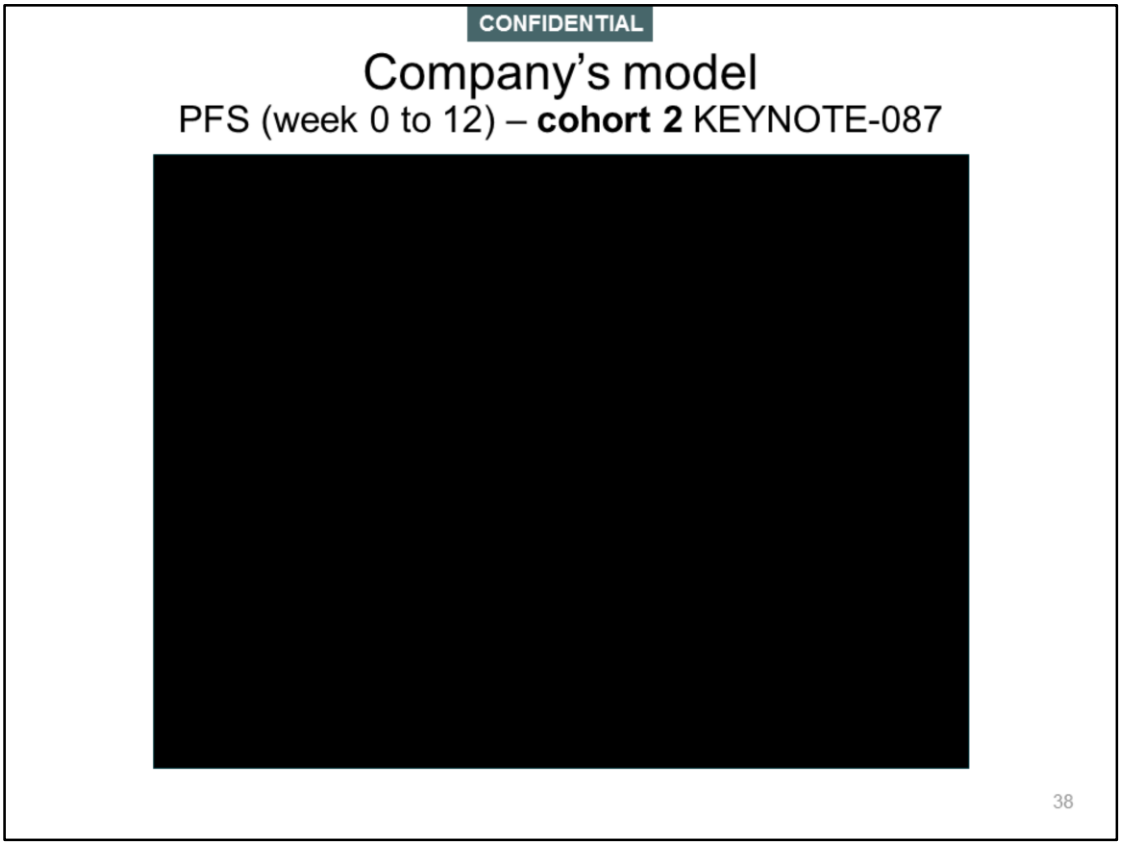
A PFS HR from week 12 to end of follow-up could not be estimated given the low number of events post week 12 observed in Cheah et al. Therefore, weeks 0 to 12 were not used to estimate the effect of treatment as it would double count patients if the week 0 to end of follow-up HR was applied after the week 0 to 12 HR. From company submission, section 5.2.2 (page 141)



Source: Company submission, section 5.3.1, figure 16 (page 154)

Summary of goodness-of-fit of the survival models for cohort 1 can be found in the company submission (table 58, page 153)

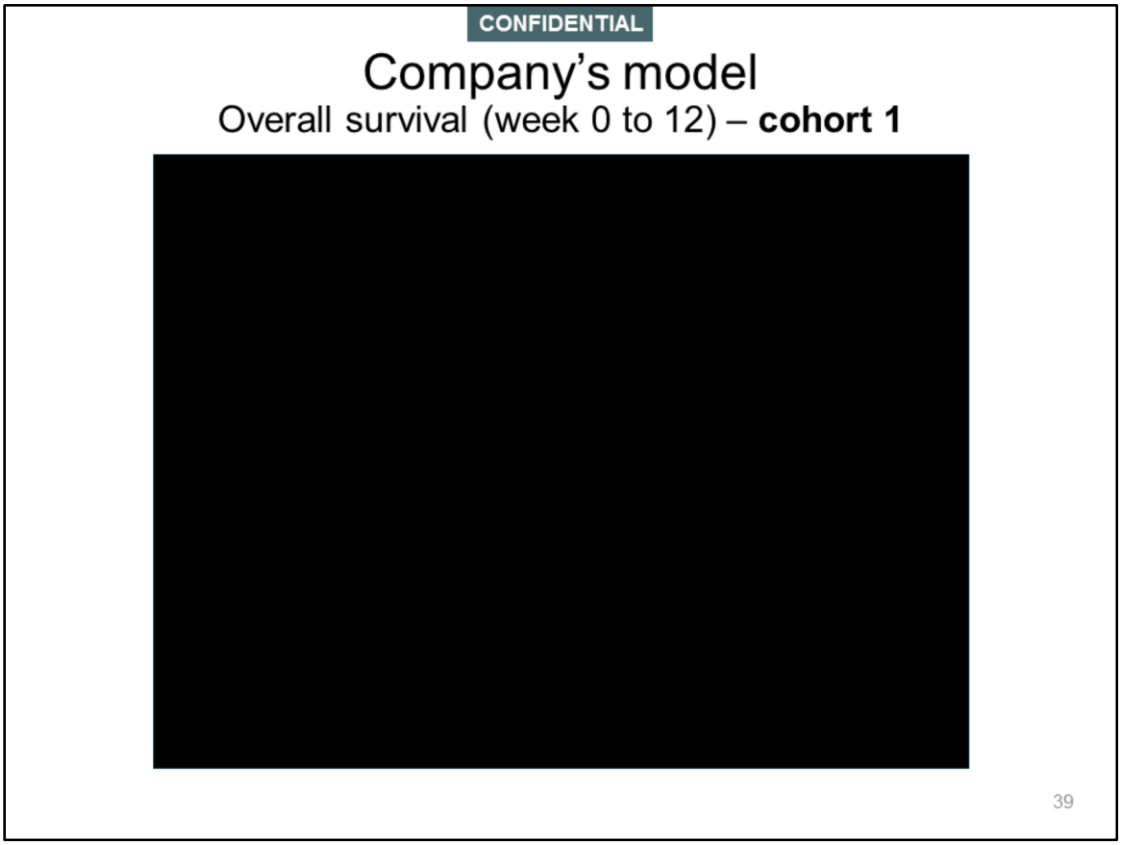
Log-logistic used in the base case model as it had the best statistical fit (lowest AIC/BIC) and predicted the most comparable rate of patients progression-free compared to the observed data at week 12. From company submission, section 5.3.1 (page 154)



Source: Company submission, section 5.3.1, figure 17 (page 156)

Summary of goodness-of-fit of the survival models for cohort 1 can be found in the company submission (table 59, page 155).

The generalised gamma had the best fit both statistically and visually compared to all other distributions, and was applied in the base case model. However, it overestimated the proportion progression-free at week 12 compared to the observed data; therefore, the Weibull was considered during scenario analysis as it predicted a lower proportion of patients progression-free at week 12 and had the third best statistical fit (AIC/BIC). From company submission (page 156)

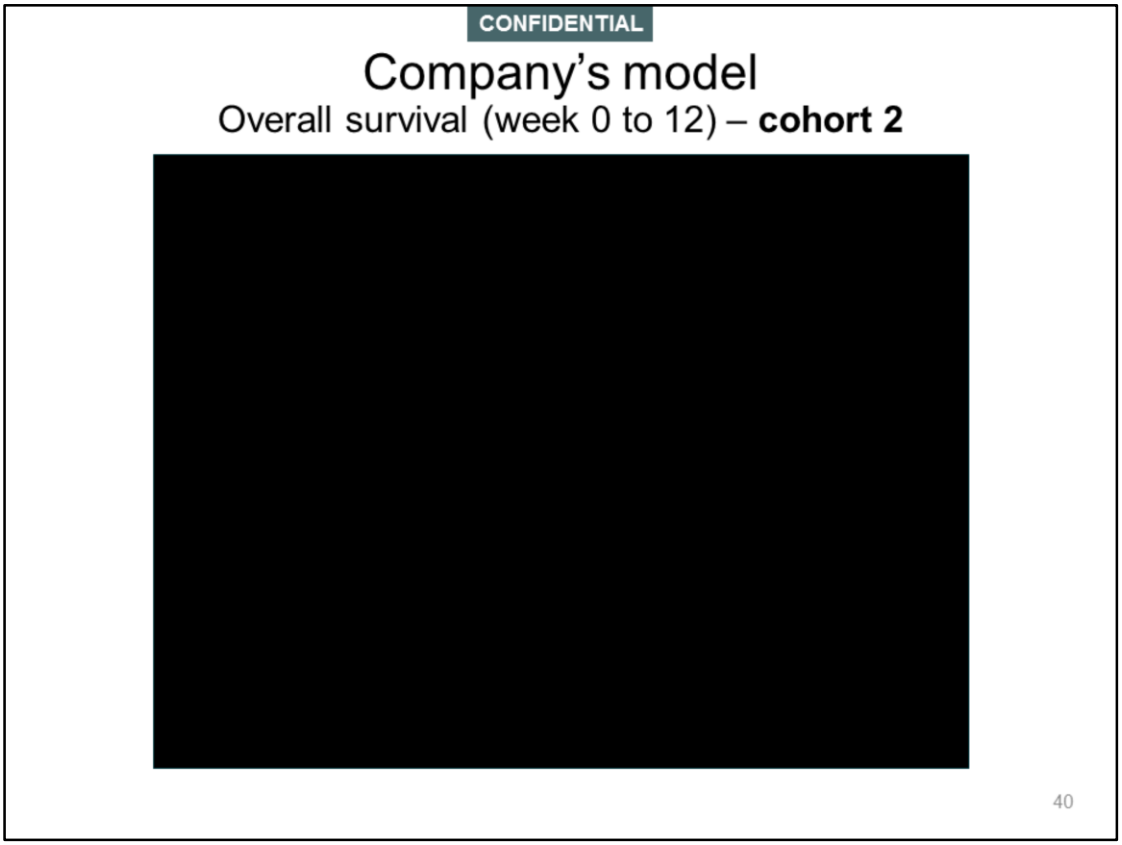


Source: Company submission, section 5.3.1, figure 18 (page 157)

Parametric models fitted to all observed data from KEYNOTE-087. Summary of the goodness of fit qualities of the survival models for cohort 1 can be found in the company submission, table 60 (page 157).

It was assumed that overall survival on SOC would be equivalent (as any HR estimated from an indirect comparison with Cheah et al. would have significant uncertainty). From company submission, section 5.3.1 (page 156).

No meaningful difference in statistical fit, visual fit or predicted number patients alive at 12 weeks. Log-normal was used in the base case (it predicted the highest rate of mortality at week 12). From company submission, section 5.3.1 (page 158).



Source: Company submission, section 5.3.1, figure 19 (page 159)

Summary of the goodness of fit qualities of the survival models for cohort 2 can be found in the company submission, table 61 (page 158)

No meaningful difference in statistical fit, visual fit or predicted number patients alive at 12 weeks. The exponential model was used in the base case (it predicted the highest rate of mortality at week 12). From company submission, page 159.

Company's model
Pre-week 12 analysis: ERG's critique

- Models were fitted to all available data from entire study period (not just first 12 weeks)
- Few events occurred in first 12 weeks; therefore fitted curves likely to have been influenced more by the post-12 week period
- ERG considered obtaining pre-12 week OS and PFS from entire study data to be questionable
  - Particularly apparent for cohort 2 progression-free survival model fit after week 11
  - Suggestion that Weibull over predicts the proportion of patients progression-free at week 12 to a lesser extent than generalised gamma not accepted by ERG

Overall survival

- Company used pre-12 week OS models that predicted highest mortality at week 12, disregarding statistical fit
- ERG preferred to use models with best statistical fit

41

Source: ERG report, section 5.2.6.1 (page 71)

Models used in ERG base-case:
Overall survival (pre-week 12)
Cohort 1: Exponential
Cohort 2: Lognormal
From ERG report, section 5.3 (page 99)

CONFIDENTIAL

Company's model

Response rates at week 12

- Cohort specific response rates at week 12 obtained from KEYNOTE-087
- SOC response rates estimated using ORs from naïve indirect comparison (CR/PR at 12 weeks in KEYNOTE-087 compared to CR/PR as best overall response in Cheah et al.):

Response	Cohort 1	Cohort 2
	Mean (SE)	Mean (SE)
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]

42

Source: Company submission, table 63 (page 160); Corrected table 62 from the company submission provided in clarification response

Results of naïve indirect comparison presented in earlier slides (slides 24 and 25)

CONFIDENTIAL

Company's model

Uptake of alloSCT at week 12 (conditional on response)

- KEYNOTE-087 data not used in model. Uptake of alloSCT was low overall ([redacted]); but was higher among UK participants ([redacted] cohort 1; [redacted] cohort 2)
- The proportion of people having alloSCT at 12 weeks, depending on their response to treatment, was estimated from clinician surveys:

Response	% people expected to receive alloSCT		
	MSD survey mean (n=16)	Alternative survey mean*	Overall mean
CR	56.79%	[redacted]	[redacted]
PR	43.93%	[redacted]	[redacted]
SD	18.36%	[redacted]	[redacted]

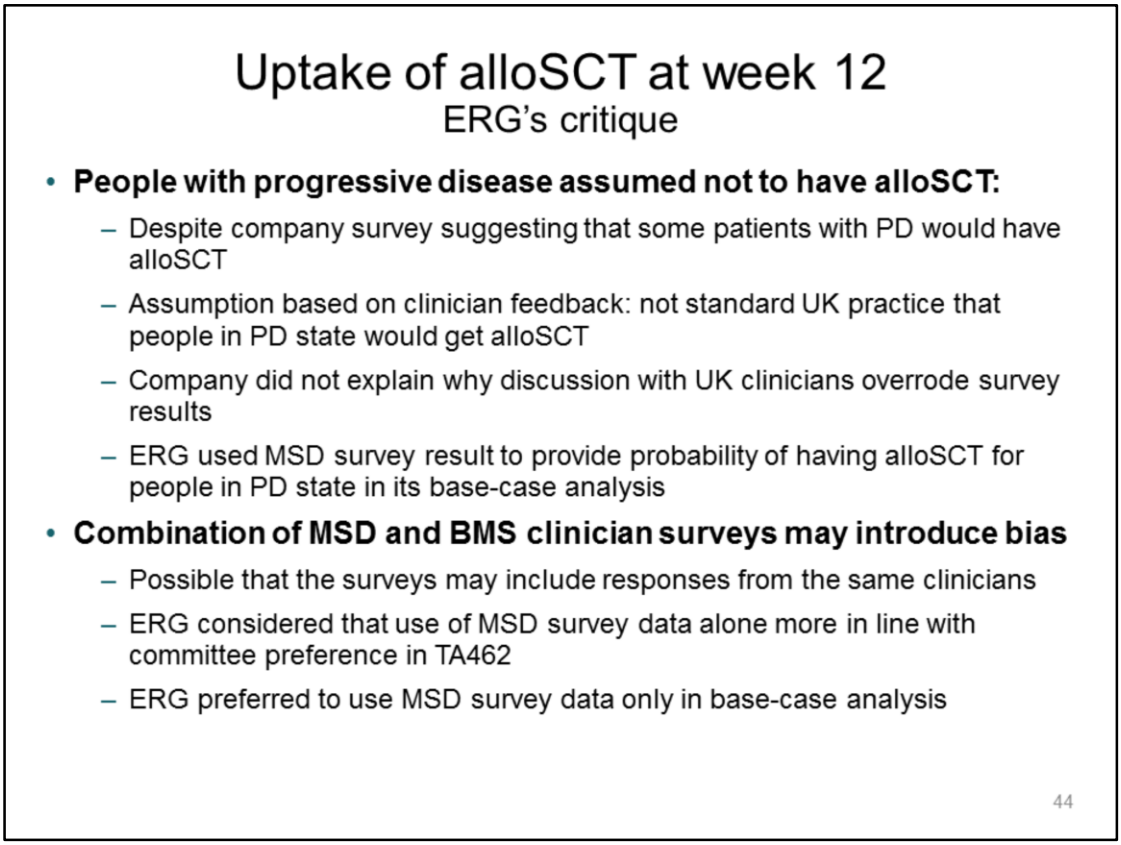
* alternative clinician survey completed by Bristol-Myers Squibb and presented marked as academic in confidence in presentation for TA462

43

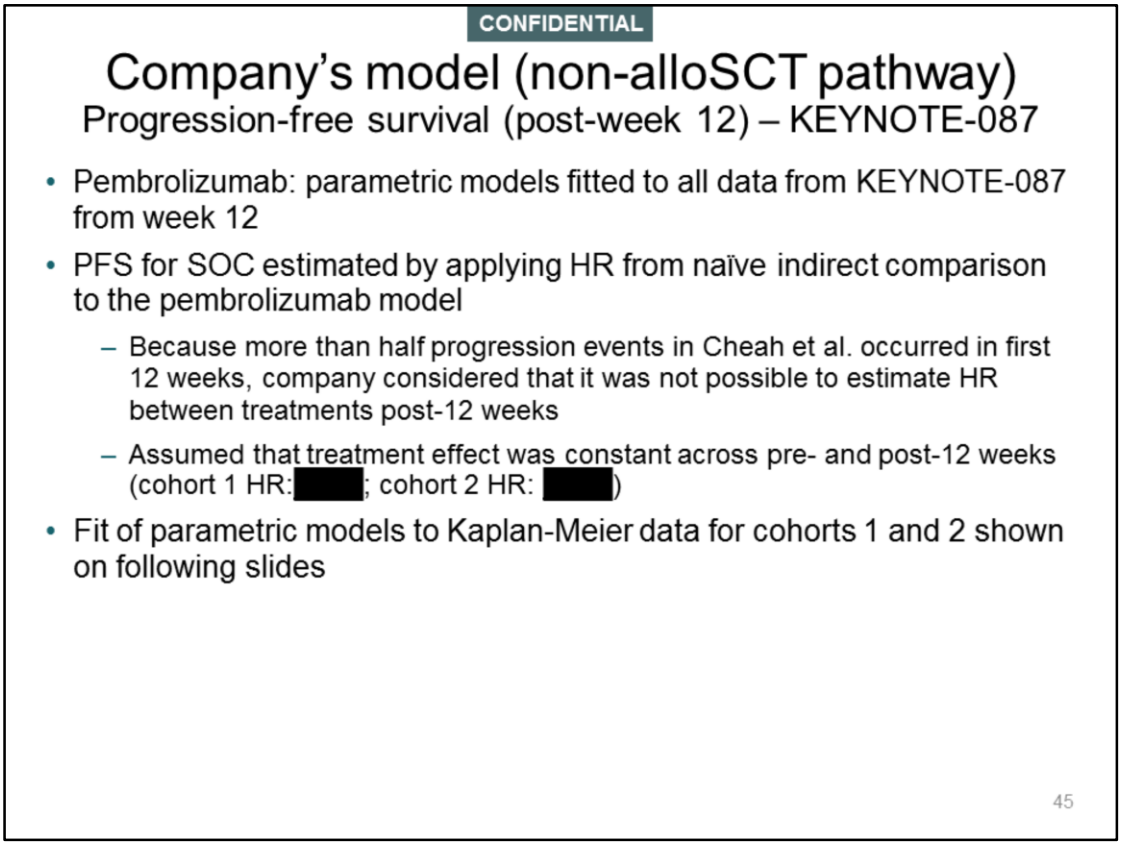
Source: Company submission, section 5.3.1 (pages 160 and 161); Response to clarification questions; ERG report, section 5.2.6 (page 73)

Further detail on the clinician survey can be found in the company submission, section 4.11.1 (pages 102 onwards)

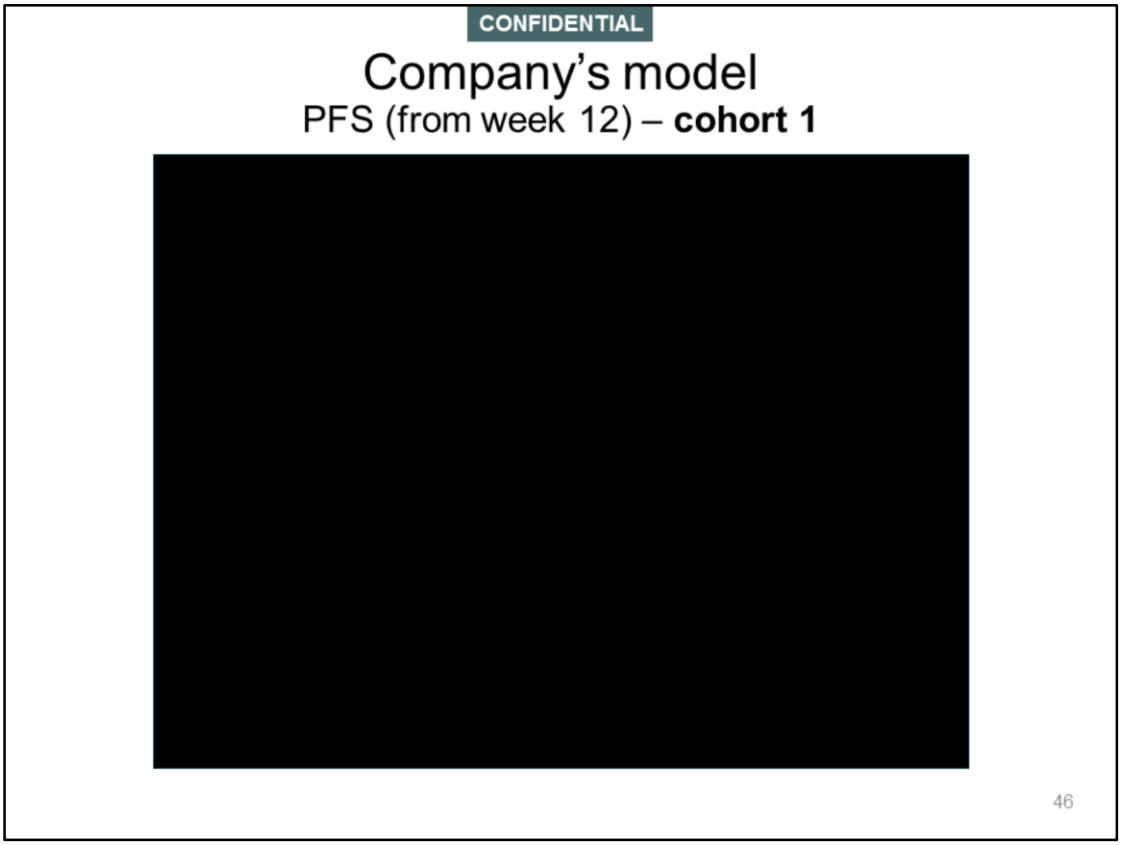
Assumed that people in a progressed disease (PD) state did not have alloSCT. The MSD clinical survey did have responses that suggested alloSCT may be done for people in PD state; however, following further discussion with UK clinicians, alloSCT was not applied in this state because it was not thought to be standard UK clinical practice. From Company submission, section 5.3.1 (page 160 and 161)



Source: ERG report, section 5.2.6 (pages 74 and 75)



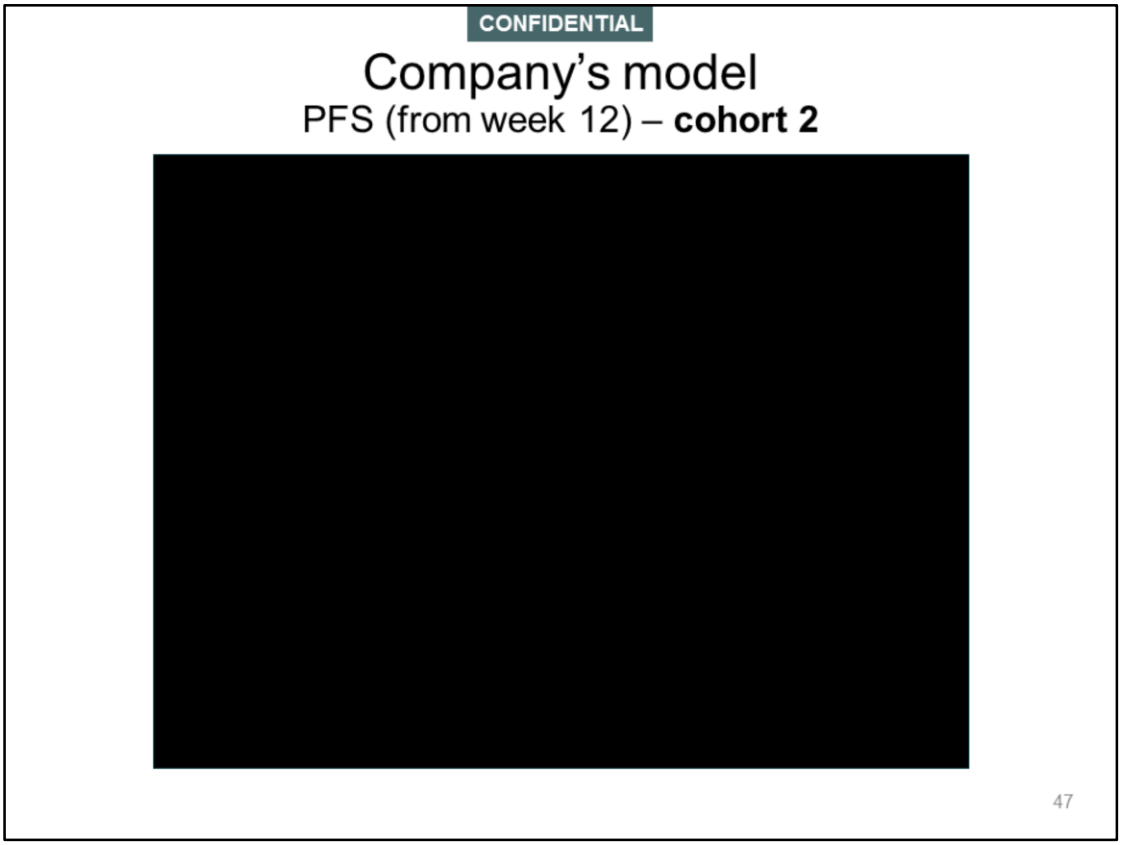
Source: Company submission, section 5.3.1 (page 162).



Source: Company submission, section 5.3.1, figure 20 (page 164)

Summary of goodness-of-fit of the survival models for cohort 1 can be found in the company submission (table 65, page 163)

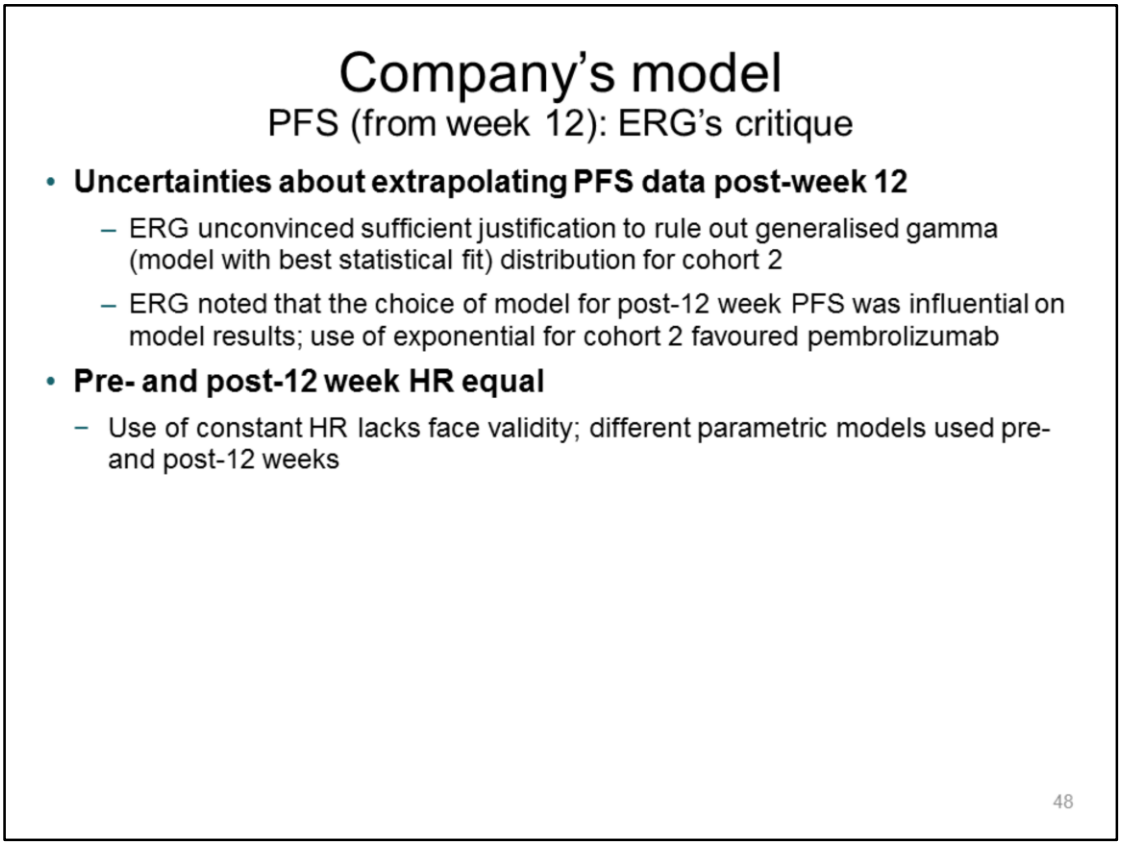
Exponential model used in base case because it had the best statistical fit, it provided the closest estimates to the median and 1-year PFS of the observed data and it followed a hazard rate over time consistent with that observed within Cheah et al. From company submission, section 5.3.1 (page 164)



Source: Company submission, figure 21 (page 167)

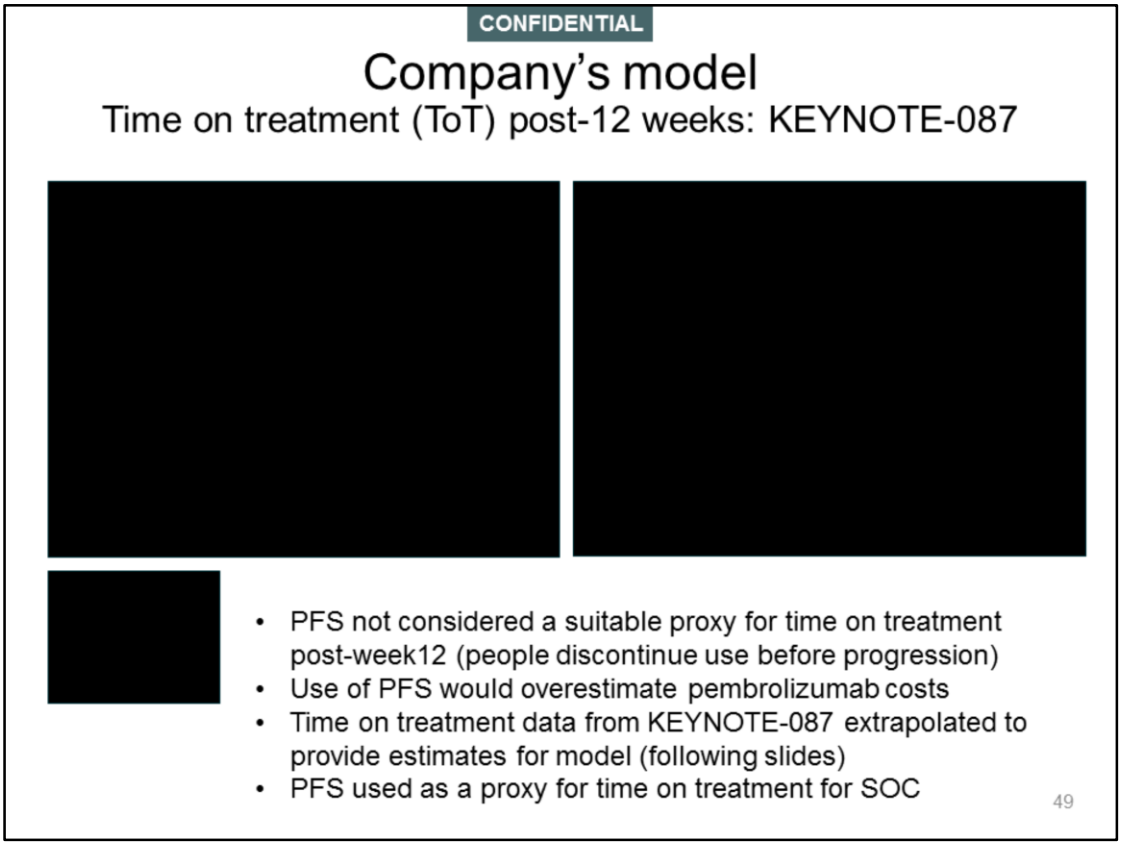
Summary of goodness-of-fit of the survival models for cohort 1 can be found in the company submission, table 66, (page 166).

Generalised gamma was best performing model according to AIC and BIC. However final drops in KM curve (from month 11) stated to be associated with considerable uncertainty because of low patient numbers (n=3). All models underestimated the median and 1-year PFS compared to KEYNOTE-087 data; particularly the generalised gamma. Therefore, despite the superior visual fit to the tail of the Kaplan-Meier data, this model was not used for analysis. Exponential model was used in base case analysis and the Gompertz was used in a scenario analysis. From company submission, section 5.3.1 (page 167).



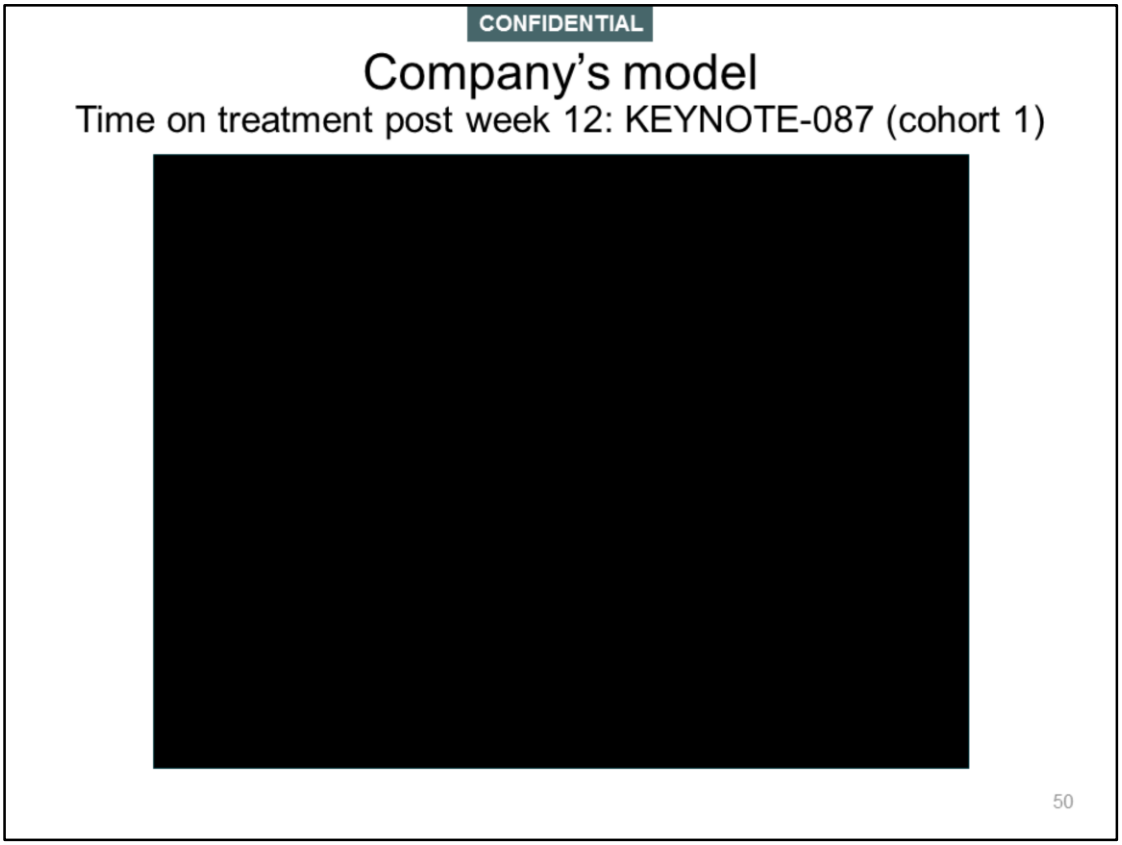
Source: ERG report, section 5.2.6 (pages 76 and 77)

Use of alternative parametric survival models for post-week 12 PFS investigated by ERG in exploratory analysis (results presented on later slides). From ERG report, section 5.3.2 (page 100)



Source: Company submission, figures 24 and 25 (page 175)

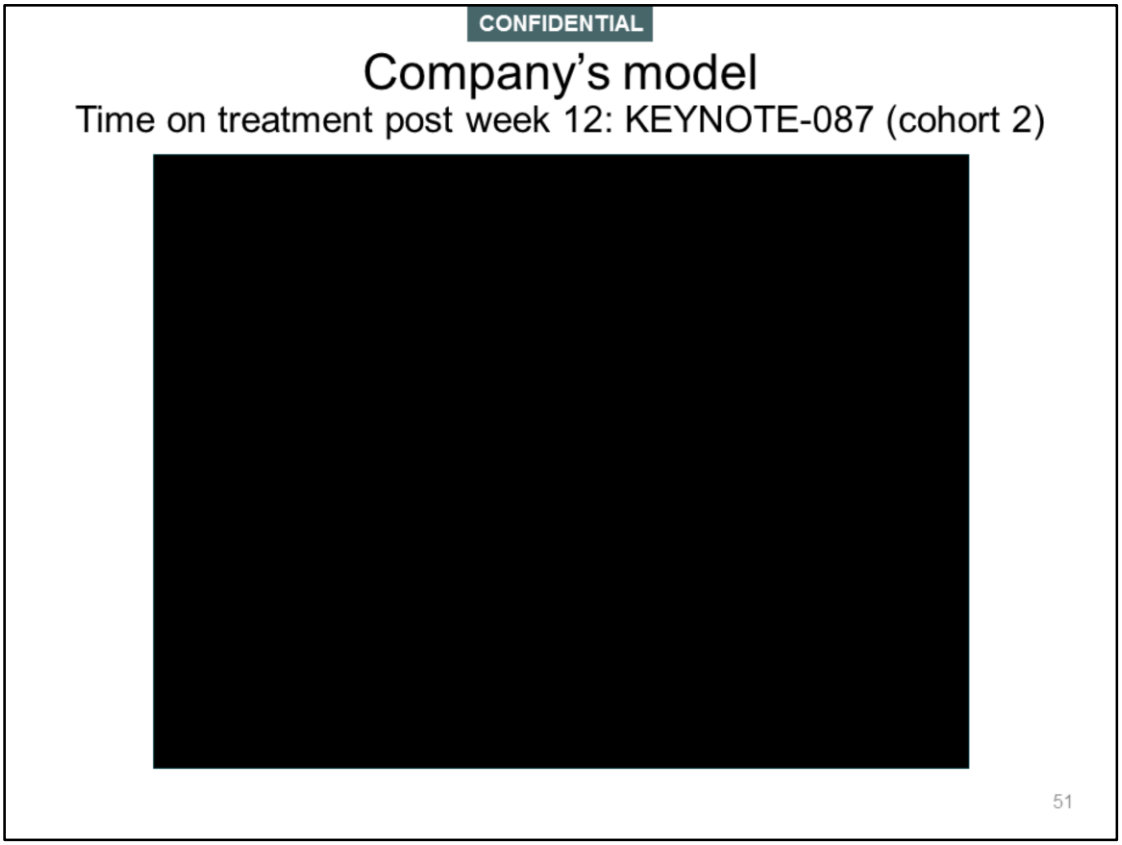
Treatment is discontinued for people receiving alloSCT in the model.
For people not having alloSCT, time to treatment discontinuation data from KEYNOTE-087 was used for post-week 12 for people having pembrolizumab (model fitting on following slides). For SOC post-week 12, progression-free survival was used as a proxy for time on treatment.
Pre-week 12, progression-free survival was used as a proxy for time to treatment discontinuation for both SOC and pembrolizumab. From ERG report, section 5.2.6.6 (page 80)



Source: Company submission, section 5.3.1, figure 26 (page 178)

Summary of goodness-of-fit of the models for cohort 1 can be found in the company submission (table 70, page 177)

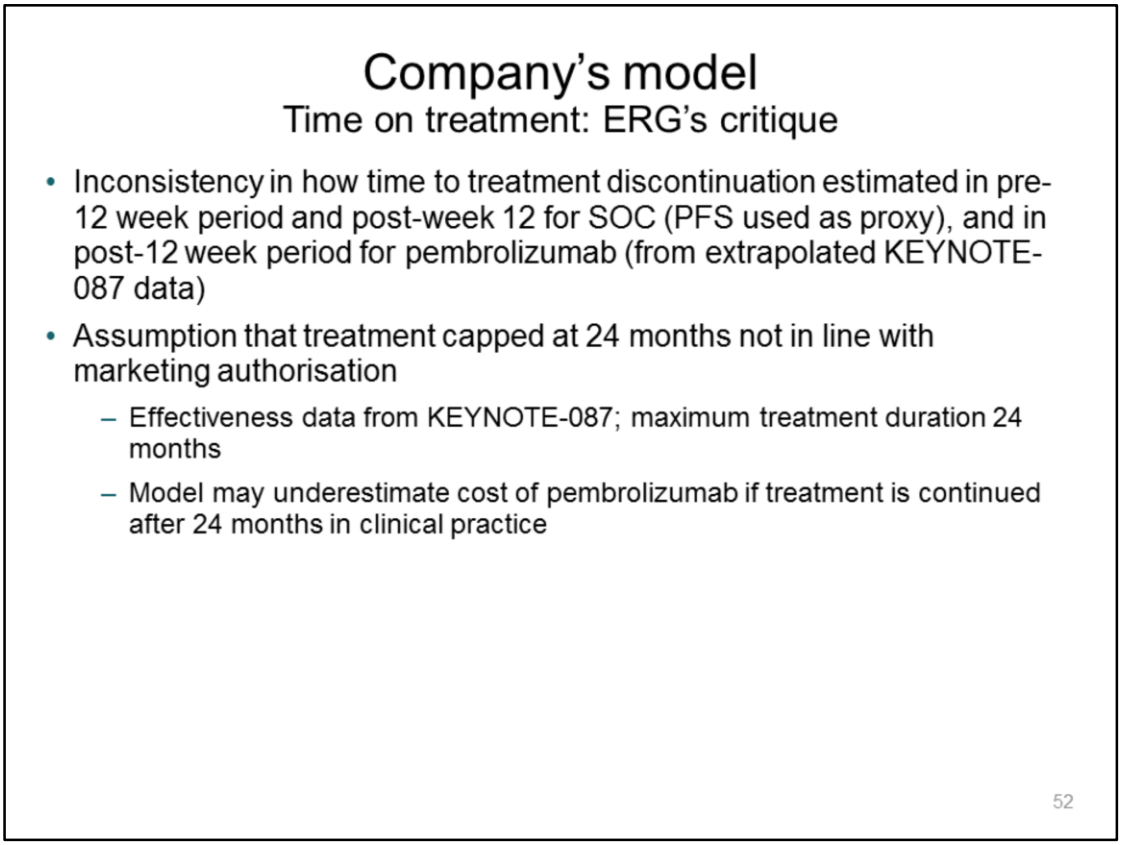
The models had similar medians, restricted means, statistical and visual fits to the KEYNOTE-087 data. The exponential model was used for the base case analysis because it had the best statistical fit and this model has been used for the base case PFS. From company submission, section 5.3.1 (page 178)



Source: Company submission, section 5.3.1, figure 27 (page 180)

Summary of goodness-of-fit of the models for cohort 2 can be found in the company submission (table 71, page 179)

All parametric models (except log-normal) had similar statistical fits to the observed data. The exponential model was used in the base-case analysis for consistency with the base-case PFS distribution. From company submission, section 5.3.1 (page 180)



Source: ERG report, section 5.2.6.6 (page 80)

Company's model

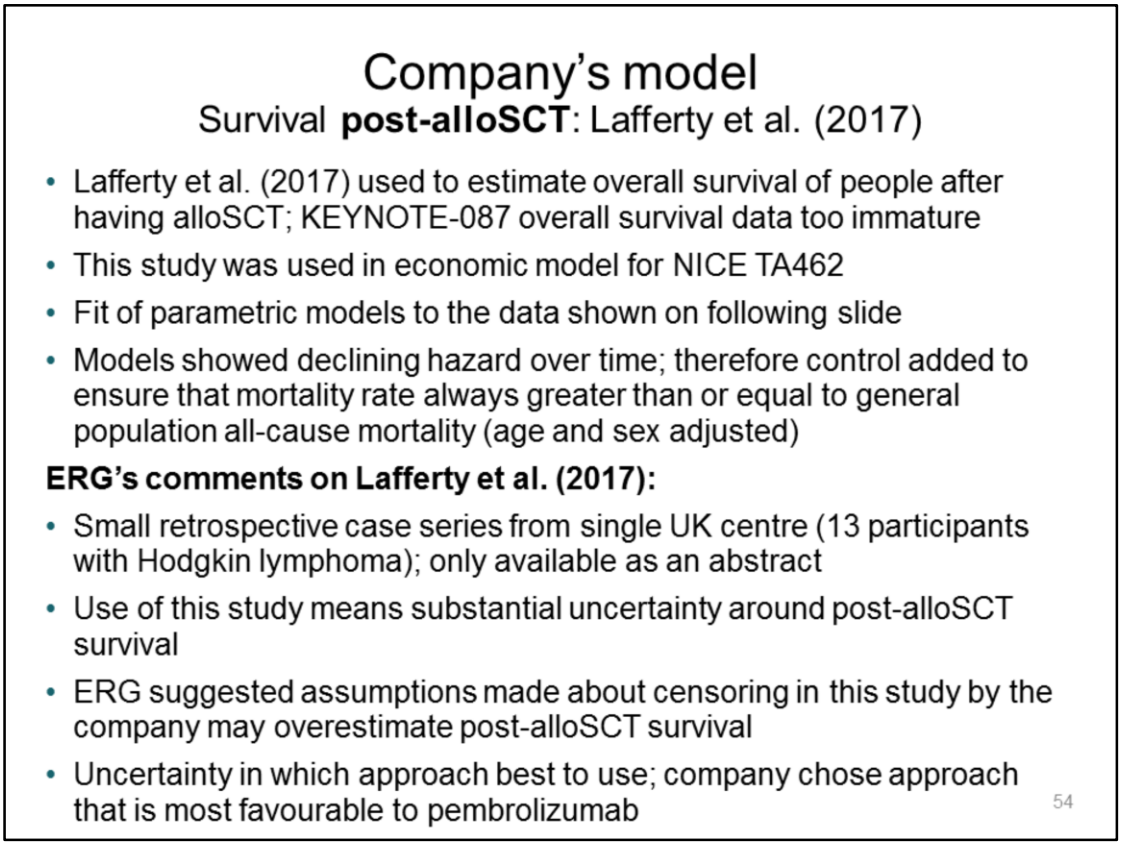
Mortality (pre- and post-progression) in **non-alloSCT pathway**

- **Mortality pre-progression (for people who have not had alloSCT):**
  - Limited survival data available from KEYNOTE-087 for participants who have not progressed.
  - Mortality rates for general population (adjusted for age and sex) used in model
  - **ERG comment:** Uncertain about impact of this assumption on model outcomes
- **Post-progression survival (for people who have not had alloSCT):**
  - Number of participants who have progressed in KEYNOTE-087 judged too small to estimate robust post-progression survival
    - ERG agreed that KEYNOTE data too immature to be used in present analysis (based on provided scenario analysis using KEYNOTE data)
  - Cheah et al. (2016) used as most appropriate data source
  - No post-progression survival benefit for pembrolizumab assumed

53

Source: Company submission (pages 168 to 170); ERG report, section 5.2.6 (page 77)

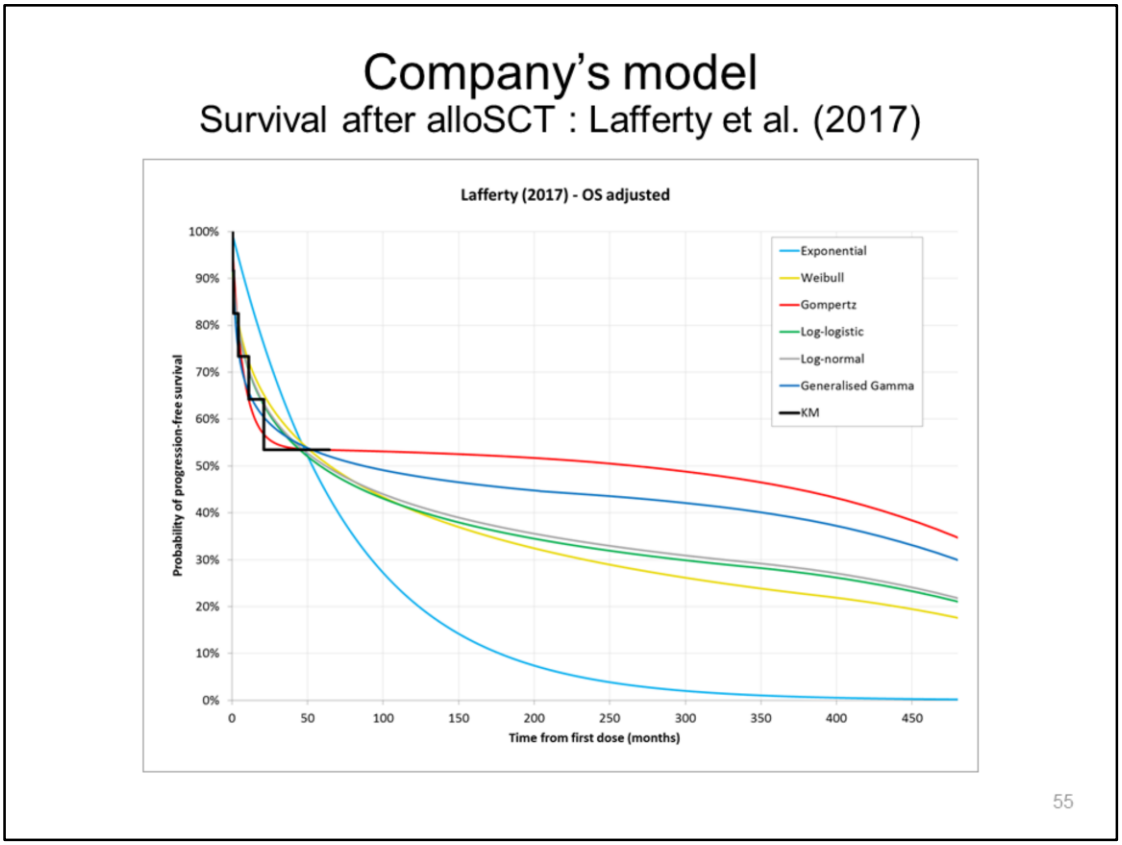
The ERG highlighted inconsistency in the choice of data sources for survival post-week 12, which was justified by the company because KEYNOTE-087 overall survival data are too immature. From ERG report, section 5.2.6 (page 70). This relates to both mortality in the non-alloSCT pathway (this slide) and the alloSCT pathway (next slide)



Source: Company submission, section 5.3.1 (page 170); ERG report, section 4.2.3 (page 44); section 5.2.6.4 (pages 77 to 79)

Figure 5.4 in the ERG report (page 79) compares the ERG's and company's approach to estimating post-alloSCT overall survival based on Lafferty et al. data

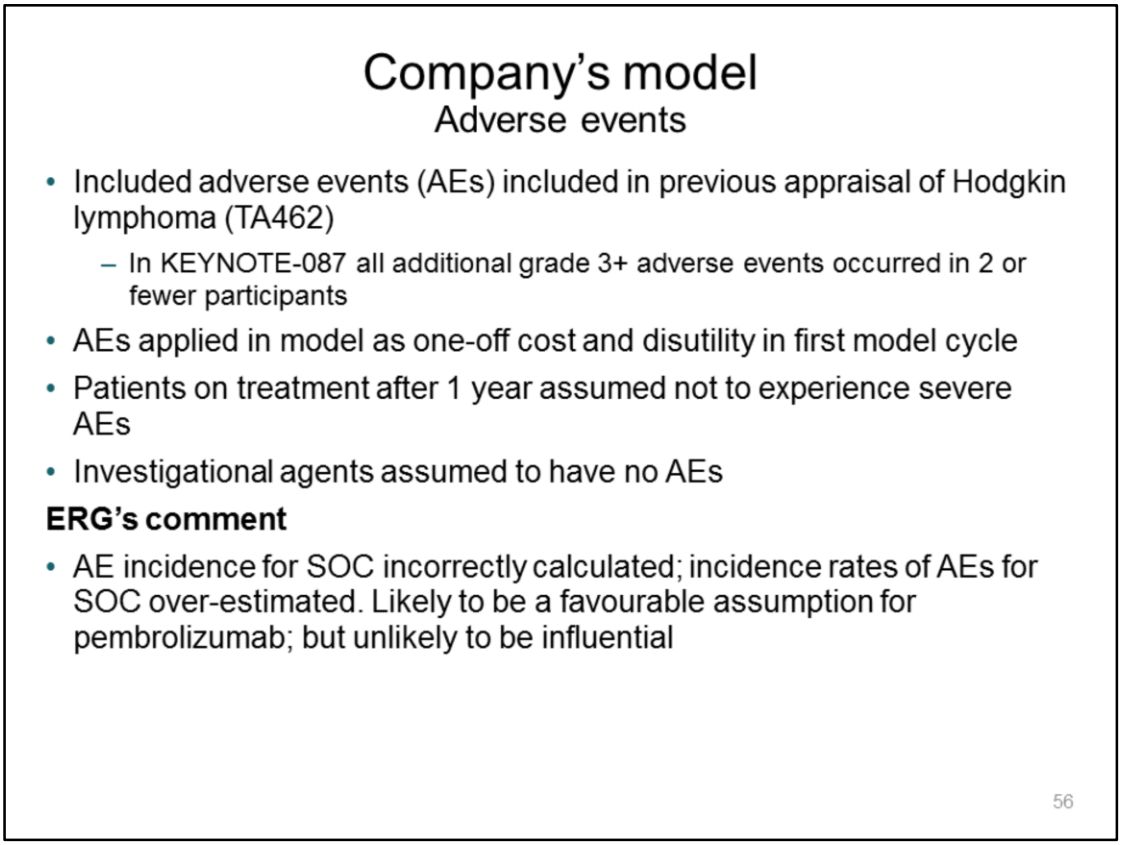
NICE TA462: Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma (2017)



Source: Company submission, figure 23 (page 174): OS after alloSCT adjusted for all-cause mortality extrapolations

Summary of goodness-of-fit of the survival models can be found in the company submission table 68 (page 171) and table 69 (page 173)

Weibull distribution used in the company's base case despite not having the best statistical fit (it had the 5th best fit), because (1) generalised gamma (which had the best fit) predicted infinite survival beyond 150 months and had to be adjusted, (2) there were only small difference in the AIC/BIC scores, (3) the ERG in NICE appraisal TA462 considered the log-normal and Weibull as most clinically plausible and (4) Weibull was suggested to be a conservative option (lowest mean survival and percentage alive at 40 years). The log-normal distribution was used in a scenario analysis. From ERG report, section 5.2.6.4 (pages 77 and 78) and company submission (page 174)



Source: Company submission, section 5.3.5 (page 182); ERG report, section 5.2.7 (pages 80 and 81)

Further detail on incidence of adverse events included in model can be found in tables 73 and 74 (page 184) of the company submission

CONFIDENTIAL

Company's model

Health state utility values

Health state		Utility value
Progression-free	Pembrolizumab (cohort 1)	[redacted]
	Pembrolizumab (cohort 2)	[redacted]
	Standard of care	[redacted]
Progressed		[redacted]

Progression-free

- Calculated from utility values from KEYNOTE-087 at week 12
- Response specific utility values (CR/PR/SD) weighted by response rates from KEYNOTE-087 (for pembrolizumab) or Cheah et al. (for standard of care) to calculate progression-free utility value

Progressed disease

- PD utility from KEYNOTE-087 ([redacted]) not used. Company stated that week 12 utility observation may not capture longer-term disutility associated with progression
- Value calculated based on utility decrement (0.33) between SD and PD in Swinburn et al. (2015); applied to KEYNOTE-087 SD utility

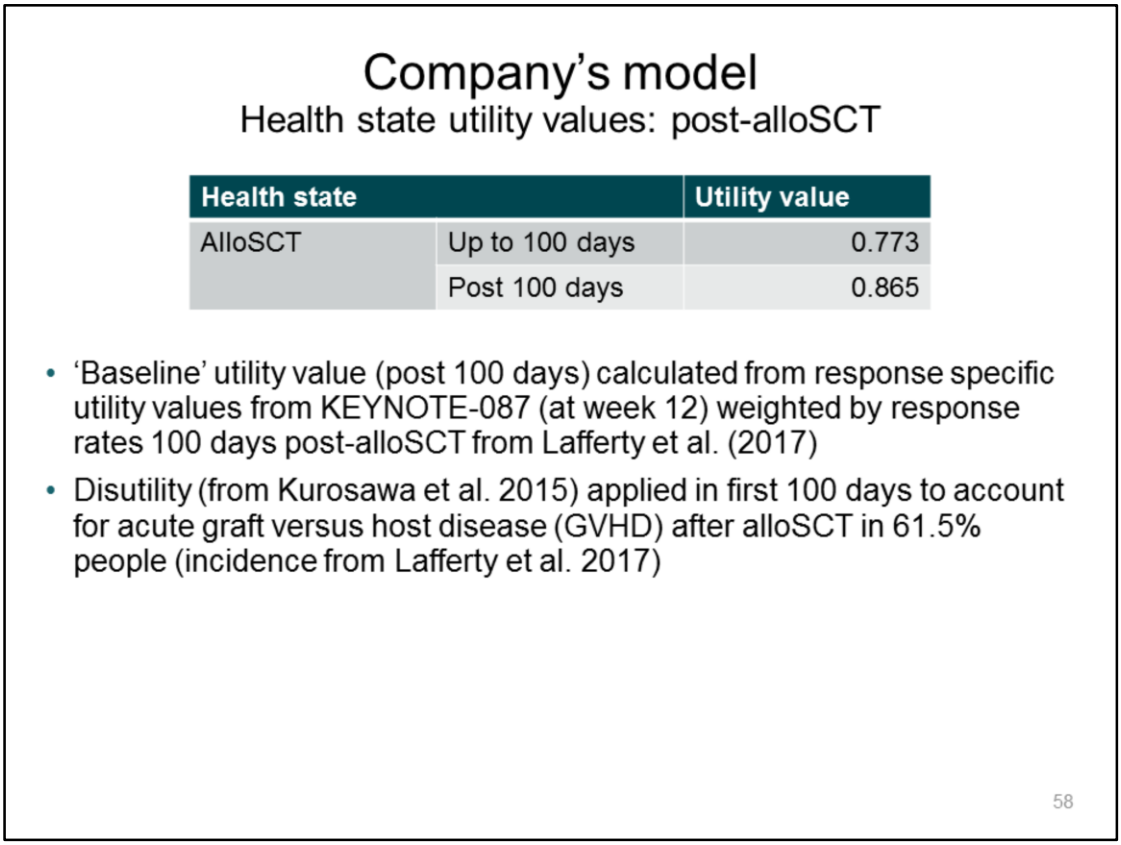
57

Source: Company submission, section 5.4.7, table 80 (page 193); ERG report, section 5.2.8 (page 81)

Swinburn et al. (2015) Health utilities in relation to treatment response and AEs in RRcHL and systemic anaplastic large cell Lymphoma

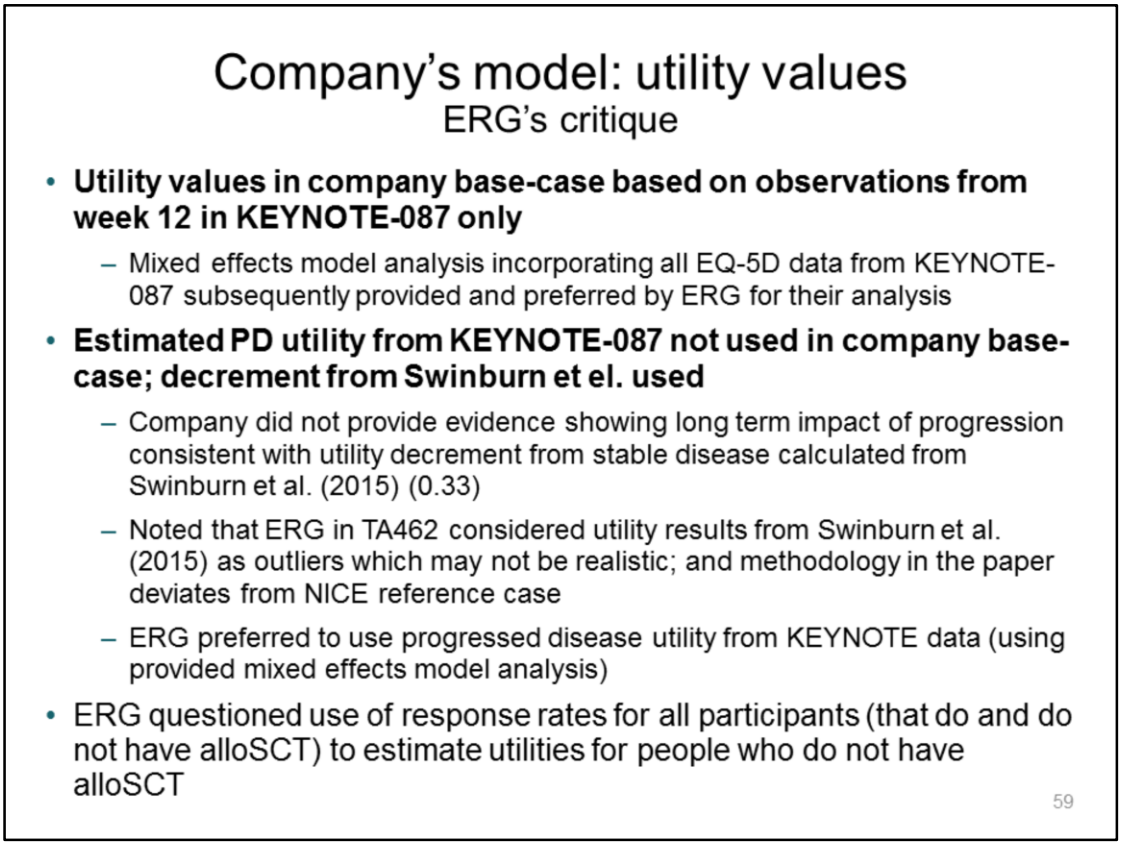
Disutility from adverse events were also applied in the model. Details can be found in the company submission, section 5.4.6 (starts page 189). Overview of adverse event disutilities presented in ERG report, section 5.2.8, table 5.12 (page 82)

Age related utility decrements were applied in all health states. Details can be found in the company submission, section 5.4.8 (page 194)



Source: Company submission, section 5.4.6.1 (page 192), section 5.4.8 (pages 193 and 194)

Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. Bone Marrow Transplant. 2015;50(9):1241-9



Source: ERG report, section 5.2.8 (pages 81 to 83)

Detail on utilities estimated from mixed effects model using all observed EQ-5D data from KEYNOTE-087 can be found in the ERG report, section 5.2.8 (page 83)

CONFIDENTIAL

Company's model: utility values

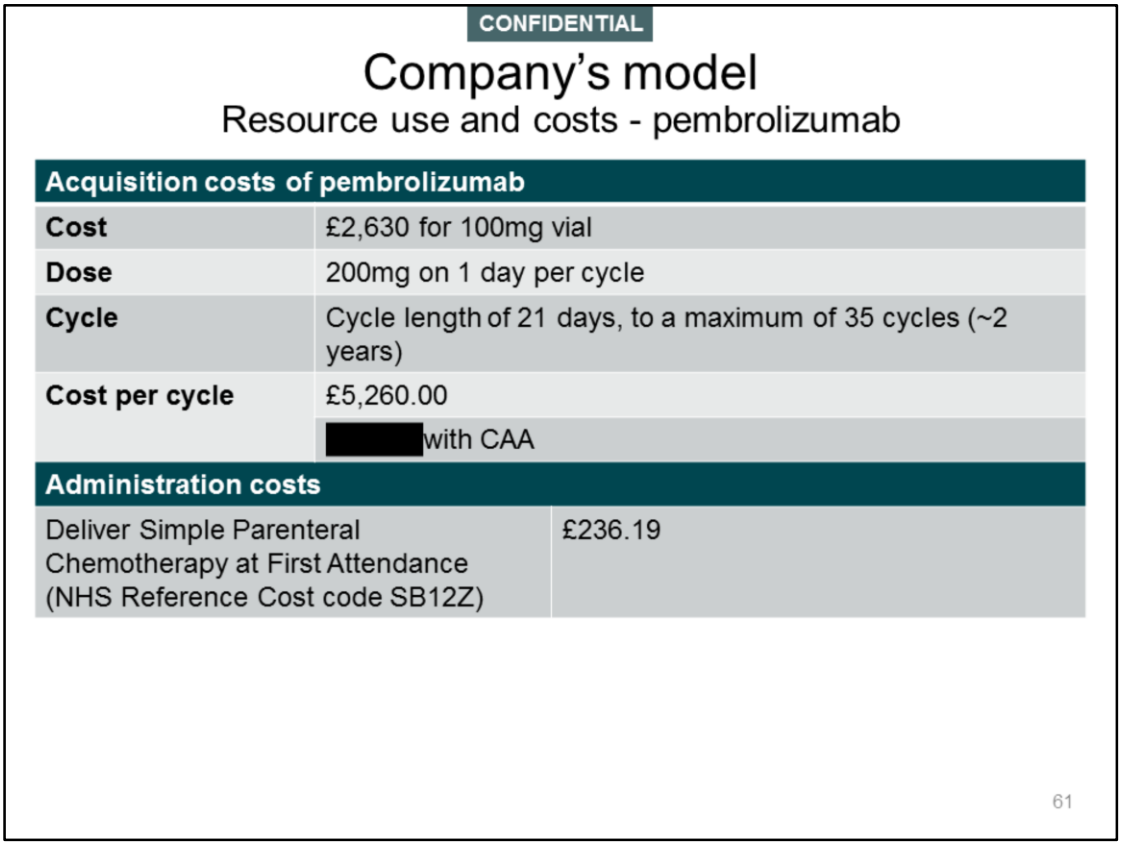
ERG's critique (cont.)

- ERG preferred to use utility data from Kurosawa et al. (rather than the disutility from this study applied to KEYNOTE derived utility) to account for GVHD in first 100 days post-alloSCT
- Overview of utilities used in company and ERG base-case analyses:

Health state		Company base-case	ERG base-case
Progression-free (first 12 weeks)	Pembrolizumab cohort 1	[redacted]	[redacted]
	Pembrolizumab cohort 2	[redacted]	[redacted]
	SOC	[redacted]	[redacted]
Progression-free (after first 12 weeks; no alloSCT)	Pembrolizumab cohort 1	[redacted]	[redacted]
	Pembrolizumab cohort 2	[redacted]	[redacted]
	SOC	[redacted]	[redacted]
Progressive disease		[redacted]	[redacted]
Post-alloSCT (first 100 days)		0.773	0.708
Post-alloSCT (post 100 days)		0.865	0.800

60

Source: ERG report, section 5.2.8, table 5.14 (pages 84 and 85)



Source: Company submission, section 5.5.5.1 (pages 199 to 201)

Company's model
Resource use and costs - SOC

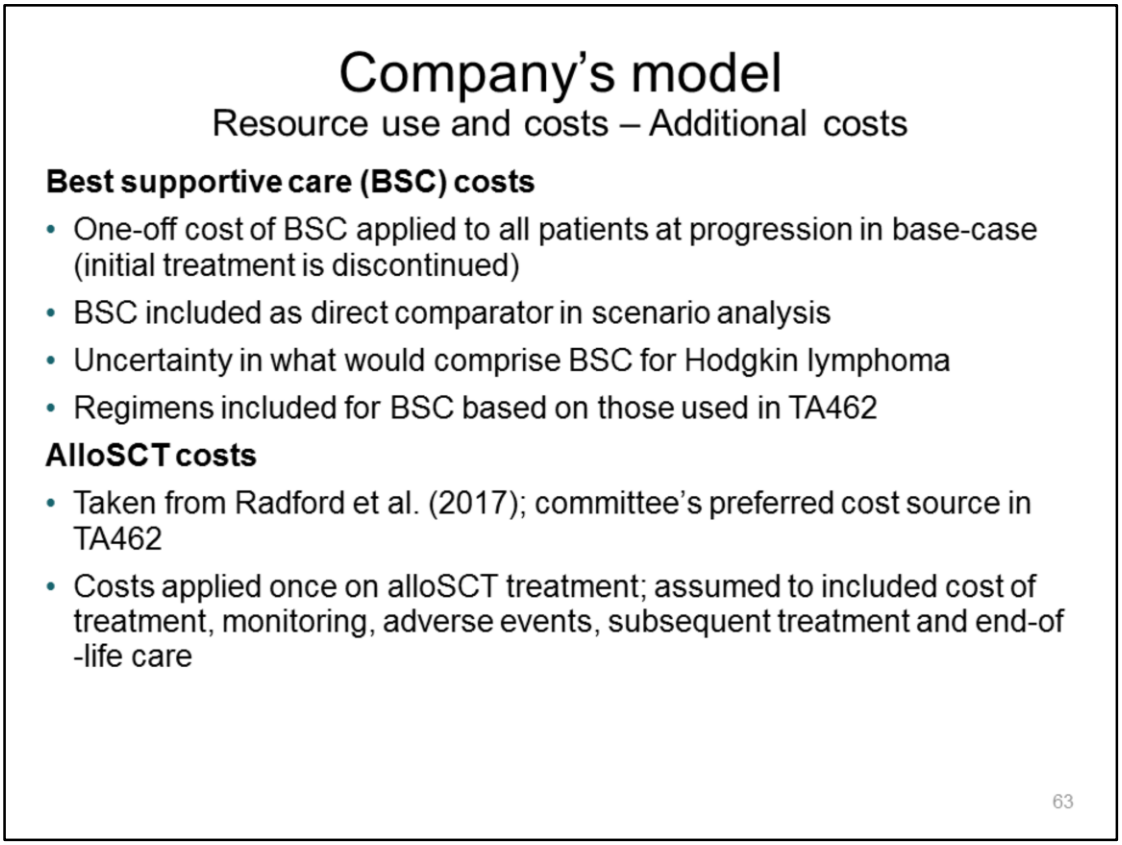
Component of SOC	Percentage*	Acquisition cost/per cycle	Administration cost/per cycle
Chemotherapy *Regimens: ASHAP, DHAOx, DHAP, ESHAP, GDP , GEM-P, GVD, ICE, IGEV, IVE, IVOx, MINE*	38.5% (equally split per regimen; i.e. 3.2% each)	Varies between regimens (from £63.32 to £2,183) Cycle length also varies between regimens	Varies between regimens (from £383.13 to £1,367.43)
Bendamustine	18.5%	£123.30	£383.13
Investigational agents	43.1%	Acquisition and administration costs of investigational agents were assumed to be £0	

* Proportions obtained from treatments in Cheah et al. (excluding BV-retreatment, autoSCT and Other) and assuming all patients not treated with bendamustine or investigational agents were distributed equally between the chemotherapy regimens.

62

Source: Company submission, section 5.5.5.1 (page 202 onwards) ; ERG report, section 5.2.9 (page 85)

Further details of SOC costs (tables 91 and 92; pages 206 and 207) and cycle lengths of chemotherapy treatments can be found in the company submission, section 5.5.5.1 (page 202 onwards) and in the ERG report, section 5.2.9 (page 85)



Source: Company submission, section 5.5.6 (page 208); section 5.5.8 (page 214); ERG report, section 5.2.9 (page ~~48~~ 85 – in ~~draft~~ report)

Further detail can be found in company submission (section 5.5.6) and the ERG report, section 5.2.9, table 5.16 (page 86)

Details on costs applied for adverse events and terminal care can be found in the company submission, section 5.5.6 (pages 211 to 212) and section 5.5.7 (pages 212 to 214) – and in the ERG report, section 5.2.9 (pages 88 and 89)

Radford et al. (2017) Treatment pathways and resource use associated with recurrent Hodgkin lymphoma after autologous stem cell transplantation. Bone Marrow Transplant; 52(3):452-4

Company's model: Resource use and costs
ERG's critique

- **All chemotherapy regimens assumed to contribute equally to SOC**
  - ERG cite previous report for TA462 which suggests lower price chemotherapy regimens are most commonly used in this population; and suggest that SOC costs are likely to be overestimated
- **One-off cost applied on alloSCT treatment**
  - ERG considered alloSCT costs to be under-estimated
  - In TA462, one-off cost was only applied in scenario analysis; monthly costs for subsequent treatment and monitoring were applied
  - ERG have applied monitoring costs over life time horizon in their analysis

64

Source: ERG report, section 5.2.9 (pages 87 to 89)

Company's base case results

Deterministic (with CAA)

Treatment		Total		Incremental		ICER
		Costs	QALYs	Costs	QALYs	
Cohort 1	SOC	£52,017	3.223	-	-	
	Pembrolizumab	£107,459	4.497	£55,442	1.274	**£43,511**
Cohort 2	SOC	£51,424	3.200	-	-	
	Pembrolizumab	£93,732	4.072	£42,308	0.871	**£48,571**

ERG's comments

- Main benefit of pembrolizumab from QALY gains after week 12 for people who have alloSCT
  - Accounts for 71% (cohort 1) and 78% (cohort 2) incremental QALYs
- BSC not included as comparator in base case; therefore pembrolizumab could not be compared to all relevant alternatives at the same time

65

Source: Company submission, section 5.7.2 (page 219); ERG report, section 5.2.10 (page 89)

A 'corrected base-case' model was provided by the company to replace their initial submitted model (which contained an error)

Breakdown of QALYs generated by pembrolizumab and SOC in company's base-case model can be found in the ERG report, section 5.2.10, tables 5.19 and 5.20 (pages 89 and 90)

Company's sensitivity analyses
Probabilistic and deterministic (with CAA)

Probabilistic sensitivity analysis

Treatment	ICER (versus SOC)	Probability of cost-effectiveness of pembrolizumab compared with SOC		
		Maximum acceptable ICER		
		£20,000/ QALY	£30,000/ QALY	£50,000/ QALY
Pembrolizumab – cohort 1	£43,653	1.1%	20.5%	60.1%
Pembrolizumab – cohort 2	£50,894	1.4%	16.1%	50.4%

Deterministic sensitivity analysis

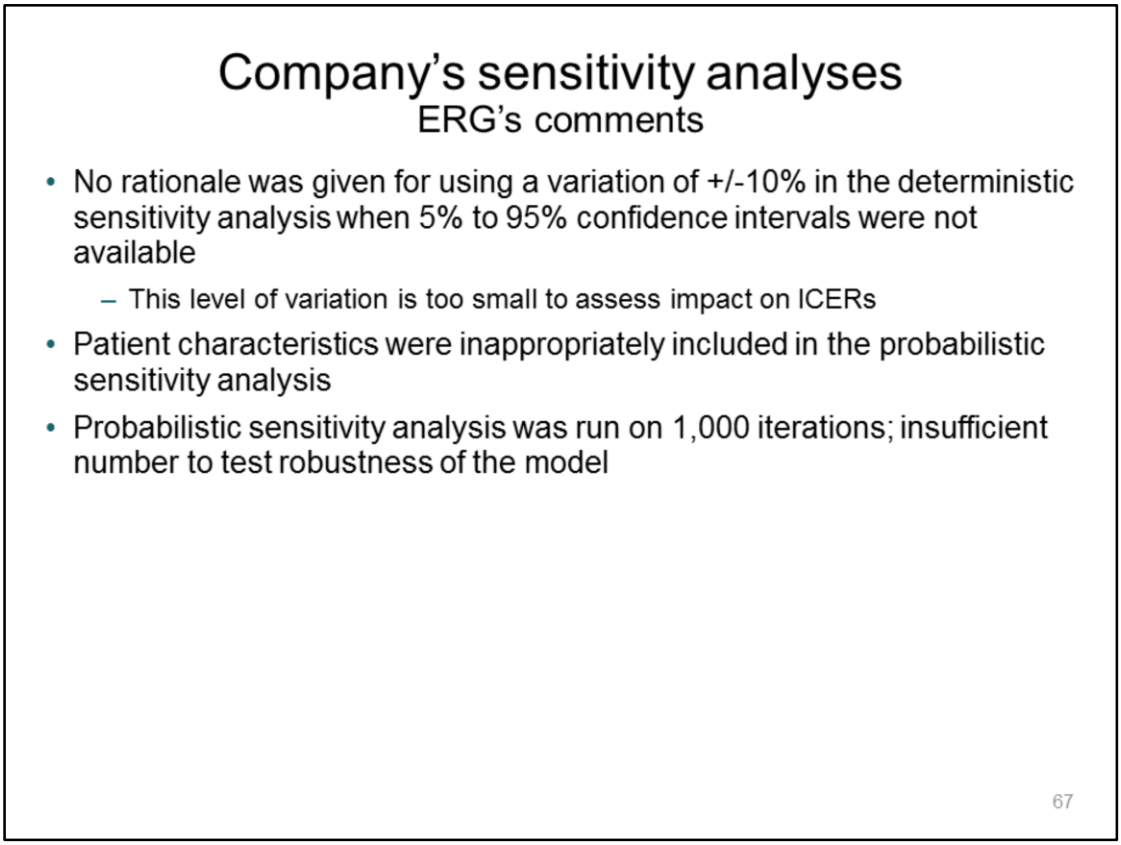
- Most influential model inputs: discount rate applied to outcomes, odds ratios applied to CR and PR at week 12
- In most scenarios ICER for pembrolizumab versus SOC was below £50,000/QALY

66

Source: Company submission, section 5.8.1 (pages 229 to 231); section 5.8.2 (pages 232 to 234); ERG report, section 5.2.11 (page 90)

Cost-effectiveness acceptability curves can be found in the company submission, section 5.8.1, figures 36 and 37 (page 231)

Tornado diagrams presenting results of deterministic sensitivity analysis can be found in the company submission, section 5.8.2 (page 234)



Source: ERG report, section 5.2.11 (page 94)

Company's scenario analyses (with CAA) (1)

Scenario	Cohort	ICER
		Pembrolizumab versus SOC
Scenario 1 BSC as comparator	Cohort 1	£44,161
	Cohort 2	£49,387
Scenario 2a 100% people with CR, PR or SD response at week 12 have alloSCT	Cohort 1	£23,564
	Cohort 2	£24,492
Scenario 2b Proportion of people with PR response at week 12 who have alloSCT taken from MSD survey	Cohort 1	£47,957
	Cohort 2	£56,677
Scenario 3 Values from MAIC (instead of naïve indirect comparison) used in model	Cohort 1	£36,423
	Cohort 2	£41,087

68

Source: Company submission, section 5.8.3 (pages 235 and 236)

Company's scenario analyses (with CAA) (2)

Scenario	Cohort	ICER
		Pembrolizumab versus SOC
Scenario 4a Weibull model used for PFS (weeks 0 to 12) in cohort 2	Cohort 2	£47,410
Scenario 4b Gompertz model used for PFS (week 12 onwards) in cohort 2	Cohort 2	£52,562
Scenario 4c Lognormal model fitted to post-alloSCT survival data from Lafferty et al.	Cohort 1	£42,075
	Cohort 2	£46,812
Scenario 5 Time horizon of 50 years	Cohort 1	£42,651
	Cohort 2	£47,516

69

Source: Company submission, section 5.8.3 (pages 235 and 236)

ERG's comments: Model validation

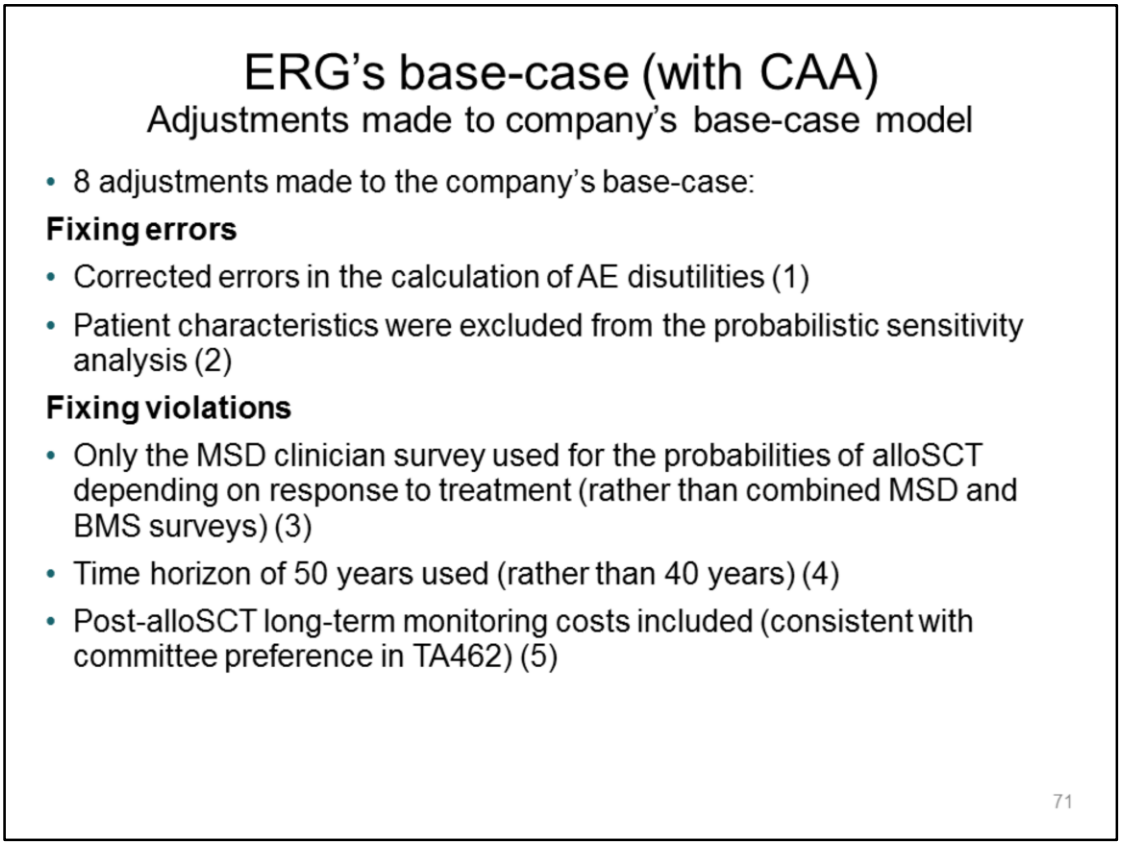
- ERG was unable to reproduce figures from a table in the company submission that compared model outcomes with those from trials
- Patients with no evaluated response were assumed to have stable disease; this probably leads to an overestimation of patients in this state
- No cross validation of model assumptions, structure and outcomes compared to TA462 was carried out
  - Different model structure used in TA462: 3 heath states (progression-free, progressed, dead) in a semi-Markov model
  - Progression post-alloSCT is incorporated in modelling in TA462 (this is not allowed in the current assessment model)
  - Different assumptions were made about the composition of SOC treatment in TA462 modelling
  - In TA462 modelling patients may receive alloSCT after 6 months
- Higher total QALYs (almost doubled) and costs (more than doubled) generated by modelling for SOC (cohort 1) in this assessment compared to TA462

70

Source: ERG report, section 5.2.12 (pages 95 and 96)

Table comparing model and trial outcomes can be found in the company submission, table 102 (page 220) and in table 5.24 in the ERG report, section 5.2.12.4 (page 95)

Table comparing SOC results from TA462 and the current assessment can be found in the ERG report, section 5.2.12, table 5.25 (page 96)

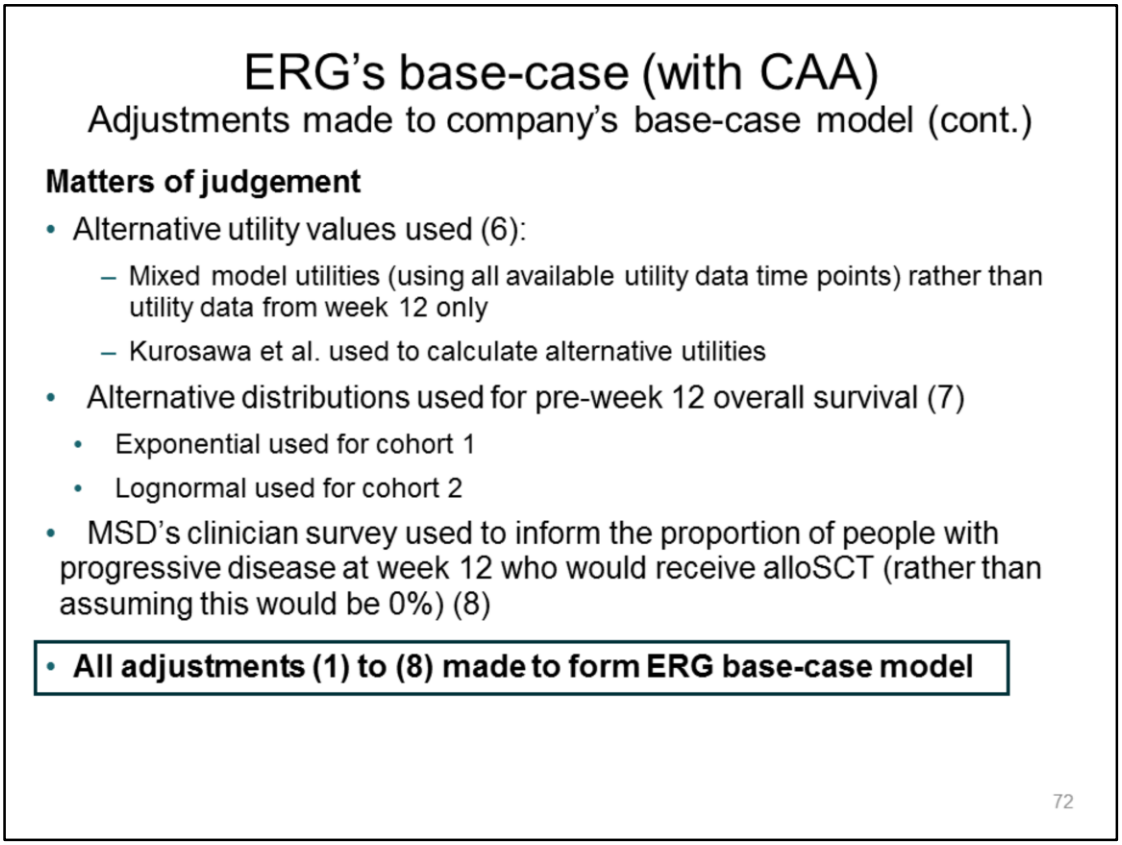


Source: ERG report, section 5.3 (page 98)

Adjustments made by ERG divided into 3 categories:

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model were the ERG considers that reasonable alternative assumptions are preferred)

From ERG report, section 5.3 (page ~~60~~ 96 in ~~draft~~ report)



Source: ERG report, section 5.3 (pages 98 and 99)

Further details on alternative utilities used by the ERG can be found on earlier slides (slide 60) and in the ERG report, section 5.2.8

For pre-week 12 overall survival, in its base-case the company used the models with the highest mortality at week 12. The ERG have used the models with the best statistical fit (see slides 39 and 40 for fit of parametric models)

Most influential adjustments made by ERG (in descending order): (1) use of alternative utility values, (2) use of MSD clinician survey data only to inform uptake of alloSCT dependent on response, (3) allowing patients in progressed disease state to have alloSCT

ERG's base-case - deterministic (with CAA)

Effects of ERG's adjustments

Adjustment	ICER (pembrolizumab versus SOC)	
	Cohort 1	Cohort 2
Company base-case	£43,511	£48,571
Fixing errors **(1) and (2)**	£43,262	£48,178
MSD survey only used for alloSCT probabilities **(3)***	£48,363	£55,478
50 year time horizon **(4)***	£42,412	£47,141
Monitoring costs included post-alloSCT **(5)***	£43,927	£48,908
Alternative utility values **(6)***	£52,705	£59,223
Alternative pre-week 12 OS distributions **(7)***	£43,262	£48,236
Proportion of alloSCT in PD state taken from MSD survey **(8)***	£46,841	£53,508

* Condition on fixing errors (1) and (2)

73

Source: ERG report, section 6, tables 6.1 and 6.2 (pages 103 and 104)

ERG's base-case results (with CAA)

ERG base-case (deterministic) – combines adjustments (1) to (8)

Treatment		Total		Incremental		ICER
		Costs	QALYs	Costs	QALYs	
Cohort 1	SOC	£50,913	3.535	-	-	
	Pembrolizumab	£107,998	4.460	£57,085	0.925	**£61,705**
Cohort 2	SOC	£50,609	3.541	-	-	
	Pembrolizumab	£93,095	4.118	£42,486	0.577	**£73,594**

ERG base-case (probabilistic)

Treatment	ICER (versus SOC)	Probability of cost-effectiveness of pembrolizumab compared with SOC	
		Maximum acceptable ICER	
		£30,000/ QALY	£50,000/ QALY
Pembrolizumab – cohort 1	£64,186	18%	42%
Pembrolizumab – cohort 2	£78,696	21%	40%

74

Source: ERG report, section 5.3.1 (page 99); section 6, table 6.1 (page 103), table 6.2 (page 104)

Cost effectiveness acceptability curves for the ERG base-case can be found in the ERG report, section 5.3.1 (pages 99 and 100)

ERG's base-case model

Further exploratory analysis (deterministic)

Exploratory analysis	Cohort	ICER
		Pembrolizumab versus SOC
Exploratory analysis 1a Alternative parametric survival models: • Cohort 1: Gompertz used for post-week 12 PFS • Cohort 2: Gompertz used for post-week 12 PFS	Cohort 1	£68,966
	Cohort 2	£87,401
Exploratory analysis 1b Alternative parametric survival models: • Cohort 2: Generalised gamma used for post-week 12 PFS	Cohort 2	£90,152
Exploratory analysis 2 MAIC used instead of naïve indirect treatment comparison for PFS hazard ratios and response rates at week 12	Cohort 1	£54,466
	Cohort 2	£60,372

75

Source: ERG report, section 5.3.2 (page 100); section 6, tables 6.3 and 6.4 (pages 104 to 105)

Alternative parametric models for post-week 12 PFS

Cohort 1 (see slide 46)

Exponential used by company in base case based on best statistical fit. Gompertz had statistical fit within 2 AIC points and ERG considered it informative to explore the use of this model in further analysis.

Cohort 2 (see slide 47 and 48)

Generalised gamma had best statistical fit for post-week 12 PFS. The ERG noted that choice of post-week 12 PFS model in cohort is very influential – and that company's choice of exponential favoured pembrolizumab. From ERG report, section 5.2.6.3 (page 77)

ERG's base-case

Further exploratory analysis (cont.)

Exploratory analysis	Cohort	ICER
		Pembrolizumab versus SOC
Exploratory analysis 3 Removal of 24 months cap on time to treatment discontinuation for pembrolizumab	Cohort 1	£78,992
	Cohort 2	£79,284
Exploratory analysis 4 Lower post-alloSCT utility used (PD utility) to explore impact of not considering PD after alloSCT in company base-case model	Cohort 1	£63,420
	Cohort 2	£75,835
Exploratory analysis 5 Use of alternative assumptions to extrapolate post-alloSCT OS from Lafferty et al. (2017)	Cohort 1	£78,204
	Cohort 2	£95,712

76

Source: ERG report, section 5.3.2 (page 100); section 6, tables 6.3 and 6.4 (pages 104 and 105)

ERG's conclusions: Cost-effectiveness

- Company's economic model meets NICE reference case, except (1) time horizon (40 years) is too short, and (2) BSC – a comparator included in the scope – was excluded from the base-case
  - BSC was excluded because of lack of evidence; accepted by committee for TA462
- Major limitation is model structure: implausible assumption that people could only be eligible for, and receive, alloSCT 12 weeks after starting treatment
  - Response may occur later than 12 weeks
  - There will be a delay between decision to pursue alloSCT and when the procedure is done
  - Assumption lacked appropriate justification and differs from how alloSCT treatment was incorporated in TA462
  - Model structure meant different parametric models had to be fitted for pre- and post-12 weeks, adding additional uncertainty
- Impact of limitations due to model structure on outcome is unknown

77

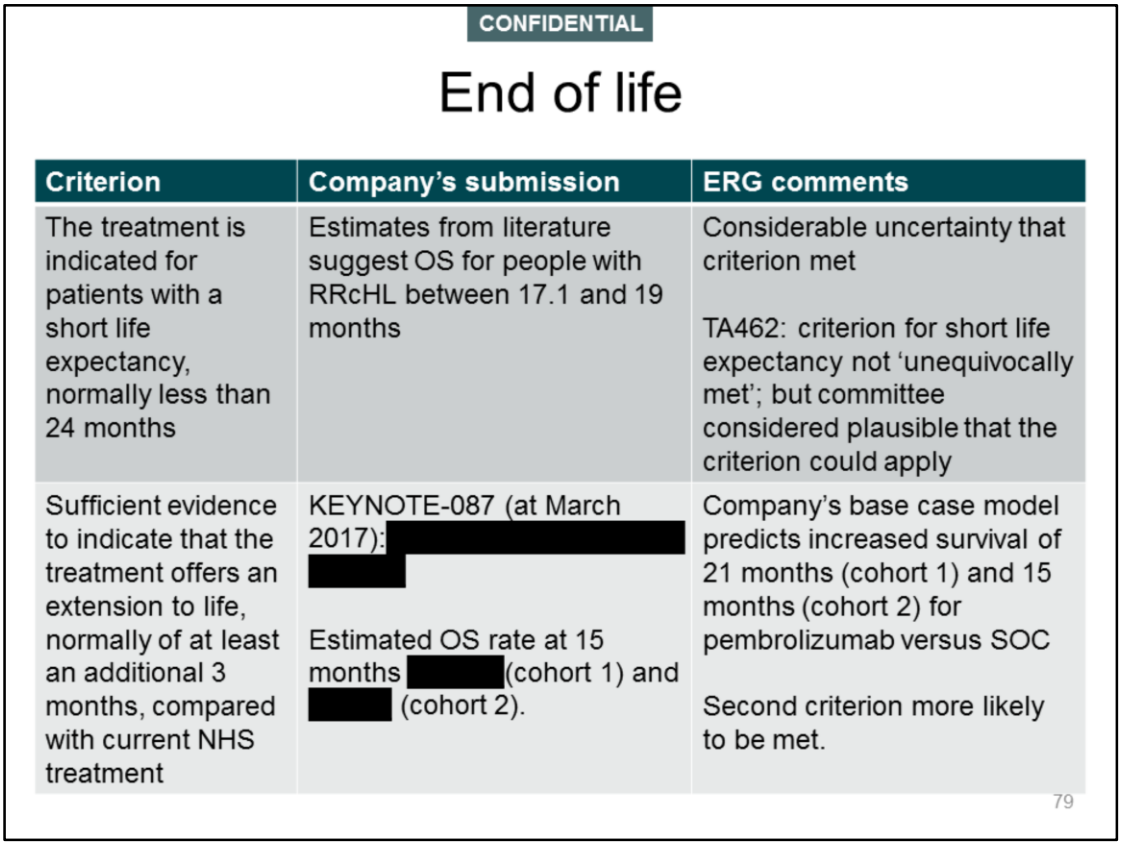
Source: ERG report, section 5.4 (pages 100 to 102)

ERG's conclusions: Cost-effectiveness (cont.)

- Comparator is a mixed population of cohorts 1 and 2
  - May be non-favourable for pembrolizumab (compared to SOC) for cohort 1 but potentially favourable for cohort 2
- Lifting the assumed capping of pembrolizumab at 24 months significantly increased ICERs, as did the use of alternative assumptions when extrapolating post-alloSCT OS data from Lafferty et al. (2017)
- Use of alternative models to extrapolate PFS post-week12 also had a large effect on ICERs
- Use of MAIC rather than naïve indirect comparison decreased ICERs
- None of the ERG's alternative scenarios resulted in an ICER below £50,000 per QALY gained
- Uncertainty about the cost effectiveness of pembrolizumab remains substantial

78

Source: ERG report, section 5.4 (pages 100 to 102)



Source: Company submission, section 4.13.3, table 51 (page 129); ERG report, section 7 (page 106)

TA462: Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma (2017) NICE technology appraisal guidance 462, published 26 July 2017

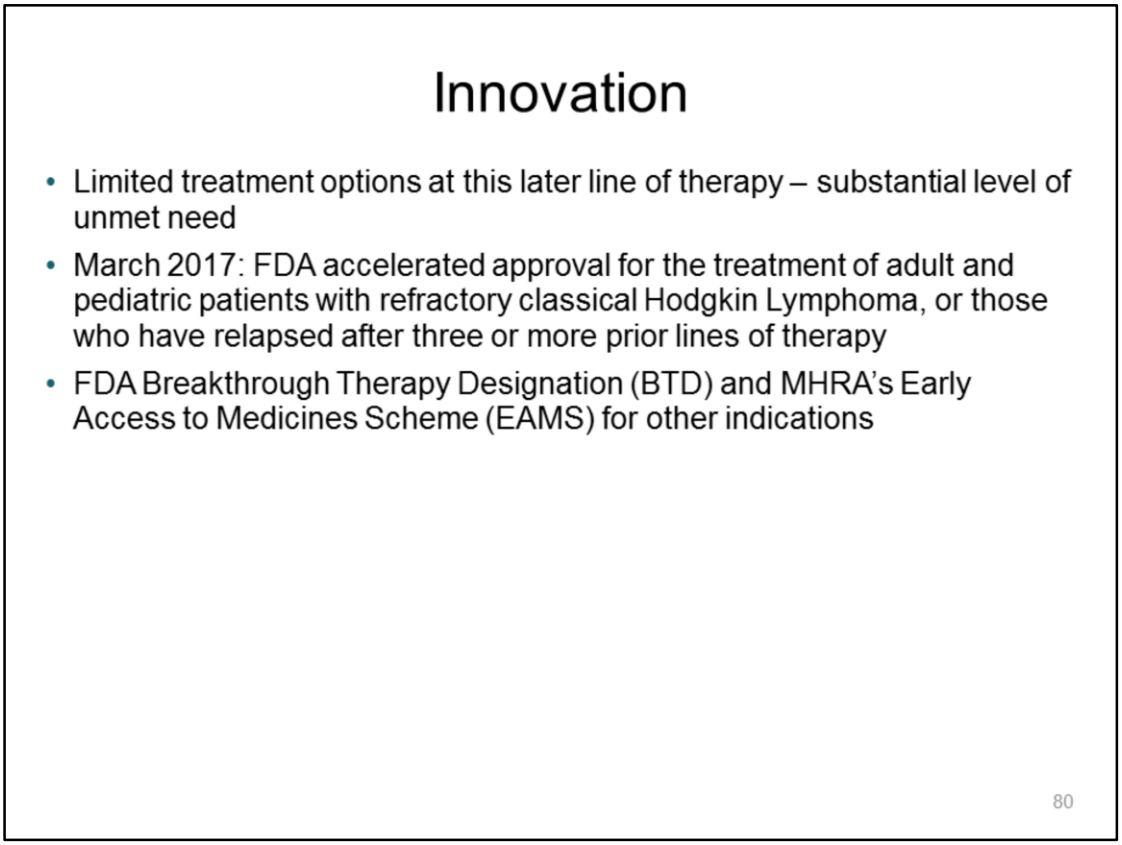
From TA462:

End-of-life considerations

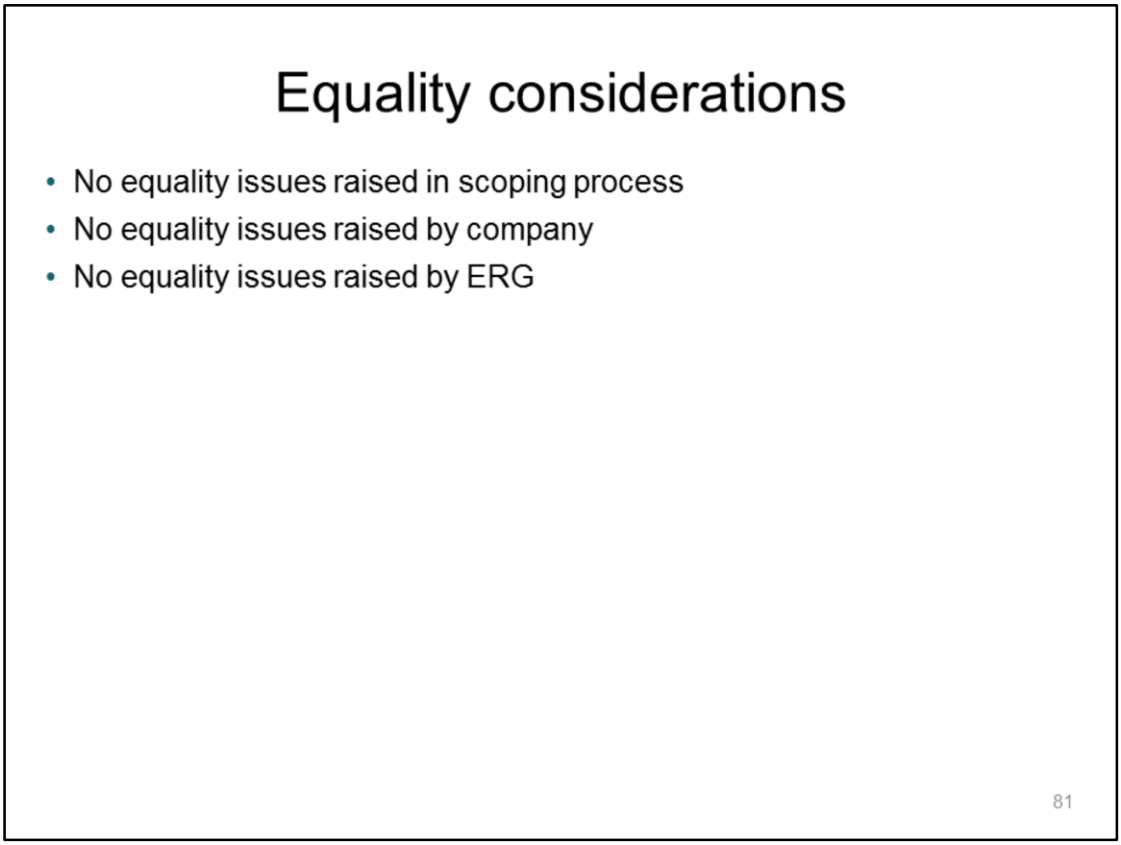
4.23 The committee considered the advice about life-extending treatments for people with a short life expectancy in NICE's final Cancer Drugs Fund technology appraisal process and methods. The company made the case that nivolumab met the criteria for life-extending treatments for people with a short life expectancy (normally less than 24 months). The committee noted that the company's modelling predicted a mean overall survival in the comparator treatment arm of more than 24 months. However, the committee also considered the data from the Haematological Malignancy Research Network provided by the company in response to consultation, which showed shorter survival and suggested that the Cheah study may have been optimistic. The committee acknowledged that nivolumab did not unequivocally meet the criterion for short life expectancy but that it was plausible that the criterion could apply. It therefore agreed that on balance, nivolumab met

the criterion for short life expectancy, and that it would take this into account in its decision-making.

4.24 The committee also discussed whether there was sufficient evidence to show that the treatment offers an extension to life of at least an additional 3 months compared with current NHS treatment. The committee noted that the cost-effectiveness analysis from which the survival benefit of nivolumab could be inferred did not reflect the committee's preferred analysis, and that because of the immaturity of the trial data and the lack of UK comparator data, all the estimates were uncertain. However, it concluded that based on the evidence presented, nivolumab met the criterion for extending life by at least an additional 3 months.



Source: Company submission, section 2.5 (pages 33 and 34)



Source: Company submission, section 3.8 (page 41)

Authors

- **Thomas Walker**
  Technical Lead
- **Rebecca Albrow**
  Technical Adviser
- With input from the Lead Team (Adrian Griffin, David Evans and Mohit Sharma)

82

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's Lymphoma

ID 1062

Merck Sharp & Dohme

Company evidence submission

September 2017

File name	Version	Contains confidential information	Date
		Yes	

Abbreviations

Abbreviation	Definition
ABVD regimen	Doxorubicin, bleomycin, vinblastine and dacarbazine
AE	Adverse event
AEOSI	Adverse events of special interest
AIC	Akaike Information Criterion
AlloSCT	Allogeneic Stem Cell Transplant
ASaT	All Subjects as Treated
ASCO	American Society of Clinical Oncology
ASHAP	doxorubicin, methylprednisolone, cytarabine, cisplatin
AUC	Area under the curve
AutoSCT	Autologous Stem Cell Transplant
AWMSG	All Wales Medicine Strategy Group
BCSH	British Committee for Standards in Haematology Guidelines
BEACOPP regimen	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone
BIC	Bayesian information criterion
BICR	Blinded independent central radiologists'
BNF	British National Formulary
BOR	Best Overall Response
BSA	Body surface area
BSC	Best Supportive Care
BTD	Breakthrough Therapy Designation
BV	Brentuximab Vedotin
CAA	Commercial access agreement
CDF	Cancer Drugs Fund
CEA	Cost-Effectiveness Analysis
cHL	classical Hodgkin Lymphoma
CHOP	cyclophosphamide, doxorubicin, prednisolone, vincristine
CI	Confidence Interval
CPS	Combined positive score
CR	Complete response
CSR	Clinical Study Report
CT	Computer Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DAEs	Discontinuations due to adverse-events
DHAOx	dexamethasone, cytarabine, oxaliplatin
DHAP	dexamethasone, cytarabine, cisplatin
DOR	Duration Of Response
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal Growth Factor Receptor
EMA	European Medicines Agency

EORTC-QLQC30	European Organisation for Research and Treatment Cancer Quality of Life Questionnaire
EQ-5D	EuroQoL 5 Dimensions
ERG	Evidence Review Group
ESHAP	etoposide, methylprednisolone, cytarabine, cisplatin
ESMO	European Society for Medical Oncology
FAS	Full analysis set
FDA	Food and Drug Administration
GDP	gemcitabine, dexamethasone, cisplatin
GEM-P	gemcitabine, cisplatin, methylprednisolone
GVD	gemcitabine, vinorelbine, liposomal doxorubicin
GVHD	Graft Versus Host Disease
HL	Hodgkin Lymphoma
HR	Hazard Ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICE	ifosfamide, carboplatin, etoposide
ICER	Incremental cost-effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
IGEV	ifosfamide, gemcitabine, vinorelbine
IRG	Independent Review Group
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITT	Intention-to-treat
IV	Intravenous
IVE	ifosfamide, epirubicin, etoposide
IVOx	ifosfamide, etoposide, oxaliplatin
KM	Kaplan-Meier
LY	Life Year
mAB	monoclonal antibody
MAIC	Matched Adjusted Indirect treatment comparison
MINE	mitoxantrone, ifosfamide, vinorelbine, etoposide
MK-3475	Pembrolizumab - *Keytruda*®
MSD	Merck Sharp and Dohme Ltd
NHL	non Hodgkin Lymphoma
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PD-1	Programmed death 1 protein
PD-L1	Programmed death ligand 1
PET	Positron Emission Tomography

PFR	Progression-free rate
PFS	Progression free survival
PI	Principal Investigator
PIM	Promising Innovative Medicines
PK	Pharmacokinetics
PMitCEBO	bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient Reported Outcomes
PSS	Personal Social Services
PSSRU	Personal and Personal and Social Services Research Unit
Q3W	Every 3 weeks
QALY(s)	Quality-Adjusted Life Year(s)
RR	Response rate
RRcHL	relapsed or refractory classical Hodgkin Lymphoma
RSC	Reed-Sternberg cells
RVIG	gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine
SAE	Serious Adverse event
SCT	Stem cell transplant
SD	Stable Disease
SD	Standard Deviation
SE	Standard Error
SG	Standard Gamble
SIGN	Scottish Intercollegiate Guidelines Network's
SLR	Systematic Literature Review
SMC	Scottish Medicines Consortium
SmPC	Summary of Product Characteristics
SOC	Standard of Care
STA	Single Technology Assessment
TA	Technology Appraisal
ToT	Time on Treatment
TTO	Time trade off
UK	United Kingdom of Great Britain and Northern Ireland
VAS	Visual Analogue Scale
VAT	Value-Added Tax

Contents

Tables & Figures

Tables

FIGURES

1.0 Executive summary

Classical Hodgkin Lymphoma is a rare, localised or disseminated, malignant proliferation of cells of the lymphoreticular system. Classical Hodgkin Lymphoma is typically localised to a group of connected lymph nodes but can spread throughout the lymphatic system and in late-stage disease will metastasize to other areas of the body, most commonly the chest, neck, or under the arms.

In the UK (2014) there were 2,106 new cases of Hodgkin Lymphoma in the UK; this equates to an age standardised rate of 3.3 (95% CI 3.2-3.5) per 100,000 persons [1]. Surveillance data within the UK (England, Scotland, and Wales), as reported by Cancer Research UK, shows that the incidence of Hodgkin Lymphoma follows a bimodal age distribution, with the first peak in young adults (20-24 years) and the second in older males and females (75-79 years); around half of diagnoses (50%) were reported in persons aged 45 years and over [1].

The literature suggests that patients who are described as relapsed/ refractory have poor prognosis compared with their counterparts who respond to therapy. A single retrospective trial of 81 patients with relapsed/ refractory disease showed that of those who failed autologous stem cell transplant (autoSCT), 96% had relapsed within two years [2]. The five year survival among these patients was markedly lower than those reported by Cancer Research UK at less than 20% [2], demonstrating the high level of unmet need within this difficult to treat patient group.

The treatment of patients with Hodgkin Lymphoma varies according to a number of factors. In those patients who do not achieve long term remission salvage therapy may include chemotherapy and/or radiotherapy with the intent to enable autoSCT, which is regarded as potentially curative [3].

The cost-effectiveness of pembrolizumab was evaluated through the development of a cohort based model composed of a short-term decision-tree to predict response and alloSCT uptake of the population during the first 12 weeks of treatment and a set of Markov state transition models to predict the lifetime survival of patients from Week 12 to death, conditional on alloSCT uptake or continued use of pembrolizumab or SoC. The model projected health outcomes (i.e. OS and PFS) to estimate patients' health-related quality of life (HRQoL) and costs. Quality-adjusted life years (QALYs) were estimated by considering utility derived from EQ-5D data collected in KEYNOTE-087 trial. Clinical and economic outcomes were projected

over a 40-year time horizon to cover the anticipated lifetime of the population initiating this late line therapy and assessed as part of this submission.

The results demonstrate that pembrolizumab, as an end of life therapy; can be considered a cost-effective use of NHS resources. The model estimates that patients treated with pembrolizumab gain 1.274 and 0.871 additional QALYS compared to UK SoC in cohorts 1 and 2 respectively. The incremental cost-effectiveness ratio (ICER) when comparing pembrolizumab to UK SoC is £43,511 and £48,571 for cohorts 1 and 2 respectively (discounted). The probability of pembrolizumab being the most cost-effective treatment at a threshold of £50,000 per QALY gained is therefore 60% and 50% respectively.

The availability of pembrolizumab as a treatment option in England, for patients with RRcHL, will represent a step-change in the treatment options available and provide patients and clinicians with a transformative new treatment alternative.

1.1 Statement of decision problem

The decision problem addressed in the submission is presented in the below.

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with relapsed or refractory classical Hodgkin Lymphoma who have received: • autologous stem cell transplant and brentuximab vedotin • brentuximab vedotin when autologous stem cell transplant is not a treatment option.	As per final scope	Not applicable
Intervention	• Pembrolizumab	• Pembrolizumab	• Pembrolizumab
Comparator (s)	• Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin • Best supportive care.	Standard of care as per Cheah et al. 2016) including: • Investigational agent • Gemcitabine • Bendamustine • Other alkylatory • BV retreatment • Platinum based	Cheah et al. 2016 reported outcome data for a mix of chemotherapy regimens and was preferred by the ERG in TA462. To separate individual regimens survival outcome data would not have been possible in the absence of individual patient level data and hence conservatively MSD have included all survival outcomes reported here.

		• autoSCT • Other	
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life.	As per final scope, with the exception of long term overall survival data. The model structure utilised OS data from week 0-12 from KEYNOTE-087, response rates at week 12, PFS from week 12 onward and external literature OS sources for post alloSCT survival	At follow up (15.9 month), [redacted] . Hence all available data from KEYNOTE-087 has been utilitsed where possible.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or	As per final scope	Not applicable

	comparator technologies will be taken into account.		
Subgroups to be considered	If the evidence allows, a scenario analysis including allogeneic stem cell transplant as a subsequent treatment after pembrolizumab or its comparators will be considered. This should reflect the proportion of people who proceed to allogeneic stem cell transplant after each treatment, as well as the costs and quality-adjusted life year benefits of the procedure		
Special considerations including issues related to equity or equality	Not applicable	Not applicable	Not applicable

1.2 Description of the technology being appraised

Pembrolizumab is a highly selective humanised monoclonal antibody against programmed death-1 (PD-1) receptor, which exerts dual ligand blockade of the PD-1 pathway, including PD-L1 and PD-L2, on antigen presenting tumour cells. By inhibiting the PD-1 receptor from binding to its ligands, pembrolizumab activates tumour-specific cytotoxic T lymphocytes in the tumour microenvironment and reactivates antitumour immunity (see section 2.1).

The route of administration for pembrolizumab is IV infusion, over a 30-minute period. The anticipated licensed dosing regimen for patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and Brentuximab Vedotin (BV), or who are transplant ineligible and have failed BV is 200mg Q3W.Treatment with pembrolizumab continues until disease progression or unacceptable toxicity, whichever occurs first. The list price of pembrolizumab is £2,630 per 100mg vial [redacted]

Regulatory approval by the EMA for the indication considered within this submission was granted on the 2nd May 2017. The final indication is: KEYTRUDA as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and BV, or who are transplant-ineligible and have failed BV.

The innovative nature of pembrolizumab has been recognised on a number of occasions across numerous oncology indications. Relevant to this indication, on the 14th March the US Food and Drug Administration (FDA) granted pembrolizumab Orphan Drug Designation for the treatment of HL, and Breakthrough Therapy Designation; this application also received priority review status and accelerated approval[4].

Table 2. Technology being appraised

UK approved name and brand name	Pembrolizumab (KEYTRUDA®)
Marketing authorisation/CE mark status	Pembrolizumab currently has a marketing authorisation covering the following indications: • KEYTRUDA as monotherapy is indicated for the treatment of advanced (unresectable or metastatic) melanoma in adults. • KEYTRUDA as monotherapy is indicated for the first-line treatment of metastatic non-small cell lung carcinoma (NSCLC) in adults whose tumours express PD-L1 with a ≥50% tumour proportion score (TPS) with no EGFR or ALK positive tumour mutations. • KEYTRUDA as monotherapy is indicated for the treatment of locally advanced or metastatic NSCLC in adults whose tumours express PD-L1 with a ≥1% TPS and who have received at least one prior chemotherapy regimen. Patients with EGFR or ALK positive tumour mutations should also have received targeted therapy before receiving KEYTRUDA. • KEYTRUDA as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma (cHL) who have failed autologous stem cell transplant (ASCT) and brentuximab vedotin (BV), or who are transplant-ineligible and have failed BV.
Indications and any restriction(s) as described in the summary of product characteristics	Indication to which this submission relates: • KEYTRUDA as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma (cHL) who have failed autologous stem cell transplant (ASCT) and brentuximab vedotin (BV), or who are transplant-ineligible and have failed BV.
Method of administration and dosage	200 mg every three weeks (Q3W); intravenous (IV) infusion.

1.3 Summary of the clinical effectiveness analysis

A systematic literature review was conducted to identify relevant clinical trials from the published literature (see Section 4.1).

The clinical effectiveness and safety evidence described within this submission are taken from the KEYNOTE-087 trial. This is a single arm, non-randomised, non-comparative, trial of pembrolizumab 200mg Q3W for patients with RRcHL who have failed/ or who are considered ineligible for autoSCT and who have subsequently failed treatment with BV. To align with the EMA license, and CUA approach taken within this submission, MSD has presented post-hoc analysis of the efficacy data relevant to the two populations described. This submission utilises the most recent data available; this is March 2017 for efficacy, and September 2016 for safety.

As KEYNOTE-087 is non-comparative, a SLR was undertaken to identify relevant literature to enable comparative effectiveness estimates. This was an existing SLR conducted for internal purposes, and was updated in June 2017 to meet the requirements of NICE as per the decision problem (Section 1.1). A single retrospective, observational study was identified and included (Cheah et al. 2016). In addition to this, MSD commissioned a clinical survey to provide UK specific validation of the literature.

The baseline characteristics of the patients included in KEYNOTE-087 were as expected for patients with rrcHL, who are typically heavily pre-treated, and can be considered representative of the patients who are anticipated to receive pembrolizumab in UK clinical practice (see Section 4.5).

The efficacy results of KEYNOTE-087 demonstrate the substantial benefit of pembrolizumab in patients with RRcHL who have received prior therapy with BV and in some cases an autoSCT following first line chemotherapy. Results presented are regardless of PDL-1 expression. The post-hoc analysis results for ORR (primary objective) was [redacted] in Cohort 1 and [redacted] in Cohort 2. This high response rate has translated into a lower incidence of progression and extended survival. The PFS among patients with an overall response at Week 12 onwards, as per the post-hoc analysis, was [redacted] months for Cohort 1 and [redacted] months for Cohort 2. Note that OS data at this time is immature. However, the rate of OS at 3 to 12 months is in excess of [redacted] across Cohort 1 and 2.

The observed safety profile of pembrolizumab was as expected within such a heavily pre-treated patient population. Furthermore, this safety profile is consisted with the safety profile established to date, and demonstrates that pembrolizumab is well tolerated in the target population, offering favourable tolerability compared to SoC chemotherapy regimens. As reported in Section 4.12 the majority of AEs were considered low grade, and few patients discontinued treatment due to AEs of any grade. Mortality within the population was low, and of the deaths reported none were considered study drug related. In general, the frequency and severity of AEOSI during the trial were similar to the previously described characterisation of the safety profile of pembrolizumab.

In addition to efficacy and safety, a clinically meaningful improvement in PROs was also observed. Patients reported an increase from baseline using the EQ-5D and EORTC-QLQ-C30 disease specific health related quality of life questionnaires.

As the SLR did not identify any relevant trials that would have allowed the formation of a connected evidence network; all analysis were conducted using Cheah et al. 2016 and KEYNOTE-087. Cheah et al. 2016 has been previously accepted by NICE for decision making within this patient population, and is considered to be the most appropriate source of evidence within this limited patient group. To enable the committee to consider this evidence versus pembrolizumab it was necessary to conduct a naive comparison, and a MAIC. Both analyses demonstrate a statistically significant improvement for pembrolizumab versus mixed agent SoC as reported in Cheah et al. 2016. Due to a lack of granularity within the Cheah et al. 2016 it was not possible to remove the effect of investigational agents that may have impacted the results; thus the results of the analysis presented can to an extent be considered conservative (Section 4.10.14).

In summary, the results of these analyses underscore the benefit of pembrolizumab as a treatment option for this patient group, who currently face a very poor prognosis.

1.4 Summary of the cost-effectiveness analysis

The cost-effectiveness of pembrolizumab was assessed against UK SoC, in patients with RRcHL who have either been treated with an autoSCT and BV (cohort 1) or are ineligible for an autoSCT and have received BV (cohort 2).

Cost-effectiveness was evaluated through the development of a cohort based model composed of a short-term decision-tree to predict response and alloSCT uptake of the population during the first 12 weeks of treatment and a set of Markov state transition models to predict the lifetime survival of patients from Week 12 to death, conditional on alloSCT uptake or continued use of pembrolizumab or SoC. The analysis was conducted in line with the NICE reference case. A discount rate of 3.5% per annum was applied to both costs and benefits. Clinical and economic outcomes were projected over a 40-year time horizon to cover the anticipated lifetime of the population here assessed. The analysis was run using 1-week model cycle. The model projected health outcomes (i.e. OS and PFS) to estimate patients' health-related quality of life (HRQoL) and costs. Quality-adjusted life years (QALYs) were estimated by considering utility derived from EQ-5D data collected in KEYNOTE-087 trial.

Since KEYNOTE-087 was a single arm study, the clinical evidence used to populate the UK SoC arm was derived from a set of naïve indirect comparisons of pembrolizumab vs UK SoC from Cheah et al 2016, details of which are mentioned in the previous section.

In the no alloSCT part of the model, PFS for pembrolizumab and UK SoC were modelled by extrapolating KEYNOTE-087 PFS data from week 12 and applying a HR derived from the naïve indirect comparison.

In the alloSCT part of the model, OS from an external literature source was extrapolated equal outcomes assumed in both the pembrolizumab and UK SoC treatment arms.

Section 5 details the development of the de novo economic model for pembrolizumab, with Table 3 below presenting the results for the two main populations of patients with RRcHL considered in this submission.

The model estimates that patients treated with pembrolizumab gain 1.274 and 0.871 additional QALYS compared to UK SoC in cohort 1 and 2 respectively. The incremental cost-effectiveness ratio (ICER) when comparing pembrolizumab to UK SoC is £43,511 and £48,571 for cohort 1 and 2 respectively. The probability of pembrolizumab being the most cost-effective treatment at a threshold of £50,000 per gained QALY is 60% and 50% for cohort 1 and 2 respectively.

Results from multiple sensitivity analyses showed the ICER to be consistently below £50,000 per QALY (discounted, with the PAS). The inputs that most affect the cost-effectiveness results relate to the discount rate applied to outcomes and the odds ratio applied to response rates CR and PR at 12 weeks. The sensitivity analyses conducted demonstrated that the cost-effectiveness of pembrolizumab is resilient to the different sources of uncertainty assessed.

Table 3: Incremental cost-effectiveness results – Base case, main population

Technologies	Cohort	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
UK SoC	Cohort 1	52,017	4.864	3.223	-	-	-
	Cohort 2	51,424	4.832	3.200	-	-	-
Pembrolizumab	Cohort 1	107,459	6.252	4.497	55,442	1.274	43,511
	Cohort 2	93,732	5.775	4.072	42,308	0.871	48,571

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

2.0 The technology

2.1 Description of the technology

Brand name: KEYTRUDA®

Generic name: pembrolizumab

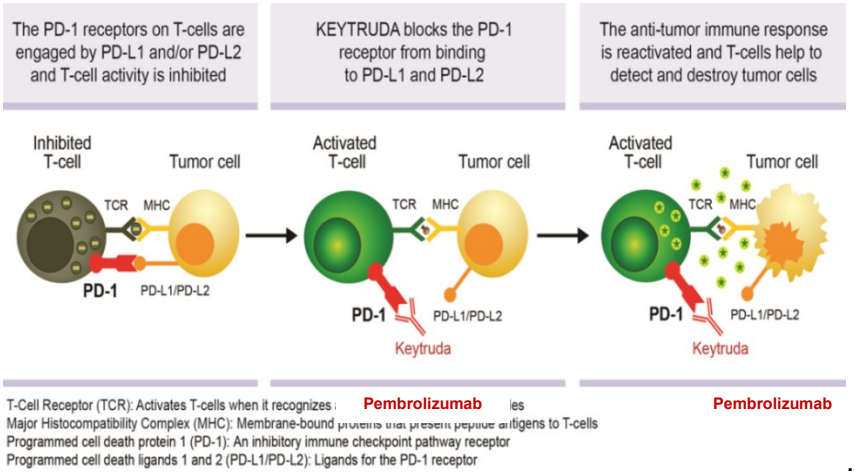
Therapeutic class: BNF Category "Other immunomodulating drugs" (08.02.04)

Brief overview of mechanism of action:

Programmed death 1 protein (PD-1) is an immune-checkpoint receptor that is expressed on antigen-presenting T cells. PD-1 acts to initiate downstream signalling, which in turn inhibits the proliferation of T cells as well as cytokine release and cytotoxicity [5]. The PD-1 ligands, PD-L1 and PD-L2, are frequently upregulated on the surface of many tumour cell surfaces [6].

Pembrolizumab (Keytruda®) is a potent and highly selective humanised monoclonal antibody (mAb) of the IgG4/kappa isotype designed to exert dual ligand blockade of the PD-1 pathway by directly blocking the interaction between PD-1 and its ligands, PD-L1 and PD-L2 which appear on antigen-presenting or tumour cells [5]. By binding to the PD-1 receptor and blocking the interaction with the receptor ligands, pembrolizumab releases the PD-1 pathway-mediated inhibition of the immune response, and reactivates both tumour-specific cytotoxic T lymphocytes in the tumour microenvironment and antitumour immunity.

Figure 1. Pembrolizumab – mechanism of action



Source: MSD data on file.

2.2 *Marketing authorisation/CE marking and health technology assessment*

2.2.1. Current UK regulatory status

As per the indication assessed within this submission, EMA marketing authorisation was granted on the 2nd May 2017

2.2.2. Approved EMA indication relevant to the UK

KEYTRUDA® as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and BV, or who are transplant-ineligible and have failed BV (Appendix 1).

2.2.3. Anticipated date of availability in the UK

Pembrolizumab (Keytruda®) has been available in the UK since 2015. For the indication under consideration, the EMA granted regulatory approval on the 2nd May 2017.

2.2.4. Summary of product characteristics

Please see Appendix 1 for the for the final summary of product characteristics (2nd August 2017)

2.2.5. Restrictions or contraindications that are included in the summary of product characteristics (SmPC) as reported in Appendix 1

As per section 4.2 of the SmPC, Keytruda® should be permanently discontinued for grade 4 toxicity; this is available in Appendix 1: Summary of product characteristics. Of particular relevance to this indication, patients who experience a Grade 4 haematological toxicity may have Keytruda® withheld until adverse reactions recover to Grade 0-1.

The SmPC also highlights there are limited data to draw conclusions for patients with classical Hodgkin Lymphoma aged ≥65 years.

The SmPC reports complications of allogeneic haematopoietic stem cell transplant (alloSCT) in patients with classical Hodgkin Lymphoma. Of 23 patients with classical Hodgkin Lymphoma who proceeded to alloSCT after treatment with pembrolizumab, 6 patients (26%) developed graft-versus-host-disease, one of which was fatal and 2 patients (9%) developed severe hepatic veno-occlusive disease after reduced-intensity conditioning, one of which was fatal. The 23 patients had a median follow-up from subsequent alloSCT of 5.1 months (range: 0-26.2 months). Until further data become available, careful consideration to the potential

benefits of haematopoietic stem cell transplant and the possible increased risk of transplant-related complications should be made case by case (Appendix 1).

2.2.6. Details of regulatory approval outside of the UK

On March 14, 2017, The U.S. Food and Drug Administration granted accelerated approval to pembrolizumab (KEYTRUDA), Merck and Co., Inc.) for the treatment of adult and paediatric patients with refractory classical Hodgkin Lymphoma (cHL), or those who have relapsed after three or more prior lines of therapy.

2.2.6 Other health technology assessments in the UK

MSD will be making a submission to the Scottish Medicines Consortium (SMC) in [redacted] as per the license indication considered within this submission.

2.3 Administration and costs of the technology

Table 4. Costs of the technology being appraised

	Cost	Source
Pharmaceutical formulation	Concentrate for solution for infusion	SmPC (see Appendix 1)
Acquisition cost (excluding VAT) *	List price: 100mg vial = £2,630. [redacted]	Department of Health
Method of administration	Intravenous infusion	SmPC (see Appendix 1)
Doses	Induction dose: 200mg	SmPC (see Appendix 1)
Dosing frequency	200mg every 3 weeks until disease progression or unacceptable toxicity	SmPC (see Appendix 1)
Average length of a course of treatment	Based on KEYNOTE-087 trial, the average time on therapy per patient: [redacted] Patients are treated with pembrolizumab 200mg Q3W during a course of treatment.	CSR KEYNOTE-087
Average cost of a course of treatment	The average cost per treatment course is: £[redacted] (based on average of [redacted] cycles) at list price	CSR KEYNOTE-087
Anticipated average interval between courses of treatments	Treatment is continued until disease progression or unacceptable toxicity leading to discontinuation	SmPC (see Appendix 1)
Anticipated number of repeat courses of treatments	Repeated treatment is not anticipated	SmPC (see Appendix 1)
Dose adjustments	No dose adjustment is expected	SmPC (see Appendix 1)
Anticipated care setting	Pembrolizumab is anticipated to be administered in a hospital setting	

* Indicate whether this acquisition cost is list price or includes an approved patient access scheme. When the marketing authorisation or anticipated marketing authorisation recommends the intervention in combination with other treatments, the acquisition cost of each intervention should be presented.

2.4 Changes in service provision and management

2.4.1 Additional tests or investigations needed

No additional tests or investigations are required further to the usual tests undertaken in current clinical practice. No diagnostic test is required to identify the population for whom pembrolizumab is indicated and no particular administration for the technology is required.

2.4.2 Main resource use to the NHS associated with the technology being appraised

Pembrolizumab is administered until disease progression or unacceptable toxicity. The main resource use to the NHS associated with the use of pembrolizumab is therefore expected to be related to the management of patients in the pre-progression period.

The administration of pembrolizumab will take place in secondary care (i.e. hospital setting) with no inpatient stay required. Patients will receive pembrolizumab in the outpatient setting on a 3-weekly cycle, with duration of administration of 30 minutes per infusion (SPC, Appendix 1) [7].

2.4.3 Additional infrastructure in the NHS

Pembrolizumab is not anticipated to require any additional infrastructure in the NHS to be put in place.

2.4.4 Extent that the technology will affect patient monitoring compared with established clinical practice in England

Pembrolizumab is expected to provide durable benefit for a proportion of patients treated. These patients can be anticipated to receive on-going follow-up including scanning.

2.4.5 Concomitant therapies administered with the technology

No concomitant therapies are required.

2.5 Innovation

Pembrolizumab represents a stepwise change in the management of patients with RRcHL following treatment with BV. Pembrolizumab, a checkpoint inhibitor, is able to interact with a patient's immune system to destroy cancer cells, as described in Section 2.1. Furthermore, given the limited treatment options available for patients at this later line of therapy (Section 6), it is expected that both clinicians and patients would value an alternative to current standard of care where outcomes are poor [8, 9]. Thus, there is a substantial level of unmet need within this patient population.

The innovative nature of pembrolizumab was first recognised by the US Food and Drug Administration (FDA) in January 2013 by granting it Breakthrough Therapy Designation (BTD) for advanced melanoma [10]. The FDA's BTD is intended to expedite the development and review of a drug that is planned for use, alone or in combination, to treat a serious or life-threatening disease or condition when preliminary clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies on one or more clinically significant endpoint [11]. Pembrolizumab has continued to be recognised for its innovation within numerous tumour types, as described:

- October 2014, FDA BTD for the treatment of patients with advanced (metastatic) NSCLC whose disease has progressed after other treatments [11]
- October 2015, FDA accelerated approval for the treatment of patients with metastatic NSCLC whose tumours express PD-L1 as determined by an FDA-approved test and who have disease progression on or after platinum-containing chemotherapy [11].
- December 2015, FDA expand pembrolizumab label to include the treatment of patients with unresectable or metastatic melanoma [12].
- August 2016, FDA accelerated approval for the treatment of patients with recurrent or metastatic head and neck squamous cell carcinoma (HNSCC) with disease progression on or after platinum-containing chemotherapy [13].
- September 2016, FDA BTD and priority review for the first-line treatment of patients with advanced non–small cell lung cancer whose tumours express PD-L1 [14].
- February 2017, FDA BTD for the second-line treatment of patients with locally advanced or metastatic urothelial cancer with disease progression on or after platinum-containing chemotherapy [15].
- March 2017, FDA accelerated approval for the treatment of adult and pediatric patients with refractory classical Hodgkin Lymphoma, or those who have relapsed after three or more prior lines of therapy [4].

- May 2017, FDA accelerated approval to pembrolizumab in combination with pemetrexed and carboplatin for the treatment of patients with previously untreated metastatic non-squamous non-small cell lung cancer [16].

In the UK, in March 2015 pembrolizumab became the first medicine to be granted positive scientific opinion under the MHRA's Early Access to Medicines Scheme (EAMS) for the treatment of unresectable or metastatic melanoma with progressive, persistent, or recurrent disease on or following treatment with standard of care [17]. Pembrolizumab received Promising Innovative Medicines (PIM) designation (EAMS Step 1) in November 2015, and in March 2016 a positive Scientific Opinion was granted (MHRA EAMS number 00025/0001) for "*the treatment as monotherapy of adults with metastatic NSCLC whose tumours express PD-L1 as determined by a validated test and who have not received prior systemic therapy and are negative for EGFR sensitising mutation and ALK translocation or whose disease has progressed on or after platinum-containing chemotherapy. Patients who have an EGFR sensitising mutation or an ALK translocation should also have had disease progression on approved therapies for these aberrations prior to receiving pembrolizumab*" [18]. EAMS aims to give earlier access to promising new unlicensed or 'off label' medicines to UK patients that have a high unmet clinical need. This validates MSD's position that pembrolizumab should be considered innovative in its potential to make a significant and substantial impact on health-related benefits in an area of high unmet need.

3.0 Health condition and position of the technology in the treatment pathway

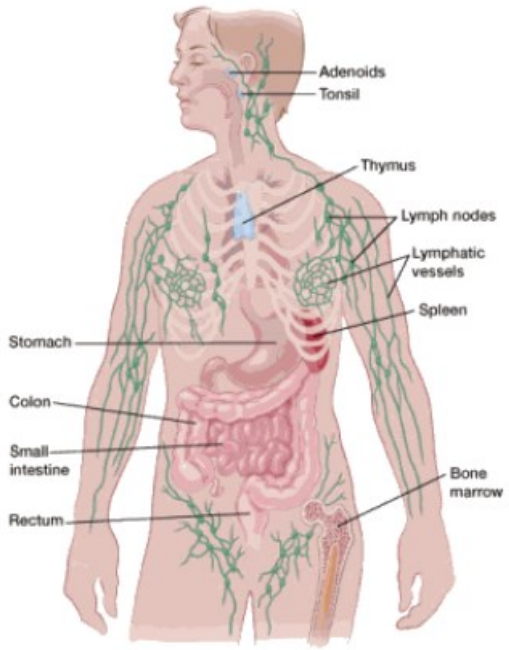
3.1: Brief overview of the disease/condition for which the technology is being used

Classical Hodgkin Lymphoma (cHL) is a rare, localised or disseminated, malignant proliferation of cells of the lymphoreticular system, occurring mostly in lymph node tissues, spleen, liver, and bone marrow [19]. Classical Hodgkin Lymphoma is typically localised to a group of connected lymph nodes but can spread throughout the lymphatic system and in late-stage disease will metastasise to other areas of the body, most commonly the chest, neck, or under the arms Figure 2 [20].

Hodgkin Lymphoma is comprised of two subgroups: Classical Hodgkin Lymphoma, which accounts for 95%, and nodular lymphocyte predominant Hodgkin Lymphoma, which is responsible for the remaining 5% of cases [21]. Classical Hodgkin Lymphoma is characterised by the presence of binucleated Reed-Sternberg cells (RSC), which are the result of clonal transformation of germinal centre B-cells, located within secondary lymph nodes of the lymphatic system [22].

Patients with cHL may present with a variety of symptoms, including swelling of lymph nodes, persistent fatigue, fevers and chills, night sweats, weight loss, loss of appetite, and itching [23, 24]. Patients are typically divided into those that have B symptoms (presence of fever, weight loss, and drenching night sweats) and those without[24]. The presence of B symptoms is associated with the development of advanced forms of disease and worse outcomes[25]. In some patients whose disease affects the lymph nodes in the chest, swelling of these nodes may press against the trachea and manifest as coughing or other breathing difficulties [23, 26].

Figure 2. Lymphatic system of the human body



Source: *Adapted from American Cancer Society* [21]

3.2: Effects of the disease/condition on patients, carers and society

Cancer Research UK reports that in 2014 there were 2,106 new cases of Hodgkin Lymphoma in the UK; this equates to an age standardised rate of 3.3 (95% CI 3.2-3.5) per 100,000 persons [1]. Utilising these observed trends, incidence rates have been extrapolated by Cancer Research UK among both male and female patients. It is expected that incidence rates may increase by 5% in the UK population overall between 2014 and 2035; this equates to 4 cases per 100,000 persons. It should be noted that age standardised incidence rates in the UK could rise by 9% in males between 2014 and 2035 (5 cases per 100,000), whilst decreasing by 1% in females during the same time period (3 cases per 100,000 persons) [1].

Surveillance data within the UK (England, Scotland, and Wales), as reported by Cancer Research UK, shows that the incidence of Hodgkin Lymphoma follows a bimodal age distribution, with the first peak in young adults (20-24 years) and the second in older males and females (75-79 years); around half of diagnoses (50%) were reported in persons aged 45 years and over [1].

Survival data for patients diagnosed with Hodgkin Lymphoma (England and Wales 2010-2011) appears promising at 91.4%, 85.0%, and 80.4% at years 1, 5 and 10, respectively[27]. However, these values should be interpreted with caution and are likely to be substantially

different in the context of the later line of therapy being considered within this submission document. The literature suggests that patients who are described as relapsed/ refractory (rr) have poor prognosis compared with their counterparts who respond to therapy. A single retrospective trial of 81 patients with relapsed/ refractory disease showed that of those who failed autoSCT, 96% had relapsed within two years [2]. This study also reported worse outcomes for those patients who relapsed within 6 months compared with those who relapsed after 6 months with a median OS of 15 month and 36 months, respectively [2]. The five year survival among these patients was markedly lower than those reported by Cancer Research UK at less than 20% [2].

To understand the economic burden of cHL a review was conducted to identify all relevant direct and indirect treatment costs; this review identified 12 studies relevant to patients with diagnoses of RRcHL. Of particular interest was a retrospective UK observation study that reported costs related to treatment, hospital stay, outpatient visits, scans, and day care visits[28]. The treatment pathways in these patients were; chemotherapy followed by alloSCT, palliative chemotherapy; chemotherapy followed by second autoSCT; and best supportive care[28]. Chemotherapy followed by alloSCT was the most expensive treatment pathway (mean cost of £110,374 per patient) followed by palliative chemotherapy (mean cost of £32,264 per patient), chemotherapy followed by second autoSCT (mean cost of £21,612 per patient) and best supportive care (mean cost of £13,288 per patient) [28]. Indirect costs were highlighted in two US publications. One study estimated, based on age standardised mortality rates and deaths from each group of Hodgkin Lymphoma, the annual indirect cost associated with life lost reached $3.2 billion in 2000 [29]. This was supported by another study that reported high societal costs per death due to relatively large proportions of patients of working age compared with other cancers [30].

In summary, the direct costs associated with the management of RRcHL are substantial, and are likely to increase with disease progression and continued therapy. Whilst, there is a lack of UK specific data relating to the indirect costs of RRcHL, it is clear that there are substantial costs affecting both patients and caregivers. This is driven by the age (working age) of patients and life lost, but also the substantial time and resource lost by caregivers.

3.3: Clinical pathway of care showing the context of the proposed use of the technology

In the absence of NICE guidelines for the treatment of RRcHL the recommendations of the British Committee for Standards (BCSH) in Haematology are relevant to UK clinical practice; these are described in Section 3.5.

The treatment of patients with Hodgkin Lymphoma varies according to a number of factors, including: disease stage, lymph node size, disease spread, and importantly the patient's age and general health, i.e. are they candidates for therapy, due to toxicity etc. [3, 31] (ref). As per the recommendations of BCSH, first-line therapy may include but it not limited to: doxorubicin, bleomycin, vinblastine and dacarbazine (ABVD regimen) with 20Gy radiotherapy; or bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone (BEACOPP regimen) [3]. In those patients who do not achieve long term remission salvage therapy may include chemotherapy and/or radiotherapy with the intent to enable autoSCT, which is regarded as potentially curative[3]. Following "salvage therapy" there may be a subset of patients who are ineligible for autoSCT; this is typically due a lack of clinical response, namely complete or partial response; or due to factors such as age or comorbidity that would prevent a transplant[3].

Historically, patients who fail therapy or who are ineligible for treatment following salvage chemotherapy have limited treatment options available. However, in April 2017 NICE recommended the use of BV (TA446) [32] among two patient populations; those who have: relapsed or refractory disease after autoSCT, or have relapsed or refractory disease after at least 2 previous therapies and they cannot have autoSCT or multi-agent chemotherapy. It should be noted that there are a number of international clinical guidelines [33] [34, 35], and recommendations from the SMC (Scotland)[36] and AWMSG (Wales)[37], which suggest that BV has improved the outcomes of many patients and is suitable for those patients later in the clinical pathway. However, for those who do not respond to BV the prognosis remains poor with little/ no treatment options. Although the license of BV does not preclude its use as a retreatment option, this was not included within the NICE recommendations [32], and it is highly likely that patients at this later line of therapy will be unable to tolerate the toxicity associated with traditional treatments, such as high dose chemotherapy.

As per the license indication, as reported in Section 2.2.2, it is expected that pembrolizumab would offer both patients and clinicians a much needed treatment option for those patients

who have failed to respond to BV. The benefits of pembrolizumab include but are not limited to:

- High levels of response (complete response and partial response)[38, 39]
- A favourable safety profile among a heavily pre-treated patient population [38, 39]
- A potential bridge to allogeneic stem cell transplant[40].
- A convenient and less toxic administration schedule, occurring once every three weeks for just 30 minutes, when compared with SoC (i.e. chemotherapy)[38, 39].
- Improved patient reported outcomes associated with disease related symptoms, functioning, and health status[38].

3.4: Information about the life expectancy of people with the disease or condition in England and the source of the data

Please refer to section 3.2

3.5: Details of relevant NICE guidance, pathways or commissioning guides related to the condition for which the technology is being used

As described in section 3.3, NICE do not currently have a clinical pathway available for the management of cHL. Summarised in Table 5 below are the most relevant UK clinical guidelines. However, there is still no clear consensus on the management of patients post BV. Although there is no clear consensus on the management of patients post BV, it would be logical to assume that patients continue to receive single agent/reduced intensity chemotherapy as decided on a patient by patient basis; this was supported by a recent advisory panel meeting held by MSD [40].

Table 5. Summary of relevant clinical guidelines for the treatment of Classical Hodgkin Lymphoma

Guideline group	Recommendations

British Committee for Standards in Haematology, 2014 [3]	RRcHL
	• rr to one prior line of therapy
	  ○ Salvage chemotherapy (one of ICE, IVE, MINE, IVOx, IGEV, GEM-P, GDP, Mini-BEAM, Dexa-BEAM, ESHAP, DHAP, DHAOx) for R/R patients eligible for HDT/ASCT
	  ○ Combined modality in patients ineligible for HDT/ASCT especially in early stage relapse and in patients who have not received prior RT, or who have relapsed outside of the initial radiotherapy field
	  ○ Salvage RT in selected patients ineligible for autoSCT, particularly in older patients with relapsed disease who lack B symptoms, have a good performance status, and have limited stage disease at relapse
	  ○ Chemotherapy and IFRT in patients who experience late relapse (>5 years after primary therapy) occurring at a localised site without B symptoms
	• rr to two prior lines of therapy
	  ○ HDT/ASCT in patients who achieve an adequate response to salvage therapy
	
	Post-ASCT failure
	• AlloSCT using a reduced intensity conditioning regimen who relapsed following autoSCT

London Clinical Alliance, 2015 [41]	RRcHL • Salvage chemotherapy in patients eligible for autoSCT (usually a platinum-based regimen such as GEM-P, ICE, ESHAP) • Modality therapy or salvage RT alone in patients ineligible for autoSCT Post autoSCT failure or autoSCT ineligible • Brentuximab vedotin in patients who are refractory to second-line therapy • Refractory patients may also be enrolled in clinical trials owing to a lack of further approved-treatment options

3.6: Details of other clinical guidelines and national policies

Please refer to Section 3.5.

3.7: Issues relating to current clinical practice, including variations or uncertainty about established practice

3.8: Equality issues

Not applicable

4 Clinical effectiveness

4.1 Identification and selection of relevant studies

4.1.1 Systematic Review

To address the decision problem outlined in the final NICE scope, MSD updated an existing Systematic Literature Review (SLR). This was designed to identify clinical trials and observational studies comparing the efficacy of pembrolizumab and relevant comparators for the treatment of patients with RRcHL.

An update to the Population, Interventions, Comparison, and Study Design (PICOS) statement occurred in June 2017. This amendment added additional criteria to identify patients with disease progression during or after treatment with BV, and to further restrict interventions so as to reflect relevant UK clinical practice. The PICOS statement for the review is presented in Table 6.

The updated SLR was designed to identify relevant studies to inform both direct and indirect comparisons between the interventions relevant to the NICE final scope. Further details are provided below.

Table 6. PICOS for review of treatment for RRcHL studies

Criteria	Description	
	Original SLR (Oct.19 and Dec. 2, 2016)	**Updated SLR (June 15 2017)**
Population	Adult cHL patients who either: failed to achieve a response to any line of therapy (refractory patients) or who have relapsed after ≥ 3 prior lines of therapy	Additional criteria added to restrict patients to those with disease progression during or after treatment with BV
Interventions	*The following targeted drugs alone or as combinations with systemic chemotherapies:* • Pembrolizumab • Brentuximab vedotin • Everolimus • Lenalidomide • Lucatumumab • Nivolumab • Ofatumumab • Panobinostat • Rituximab • Vorinostat *The following systemic chemotherapies alone or in combinations:* • Adriamycin • Bendamustine • Bleomycin • Carmustine • Cisplatin • Cyclophosphamide • Cytarabine • Dacarbazine • Etoposide • Gemcitabine • Ifosfamide • Mecholrethamine (Nitrogen mustard) • Melphalan • Mitoxantrone • Oxaliplatin • Procarbazine • Vinblastine • Vincristine • Vinorelbine *Other treatments in combination with chemotherapies:* • Dexamethasone • Methylprednisolone • Prednisone	*Additional criteria were added to reflect only those interventions considered relevant to UK clinical practice:* • Single or combination chemotherapy including drugs such as: ○ Cisplatin ○ Gemcitabine ○ Vinblastine • Best supportive care
Comparators	• Any	• Any
Outcomes	• Overall survival • Progression-free survival • Objective response • Complete response • Partial response • Treatment discontinuation due to AEs • Serious (grade 3 and above) AEs (not used for study selection)	• No change
Study design	• Randomised controlled trials • Non-randomised controlled trials • Single arm trials • Retrospective and prospective controlled observational studies • Single group observational studies	• No change

Abbreviations: AEs, adverse events; BV, brentuximab vedotin; cHL, classical Hodgkin's Lymphoma

4.1.2 Search strategy description

Separate searches were conducted for clinical trials and observational studies in the following databases using the OVID portal: Embase, MEDLINE, and Cochrane Register of Controlled Trials (clinical trials only). The Scottish Intercollegiate Guidelines Network's (SIGN) filters for randomised-controlled trials and observational studies were used in the Embase and Medline searches. The primary searches for clinical trials and observational studies were conducted on October 19 and December 2, 2016, respectively. All searches were then rerun on June 15, 2017 with terms added to restrict hits to those published in the period in between this date and the date which the primary searches were run. The full search strings are summarised in Appendix 2

The database searches were supplemented with searches of the Northern Light database that contains conference proceedings from 2010 to the present. The annual meetings of the American Society of Clinical Oncology from 2015 to 2016 and American Society of Haematology from 2014 to 2016 were searched using this database (the search strategies from each conference can be found in Appendix 2). In addition, manual searches were conducted of the WHO International Clinical Trials Registry Platform (ICTRP) to identify on-going trials (the terms searched are presented in Appendix 2).

4.1.3 Study selection

Two investigators working independently reviewed all abstracts and proceedings identified; the eligibility criteria used in the search strategy are outlined in Table 7. All citations identified as potentially relevant during abstract screening were then screened as full texts by the same two reviewers. Following reconciliation between the two investigators, a third investigator was included to reach consensus for any remaining discrepancies. Full articles were retrieved for further detailed assessment by the same reviewers. Discrepancies occurring between the two investigators were resolved by involving a third investigator and reaching consensus.

Once the list of included studies was finalised two investigators working independently extracted data for the final list of included studies. Any discrepancies observed between the data extracted by the two data extractors were resolved by involving a third reviewer and coming to a consensus. Extraction data included, but was not limited to, study characteristics, intervention details, patient baseline characteristics, outcomes, and quality assessment.

Table 7. Eligibility criteria used in the search strategy

	Inclusion criteria
Population	Adult cHL patients who either: failed to achieve a response to any line of therapy (refractory patients) or who have relapsed after ≥ 3 prior lines of therapy
Intervention	Pembrolizumab
Comparators	Single or combination chemotherapy including drugs such as: • Cisplatin • Gemcitabine • Vinblastine Best supportive care
Outcomes	Overall survival Progression-free survival Objective response Complete response Partial response Treatment discontinuation due to AEs Serious (grade 3 and above) AEs (not used for study selection)
Study design	Randomised controlled trials Non-randomised controlled trials Single arm trials Retrospective and prospective controlled observational studies Single group observational studies

4.1.4 Flow diagram of the number of studies included and excluded at each stage

Clinical trials

Original search

A total of 10,359 citations were identified through the primary clinical trial searches of Embase, Medline, and Cochrane Central Register of Controlled Trials. Of these, 2,949 were removed as duplicates, with a further 7,312 excluded during abstract screening. From the 99 citations included for full text screening, 60 were excluded: 17 for study design, 28 for population, four for outcomes, four for interventions, and seven for other reasons (e.g. publications were letters to the editor). A list of these excluded studies can be found in Appendix 3. Adding data from a clinical study report (CSR) for KEYNOTE-087 therefore gave a total of 39 citations included as a result of the primary searches. After applying the updated criteria from the relevant to the NICE decision problem, 38 of these citations were removed. The one citation that was still

relevant was the CSR for KEYNOTE-87. A list of the citations that were excluded due to the updated PICOS can be found in Appendix 3.

Updated search

For the updated search a total of 1,497 citations were identified (981 electronic databases, and 516 additional sources). After the removal of duplicates (190 citations) and screening abstracts, 14 citations were included for full-text screening. From these, 11 citations were excluded: one for population, two for outcome, two for study design and six for other (all captured in the original search). The updated search only yielded one new citation to be added into the evidence base [38]. After screening the conference proceeding of ASCO 2015-2016 and ASH 2014-2016 ASH Annual Meeting, two additional citations were added [42, 43], resulting in three citations being included from the updated search. This gave a total of four citations, representing one clinical trial (KEYNOTE-087), being included in the final review. The flow of study selection is presented using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) chart in Figure 3.

Observational studies

The primary search for observational studies was carried out using Embase and Medline via the OVID platform (n=163). Following the removal of duplicates a total of 131 citations were screened, and resulted in the inclusion of a single citations [44]. The search update resulted in three studies being included for full text screening but none of these met the inclusion criteria. After screening the conference proceedings (ASH 2014-2016 and ASCO 2015-2016), no additional citations were added. Furthermore, when screening the updated searches for both clinical trials and for observational studies, it was noted that some observational studies were not being captured in the observational search but were captured in the clinical trial search. The decision was made to re-screen the citations that were excluded by study design in the clinical trials searches (primary and update) though this did not identify and additional studies. Therefore, a single retrospective study by Cheah et al. 2016 was included [44]. The flow of study selection is presented using the PRISMA chart in Figure 4.

Figure 3. PRISMA flow diagram of clinical trials - Original and updated SLR

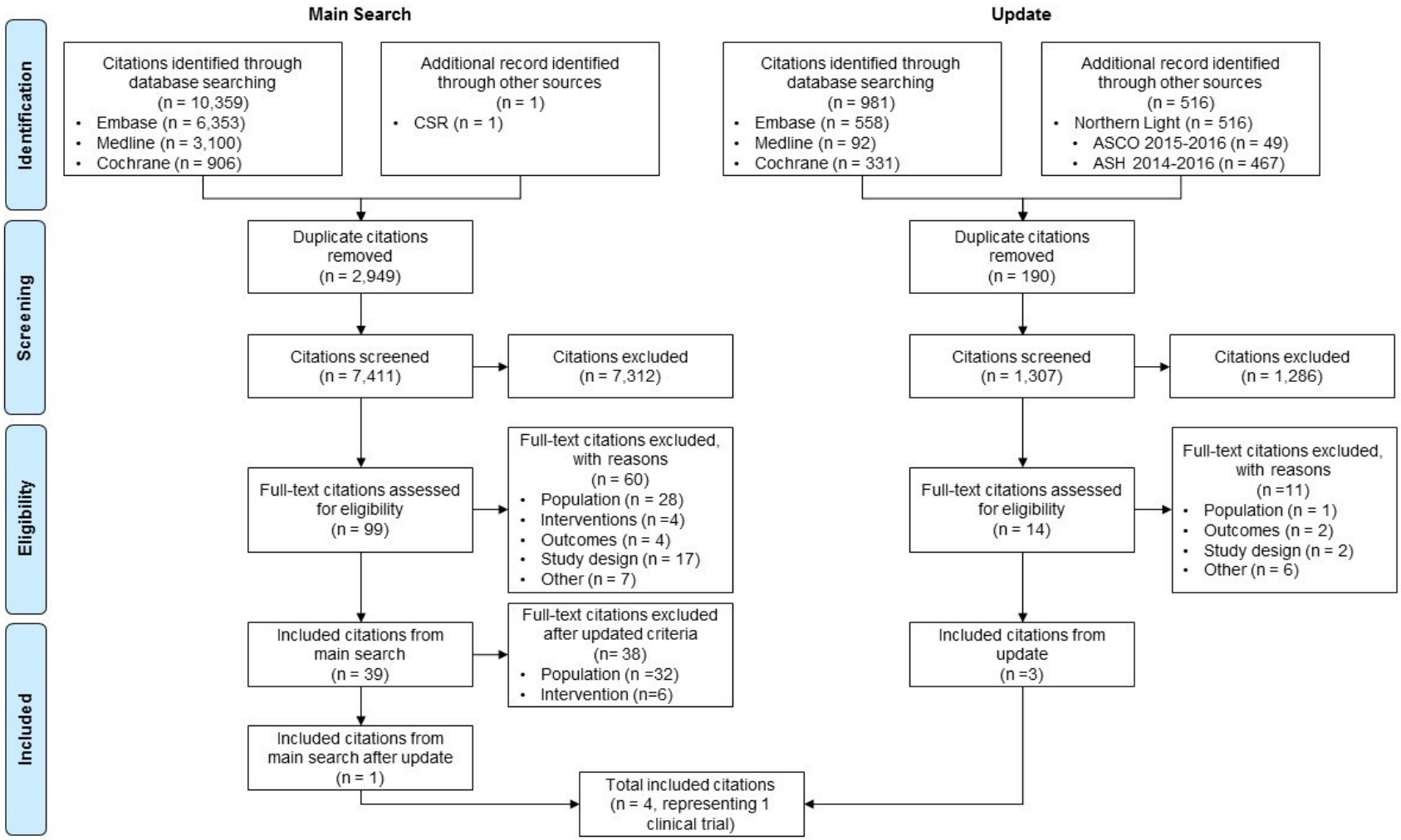
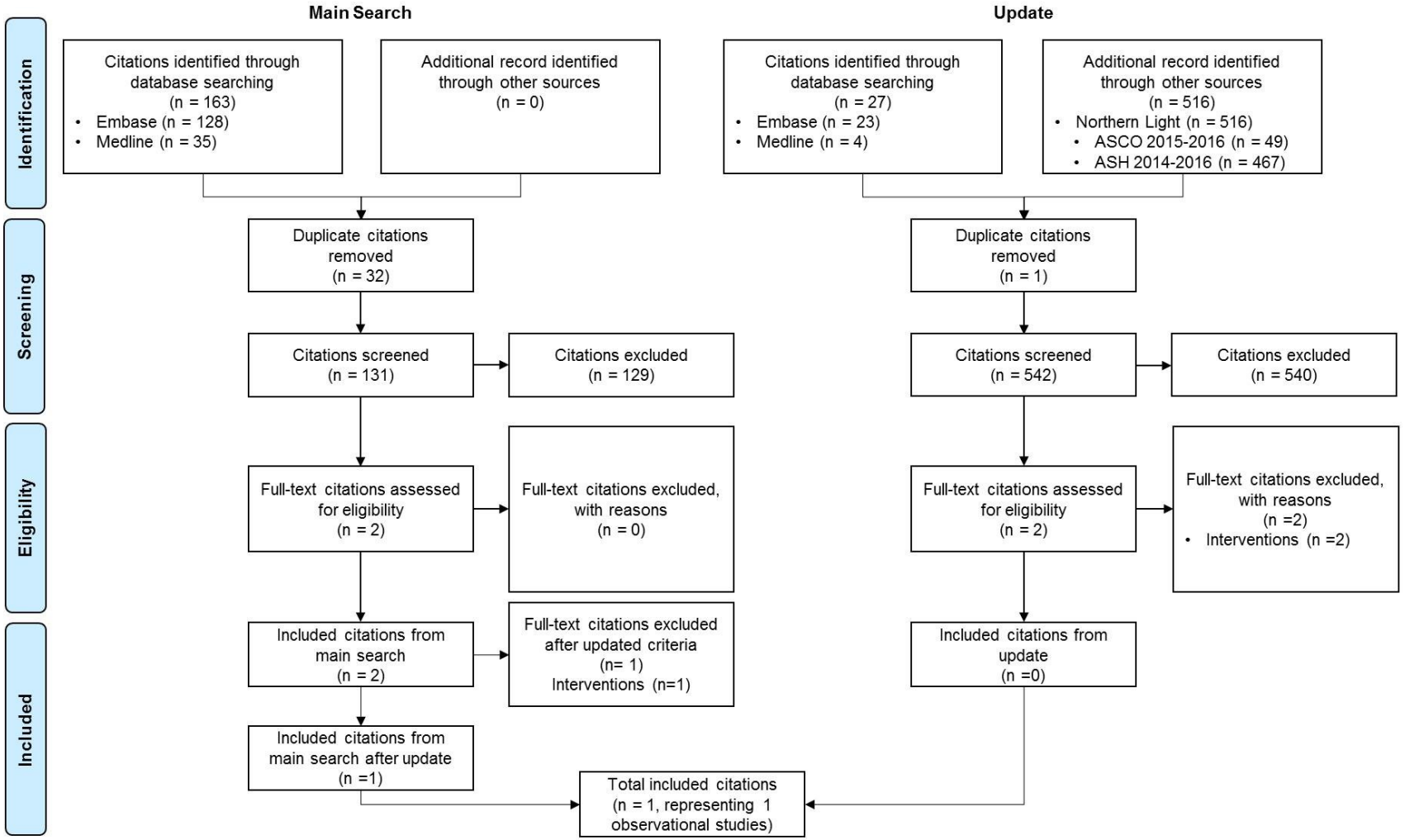


Figure 4. PRISMA flow diagram of observational studies - Original and updated SLR



4.1.5 Complete list for excluded studies

A complete reference list for excluded studies (and the reason for exclusion) has been provided in Appendix 3.

4.2 List of relevant randomised controlled trials

No randomised control evidence relevant to the decision problem was identified. The pivotal clinical trial, KEYNOTE-087, is a multicentre, single arm, multi-cohort, non-randomised trial. Therefore, Sections 4.3-4.7 relate to this pivotal clinical trial.

4.2.1 List of relevant clinical trials/ observational studies

In total two studies relevant to the decision problem were identified; this included a single arm phase II clinical trial (KEYNOTE-087) [45-48], and one retrospective observational study [44]. As no evidence providing direct comparative evidence for pembrolizumab versus comparators exists, the observational study identified within the SLR was used in the naïve and matched adjusted indirect comparisons as described Section 4.10.

4.3 Summary of methodology of the relevant trials

4.3.1 KEYNOTE-087

Trial design

KEYNOTE-087 (NCT02453594) is a phase II, multicentre, single arm, multi-cohort, non-randomised trial of pembrolizumab in patients with RRcHL. The three study cohorts included patients with RRcHL, who have failed to achieve a response or progressed after autologous stem cell transplant (auto-SCT) and have relapsed after treatment with, or failed to respond to, brentuximab vedotin post auto-SCT (Cohort 1); who were unable to achieve a complete response (CR) or partial response (PR) to salvage chemotherapy and did not receive auto-SCT, but have relapsed after treatment with, or failed to respond to, brentuximab vedotin (Cohort 2); and subjects who have failed to respond to, or progressed after, auto-SCT and have not received brentuximab vedotin post auto-SCT. These patients may or may not have received brentuximab vedotin as part of primary or salvage treatment (Cohort 3).

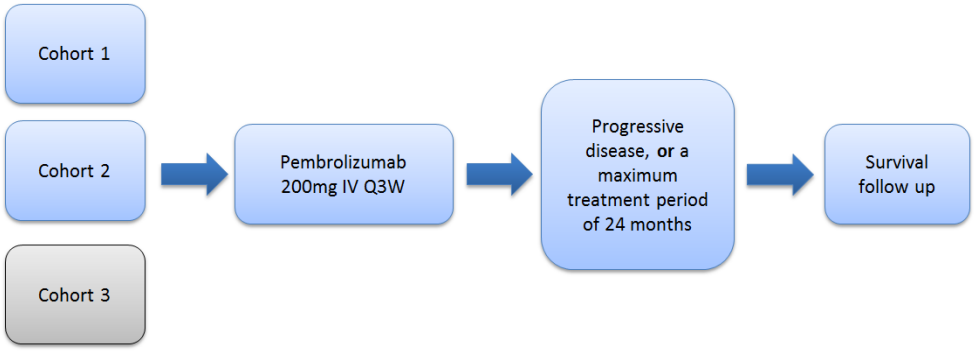
The rationale for selecting a single arm non-comparative trial is largely based on the absence of established clinical practice at this later line setting, as discussed in Section 3.5, and the limited number of eligible patients for treatment.

Please note that this submission focusses on data derived from Cohorts 1 and 2 of KEYNOTE-087; this supports the EMA license recommendation as reported in Section 2.2.2.

Patients who experienced a CR or PR or had SD were able to remain on treatment for up to 2 years (approximately 37 administrations) or until unacceptable toxicity or progression. After documented disease progression or the start of new antineoplastic therapy each patients was to be followed by telephone for overall survival (OS) until death, withdrawal of consent, or the end of the study, whichever occurred first.

At the investigators discretion patients who attained a CR may have been considered for stopping pembrolizumab after receiving a minimum of 24 weeks months of treatment with at least two doses since CR had been initially confirmed. Patients who later experienced disease progression would have been eligible for retreatment with pembrolizumab at the discretion of the investigator if: no cancer treatment was administered since the last dose of pembrolizumab, the subject met the safety parameters listed in the inclusion/exclusion criteria, and the trial was open. Patients would have resumed therapy at the same dose and schedule as at the time of initial discontinuation.

Figure 5. KEYNOTE-087 Study design



Note. Cohort 1: who failed to achieve a response or progressed after autoSCT and have relapsed after treatment with or failed to respond to BV. Cohort 2: who were unable to achieve a CR or PR to salvage chemotherapy and did not received autoSCT and have relapsed after treatment with or failed to respond to BV. Cohort 3: who failed to respond to or progressed after autoSCT and have not received BV post autoSCT. These patients may or may not have received BV as part of primary salvage treatment. Please note that cohort 3are greyed out, and due to license requirement of required BV use, are not relevant to the submission decision problem.

Eligibility criteria

The key inclusion/ exclusion criteria are provided below.

Key inclusion criteria:

In order to be eligible for participation in this trial, the subject had to:

- Be ≥18 years of age on day of signing informed consent.
- Have relapsed* or refractory* de novo classical Hodgkin Lymphoma and meet one of the following cohort inclusions:

  *Relapsed: disease progression after most recent therapy
  *Refractory: failure to achieve CR or PR to most recent therapy

  - Cohort 1: Have failed to achieve a response or progressed after autoSCT. Patients must have relapsed after treatment with or failed to respond to BV post autoSCT.
  - Cohort 2: Were unable to achieve a CR or a PR to salvage chemotherapy and did not receive autoSCT. Patients must have relapsed after treatment with or failed to respond to BV.
  - Cohort 3: Have failed to achieve a response or progressed after autoSCT and have not have received BV post autoSCT. Note: These patients may or may not have received BV as part of primary treatment, or salvage treatment.

- Have measureable disease defined as at least one lesion that can be accurately measured in at least two dimensions with spiral computerised tomography (CT) scan. Minimum measurement must be >15 mm in the longest diameter or >10 mm in the short axis.

- Be able to provide an evaluable core or excisional lymph node biopsy for biomarker analysis from an archival or newly obtained biopsy at Screening. In addition patients may provide additional biopsy at Week 12 and at the time of discontinuation due to progression. If submitting unstained cut slides, freshly cut slides should be submitted to the testing laboratory within 14 days from when the slides are cut
- Must have a performance status of 0 or 1 on the Eastern Cooperative Oncology Group (ECOG) Performance Scale

Key exclusion criteria

Patients were excluded from participating in the trial if they met any of the following key criteria:

- Has a diagnosis of immunosuppression or is receiving systemic steroid therapy or any other form of immunosuppressive therapy within 7 days prior to the first dose of trial treatment. The use of physiologic doses of corticosteroids may be approved after consultation with the Sponsor
- Has undergone prior alloSCT within the last 5 years. Patients who have had a transplant greater than 5 years ago are eligible as long as there are no symptoms of graft vs. host disease
- Has a known additional malignancy that is progressing or requires active treatment. Exceptions include basal cell carcinoma of the skin, squamous cell carcinoma of the skin, or in situ cervical cancer that has undergone potentially curative therapy.
- Has evidence of active, non-infectious pneumonitis
- Has an active infection requiring intravenous systemic therapy
- Is pregnant or breastfeeding, or expecting to conceive or father children within the projected duration of the trial, starting with the pre-screening or screening visit through 180 days after the last dose of trial treatment.
- Has received prior therapy with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-cytotoxic T-lymphocyte associated antigen-4 antibody (including ipilimumab or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways).

Settings and Location where the data were collected

This was a global study enrolling a total of 210 patients (cohort 1, n=69; cohort 2, n=81; cohort 3, n=60) between the 26th June 2015 and March 21st 2016 across 51 study sites. This included three study sites in the UK, 23 sites across Europe (France, Russia, Italy, Spain, Germany, Greece, Hungary, Sweden, and Norway), eleven in the USA, seven in Japan, four in Israel, two in Australia, and one in Canada.

There were 14 patients (Cohort 1, n=4; Cohort 2, n=10) enrolled from three UK study sites.

Trial drugs and concomitant medication

This was an open label trial, therefore the sponsor, investigator, and patient knew the treatment administered. All trial treatment was administered in the outpatient setting by qualified site personnel.

All patients received pembrolizumab 200mg via intravenous (IV) infusion as a 30 minute infusion every 3 weeks in the outpatient setting (Table 8). Treatment could be administered up to 3 days before or after the scheduled Day 1 of each cycle for administrative reasons. Interruptions from the treatment plan for greater than 3 days and up to 3 weeks were allowed, but required consultation between the Investigator and Sponsor, and written documentation of the collaborative decision on subject management. Neither dose escalation nor dose reduction of pembrolizumab was permitted in this trial.

Dose modification due to adverse events (AE) (both non-serious and serious) was permitted as outlined in Section 5.2.1.2 (page 36 of 130) of the KEYNOTE-087 protocol [49], as exposure with pembrolizumab may represent an immunological aetiology. These AEs may occur shortly after the first dose or several months after the last dose of treatment.

Table 8. Keynote-087 trial treatment

Study Drug	Dose/Potency	Dose Frequency	Route of Administration	Regimen/Treatment Period	Use
Pembrolizumab	200mg	Q3W	IV Infusion	Day 1 of each treatment cycle	experimental

Concomitant medication

Concomitant medication and or vaccination specifically prohibited in the study exclusion criteria were not allowed during the on-going trial. The decision on any supportive therapy or vaccination resided with the investigator and/or the patient's primary physician. However, the decision to continue the patient on pembrolizumab required the mutual agreement of the investigator, sponsor, and subject.

All treatments that the investigator considers necessary for a subject's welfare may be administered at the discretion of the investigator in keeping with the community standards of medical care. All concomitant medication including all prescription, over-the-counter, herbal supplements, and IV medications and fluids was recorded on the case report form. If changes

to medication occurred during the trial period, documentation of drug dosage, frequency, route, and date may also be included on the case report form. Patients were able remain on anti-coagulation therapy as long as the prothrombin time or activated partial thromboplastin time is within therapeutic range of the intended use of anticoagulants.

All concomitant medications received within 28 days before the first dose of trial treatment and 30 days after the last dose of trial treatment was recorded. Prohibited concomitant medications included:

- Antineoplastic systemic chemotherapy or biological therapy
- Granulocyte macrophage colony-stimulating factor
- Immunotherapy not specified in the protocol
- Chemotherapy not specified in the protocol
- Investigational agents other than pembrolizumab
- Radiation therapy
    - Any need for radiotherapy was considered indicative of progressive disease and resultant in discontinuation of study therapy.
- Live vaccines within 30 days prior to the first dose of trial treatment and while participating in the trial. Examples of live vaccines include, but are not limited to, the following: measles, mumps, rubella, chicken pox, yellow fever, rabies, BCG, and oral typhoid vaccine. Seasonal influenza vaccines for injection are generally killed virus vaccines and are allowed; however intranasal influenza vaccines (e.g. Flu-Mist®) are live attenuated vaccines, and are not allowed.
- Glucocorticoids for any purpose other than to modulate symptoms from an event of clinical interest of suspected immunologic aetiology.

Primary, Secondary, and Exploratory objectives

Primary objectives

The primary efficacy endpoint is the Overall Response Rate (ORR), defined as the proportion of patients in the analysis population who have complete remission (CR) or partial remission (PR) using IWG criteria (Cheson 2007)[50] at any time during the study. Response for the primary analysis was determined by blinded, independent central review (BICR), for the overall population and each cohort as outlined in Figure 5. A co-primary endpoint for safety and tolerability is reported in Section 4.12.

Secondary objectives

Within each of the three cohorts as described in Figure 5 the following objectives were considered:

- To evaluate the ORR of pembrolizumab by investigator assessment according to the IWG response criteria; and additionally by BICR using the 5-point scale according to the Lugano Classification [51].
- To evaluate the Complete Remission Rate (CRR) of pembrolizumab by BICR and by investigator assessment according to the IWG response criteria; and additionally by BICR using the 5-point scale according to the Lugano Classification
- To Evaluate Progression Free Survival (PFS) and Duration of Response (DOR) of pembrolizumab by BICR and by investigator assessment according to the IWG response criteria.
- To evaluate the overall survival (OS) of pembrolizumab

Duration of response is defined, only for the subgroup of patients who achieve CR or PR, as the time from start of the first documentation of objective tumour response (CR or PR) to the first documentation of tumour progression or death due to any cause, whichever comes first.

Exploratory end points

Within each of the three cohorts, and potentially pooled as described in Figure 5 the following objectives were considered:

- To evaluate ORR, CRR, PFS and DOR for patients who continue treatment with pembrolizumab beyond documented progression.
- To explore the pharmacokinetic (PK) profile of pembrolizumab.
- To evaluate changes in health-related quality-of-life assessments from baseline using the European Organization for Research and Treatment of Cancer (EORTC) Quality

of Life (QoL) Questionnaire C30 (QLQ-C30) and European Quality of Life Five Dimensions Questionnaire (EuroQoL EQ-5D).

- To further evaluate pembrolizumab immunogenicity and exposure of the proposed dose and dosing regimen.
- To compare the extent of pre-pembrolizumab PD-L1 expression in tumour biopsies for pembrolizumab responders versus non-responders.
- To investigate the relationship between candidate efficacy biomarkers and anti-tumour activity of pembrolizumab utilising pre and post-treatment lymph node biopsies and blood sampling.
- To explore the relationship between genetic variation and response to the treatment(s) administered. Variation across the human genome will be analysed for association with clinical data collected in this study.

Clinical procedures/ assessments

Biomarker collection

All patients were required to have either an archival formalin-fixed paraffin embedded (FFPE) tumour tissue sample or newly obtained core or excisional biopsy (FNA not adequate) to be submitted for characterisation at a central lab. Biopsy sites were to be selected so that subsequent biopsies can be performed at the same location.

Initial tumour imaging

Initial disease assessment or tumour imaging must have been performed within 28 days prior to the first dose of trial treatment. The site study team must have reviewed pre-trial images to confirm the subject had measurable disease as defined in the inclusion criteria In addition bone marrow biopsies were collected at screening.

Tumour imaging and assessment of disease

Tumour imaging could be performed using computer tomography (CT) and positron emission tomography (PET) and should have been used throughout the study. For Lymphomas that were not fluorodeoxyglucose-avid (FDG-avid) at screening, PET did not need to be repeated in follow-up assessments. Following screening, CT scans should have been repeated every 12 weeks for subsequent assessments. PET should have been repeated at Week 12, Week 24, to confirm CR or PD and as clinically indicated.

Assessment of disease

Anti-tumour activity of pembrolizumab was evaluated using the IWG response criteria by CT/PET.

The criteria were applied by the site as the primary measure for assessment of disease response and as a basis for all protocol guidelines related to disease status (e.g. discontinuation of study therapy). Disease response assessments were to occur every 12 weeks until documented disease progression, the start of new anti-cancer treatment, withdrawal of consent, death, or the end of the study, whichever occurred first. Assessment of Lymphoma B symptoms occurred with each Lymphoma disease response assessment. Anti-tumour activity of pembrolizumab was also evaluated by BICR as part of the exploratory analyses using the 5-Point-Scale per the Lugano Classification.

- Revised Response Criteria for Malignant Lymphoma. (Cheson et al, J Clin Oncol, 2007)[50]

- 5 Point-Scale per the Lugano Classification (Cheson et al, J Clin Oncol, 2014)[51]

Bone marrow biopsies were collected to confirm complete remission (in patients who had bone marrow involvement) or if clinically indicated. Blood for correlative biomarkers studies was to be collected at Screening, Week 12, and upon PD.

Immunotherapeutic agents such as pembrolizumab may produce antitumour effects by potentiating endogenous cancer-specific immune responses, which may be functionally anergic. The response patterns seen with such an approach may extend beyond the typical time course of responses seen with cytotoxic agents, and can manifest a clinical response after an initial increase in tumour burden or even the appearance of new lesions. Standard response assessment criteria may not provide a complete response assessment of immunotherapeutic agents such as pembrolizumab. Therefore in the setting where a subject assessment shows PD, study drug may have been continued, at the discretion of the PI, until the next disease response assessment provided that the patients' clinical condition was stable. However, imaging should have occurred at any time where there was clinical suspicion of progression.

After the first documentation of progression it is at the discretion of the investigator to keep a clinically stable subject on trial treatment or to stop trial treatment until repeat imaging performed 4-6 weeks later confirms progression. Clinical Stability may be defined as:

1) Absence of symptoms and signs indicating clinical significant progression of disease (including worsening of laboratory values) indicating disease progression.
2) No decline in ECOG performance status.

3) Absence of rapid progression of disease or progressive tumour at critical anatomical sites (e.g., cord compression) requiring urgent alternative medical intervention.

Patient reported outcomes

The EuroQol EQ-5D and EORTC QLQ-C30 questionnaires were administered by trained site personnel and completed electronically by the patients themselves. It was strongly recommended that all electronic PROs were administered prior to drug administration, AE evaluation and disease status notification. The electronic PROs were completed in the following order: EuroQol EQ-5D first, then EORTC QLQ-C30. PROs were to be assessed every cycle for the first five cycles and every 12 weeks thereafter until PD while the subject was receiving study treatment. PROs were to also be obtained at the Treatment Discontinuation Visit and 30-day Safety Follow-up Visit. If the Treatment Discontinuation Visit occurred 30 days from the last dose of study treatment, at the time of the mandatory Safety Follow up Visit, then the PROs did not need to be repeated.

Safety measurements

Vital signs, weight, physical examinations, ECOG performance status, electrocardiogram, and laboratory safety tests (e.g., urinalysis, complete blood count, prothrombin time/activated partial thromboplastin time, serum chemistries, auto-antibodies, thyroid function) were obtained and assessed at designated intervals throughout the study.

Adverse event/ adverse experience was defined as any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment. An AE can therefore be any unfavourable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a medicinal product or protocol specified procedure, whether or not considered related to the medicinal product or protocol specified procedure. Any worsening (i.e., any clinically significant adverse change in frequency and/or intensity) of a pre-existing condition that is temporally associated with the use of pembrolizumab was also considered as an AE.

A serious adverse experience is any adverse experience occurring at any dose or during any use of Sponsor's product that:

- Results in death;
- Is life threatening;
- Results in a persistent or significant disability/incapacity;
- Results in or prolongs an existing inpatient hospitalisation;

- Is a congenital anomaly/birth defect;
- Other important medical events.

In addition, the following events, although not serious per ICH definition, were reportable to the sponsor in the same timeframe as SAEs to meet certain local requirements. Therefore, these events are considered serious by the Sponsor for collection purposes:

- Is a cancer;
- Is associated with an overdose (this was defined as any dose exceeding 1000mg or greater for pembrolizumab).

Please see the KEYNOTE-087 study protocol for a full study flow chart that outlines all scheduled tests and assessments as described above page 47/130, Section 6.1 [49].

4.4 Statistical analysis and definition of study groups in the relevant trials

4.4.1 KEYNOTE-087[48, 49]

Primary hypothesis

Intravenous administration of single agent pembrolizumab will result in an ORR of greater than 20% in each of the three cohorts using IWG response criteria by BICR.

As per the earlier definition of ORR, final analysis will be conducted for each cohort when the last subject in that cohort has reached the Week 12 response assessment or has discontinued study therapy. Results will be presented as the point estimate and 95% 2-sided exact confidence interval (CI) using the Clopper-Pearson method which will have at least 95% coverage of the true rate. An exact binomial test will be conducted for each cohort versus a fixed control rate for each cohort.

Secondary hypotheses

Note that secondary objectives and explorative endpoints within each cohort will not involve hypothesis testing.

Secondary analyses for ORR will be performed based on investigator's (i.e. study site) assessment and by central review based on the Lugano Classification [51]. As per the primary hypothesis results will be presented as the point estimate and 95% 2-sided exact CI, separately per cohort. Additional analyses will be based on site assessment and by central review using the Lugano (2014) criteria [51].

The median overall survival, if reached, will be estimated in the given analysis population, separately by cohort. In addition, the Kaplan-Meier method will be used to estimate the survival curve, separately by cohort.

Duration of response (DOR) analysis will consist of Kaplan-Meier estimates. Duration of response data will be censored on the date of the last disease assessment documenting absence of progressive disease for patients who do not have tumour progression and are still on study at the time of an analysis, are given antitumour treatment (including stem cell transplant) other than the study treatment, or are removed from study prior to documentation of tumour progression. Duration of Response will be based upon central review according to the IWG criteria; a secondary analysis of DOR will be conducted using investigator

assessment. In addition, since stem cell transplant post-initiation of pembrolizumab is considered to be an indicator of positive efficacy rather than failure of the current treatment.

Analysis and stopping guidelines

Efficacy analysis population

The analysis of primary efficacy endpoints were based on the ASaT population, i.e., patients will be included if they receive at least one dose of study medication. Supportive analyses were conducted using the Full Analysis Set (FAS) population, which consisted of all patients who 1) received at least one dose of study medication; 2) had a baseline disease assessment, and 3) had a post baseline disease assessment OR discontinued the trial due to progressive disease/drug related AE.

Safety analysis population

The ASaT population was used for the analysis of safety data in this study. The ASaT population consists of all enrolled patients who received at least 1 dose of study treatment. At least one laboratory or vital sign measurement obtained subsequent to at least one dose of trial treatment was required for inclusion in the analysis of each specific parameter. To assess change from baseline, a baseline measurement was also required.

Sample size

Efficacy for each cohort was analysed separately and pooled. The proposed sample size for each of the three cohorts was 60 patients in the primary analysis population (ASaT), i.e. 180 patients in total. To obtain 180 total patients in the ASaT population, the protocol outlined that 190 patients would need to be enrolled, assuming that approximately 5% of enrolled patients are not treated. With 60 patients per cohort in the primary analysis population, there would be at least 93% statistical power (1-sided nominal 2.5% alpha) to detect a 40% or higher ORR for the pembrolizumab arm compared to a fixed control rate of 20% using the exact binomial test (nQuery version 2.0 software). Success for this hypothesis required at least 16/60 responses. If an interim analysis is performed within a cohort the power will be approximately 92%.

The selection of 20% as a fixed control rate was based partly on historical data in previously conducted studies in R/R Hodgkin Lymphoma prior to the approval of BV, where response rates ranged between 18%-53% (Johnston et al, 2010 [52], Feninger et al, 2011 [53], Younes et al, 2012 [54], and Moscowitz et al, 2012 [55]). However, since this study was conducted in patients who had failed treatment with BV, and to date there is no published data on the ORR in this

particular patient population. Thus, a 20% ORR may be taken as a conservative control rate considering that all patients to be enrolled in this study have failed an additional line of therapy (BV) than seen previously.

Multiplicity

The false positive rate for testing the primary efficacy endpoint was controlled at 0.025 (1-sided) within each cohort. No additional multiplicity adjustment was required because each cohort was evaluated independently.

Table 9 .Summary of Efficacy analysis for Primary and Secondary efficacy endpoints

Endpoint/Variable	Statistical Method	Analysis Population	Missing Data Approach
Primary:			
Overall Response Rate • IWG criteria (2007) ○ Central review	Exact test of binomial parameter; 2-sided 95% exact CI	ASaT/FAS	Subjects with missing data are considered non-responders
Secondary:			

Overall Response Rate
• IWG criteria (2007)
  ○ Study site
• Lugano criteria (2014)
  ○ Central review
	Point estimate; 2-sided 95% exact CI	ASaT/FAS	Subjects with missing data are considered non-responders

Complete Remission Rate
• IWG criteria (2007)
  ○ Central review
  ○ Study site
• Lugano criteria (2014)
  ○ Central review
	Point estimate; 2-sided 95% exact CI	ASaT/FAS	Subjects with missing data are considered non-responders

Progression-free survival
• IWG criteria (2007)
  ○ Central review
  ○ Study site
	Summary statistics using Kaplan-Meier method	ASaT/FAS	Censored at last assessment

Duration of Response
• IWG criteria (2007)
  ○ Central review
  ○ Study site
	Summary statistics using Kaplan-Meier method	All responders	Non-responders are excluded in analysis

Overall survival	Summary statistics using Kaplan-Meier method	ASaT/FAS	Censored at last assessment

Abbreviations: ASaT, all subjects as treated; FAS, full analysis set, CI, confidence intervals

4.5 Participant flow in the relevant trials

4.5.1 KEYNOTE-087

Number of patients eligible to enter trial

A total of 210 patients were enrolled into the KEYNOTE-087 trial and were included in the ASaT analysis population (n=210). Of relevance to this submission are Cohort 1 (n=69) and Cohort 2 (n=81). The first patient enrolled in the study in Cohort 1 was on 24th June 2015, and the last patient enrolled was on 8th February 2016. The first patient enrolled in the study in Cohort 2 was 24th June 2015, and the last patient enrolled was on 16th December 2015

Information relating to subject enrolment and baseline characteristics, as reported in Table 11, are reported from the June 2016 CSR [48]. However, the most recent efficacy update report, relevant to this submission, is based on a database cut off 21st March 2017 [45, 46].

At the time of data cut-off, 21st March 2017, there were [redacted] patients who remained on treatment from Cohorts 1 and 2, respectively (Table 10) [46]. Within Cohort 1 a total of 38 patients had discontinued therapy, of which 23.2% (n=16) was due to disease progression, 11.6% (n=8) due to complete response, and 8.7% (n=6) due to an AE; additional detail is provide below in Table 10. Within Cohort 2 a total of 60 patients had discontinued therapy, of which 37% (n=30) was due to disease progression, 9.9% (n=8) due to complete response, and 8.6% (n=7) due to physician decision (Table 10).

Table 10. Keynote-087 Subject disposition – All subjects (ASaT)[46]

	Cohort 1	Cohort 2
Subjects in population	69	81
Status for study medication in trial segment treatment		
Started	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]
Bone marrow transplant	[redacted]	[redacted]
Clinical progression	[redacted]	[redacted]
Complete response	[redacted]	[redacted]
Death	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]
Physicians Decision	[redacted]	[redacted]
Pregnancy	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]
Withdrawal by subject	[redacted]	[redacted]
Treatment on-going	[redacted]	[redacted]

Characteristics of participants at baseline[48]

Baseline characteristics for Cohorts 1 and 2 are reported in Table 11. Within cohort 1 there were slightly more male (52.2%) than female patients, and were predominantly White (82.6%). The median age was 34 years (range 19 to 64 years). At study entry the main disease subtype reported was cHL – nodular sclerosis (79.7%). Of the 69 patients in Cohort 1 virtually all patients reported an ECOG score of 0 (42%) or 1 (56.5%) with 31.9% reporting B symptoms; bone marrow involvement was low (4.3%). Overall, the patients of Cohort 1 were heavily pre-treated having received at least 3 lines of prior therapy, and a median of 4 lines (range 2 to 12). All patients had received prior treatment with BV with a mean time of relapse since autoSCT failure of 60 months.

Cohort 2, as summarised in Table 11, there were slightly more male (53.1%) than female patients, and were predominantly White (90.1%). The majority of patients were aged less than 65 years with a median age of 40 years (range 20 to 76 years). At study entry the majority of patients reported a diagnosis of cHL – Nodular sclerosis (80.2%). Of the 81 patients in Cohort 2 all patients reported an ECOG score of 0 (54.3%) or 1 (45.7%) with 32.1% reporting B symptoms; bone marrow involvement was low (6.2%). Overall, the patients of Cohort 2 were

heavily pre-treated with the majority having received at least 3 lines of prior therapy (96.3%), and a median of 4 lines (range 1 to 11). All patients had received prior treatment with BV.

Table 11. KEYNOTE-087 subject Characteristics of trial population [48]

	Cohort 1 n (%)	Cohort 2 n (%)
Subject in population	69	81
Gender		
Male	36 (52.2)	43 (53.1)
Female	33 (47.8)	38 (46.9)
Age Years		
<65	69 (100)	66 (81.5)
>=65	0	15 (18.5)
Mean	37.0	42.3
SD	10.9	17.4
Median	34.0	40
Range	19 to 64	20 to 76
Race		
American Indian or Alaska native	0	1 (1.2)
Asian	7 (10.1)	4 (4.9)
Black or African American	2 (2.9)	2 (2.5)
Missing	1 (1.4)	1 (1.2)
Multi-racial	2 (2.9)	0
White	57 (82.6)	73 (90.1)
Ethnicity		
Hispanic or Latino	6 (8.7)	5 (6.2)
Not Hispanic or Latino	48 (69.6)	65 (80.2)
Not reported	4 (5.8)	7 (8.6)
Unknown	11 (15.9)	4 (4.9)
Race Group		
White	57 (82.6)	73 (90.1)
Non-White	11 (15.9)	7 (8.6)
Missing	1 (1.4)	1 (1.2)
US region		

	Cohort 1 n (%)	Cohort 2 n (%)
US	13 (18.8)	20 (24.7)
Ex-US	56 (81.2)	61 (75.3)
Disease Subtype		
Classical Hodgkin Lymphoma – Nodular sclerosis	55 (79.7)	65 (80.2)
Classical Hodgkin Lymphoma – Mixed cellularity	9 (13.0)	10 (12.3)
Classical Hodgkin Lymphoma – Lymphocyte rich	4 (5.8)	1 (1.2)
Classical Hodgkin Lymphoma-Lymphocyte depleted	0	4 (4.9)
Missing	1 (1.4)	1 (1.2)
ECOG performance status		
0	29 (42.0)	44 (54.3)
1	39 (56.5)	37 (45.7)
2	1 (1.4)	0 (0.0)
Prior lines of therapy group		
>=3	68 (98.6)	78 (96.3)
<3	1 (1.4)	3 (3.7)
Prior lines of therapy		
Subjects with data	69	81
Mean	4.5	4.0
SD	1.7	1.7
Median	4.0	4.0
Range	2.0 to 12.0	1 to 11.0
Refractory or relapsed after 3 or more lines		
Yes	69 (100.0)	81 (100.0)
Time of relapse since SCT failure group		
>=12 months	37 (53.6)	0
< 12 months	32 (46.4)	0
Missing	0	81 (100.0)
Time of relapse since SCT failure (Months)		
Subjects with data	69	0

	Cohort 1 n (%)	Cohort 2 n (%)
Mean	60.2	NA
SD	39.6	NA
Median	12.6	NA
Range	2.5 to 247.9	NA
Brentuximab Vedotin use		
Yes	69 (100)	81 (100)
Prior Radiation		
Yes	31 (44.9)	21 (25.9)
No	38 (55.1)	60 (74.1)
Bulky lymphadenopathy		
Yes	5 (7.2)	12 (14.8)
No	64 (92.8)	69 (85.2)
Baseline B symptoms		
Yes	22 (31.9)	26 (32.1)
No	47 (68.1)	55 (67.9)
Baseline Bone marrow involvement		
Yes	3 (4.3)	5 (6.2)
No	66 (95.7)	75 (92.6)
Missing	0	1 (1.2)

4.6 Quality assessment of the relevant trials

To assess the risk of bias and quality of non-randomised trials and observational studies the Newcastle-Ottawa Scale was used [56]. This tool evaluates: selection bias in the choice of study population, exposure and outcome(s) as well as bias in the assessment of outcome(s), and length and quality of follow-up. Studies are awarded a star in each domain if they are deemed to have little to no risk of bias; this was conducted by two independent reviewers, with any disagreements resolved by a third reviewer. Further information can be found in Appendix 4, relating to the domains considered and supportive evidence required. As only two studies were considered relevant to the decision problem, both are presented in Table 12.

Table 12. Quality assessment of relevant clinical trials

Trial ID	Selection				Comparability of cohorts	Outcome		
	Representativeness of cohort	Selection of non-exposed cohort	Ascertainment of exposure	Outcome of interest		Outcome assessment	Adequate duration of follow-up	Adequate follow-up of cohort
Cheah et al, 2016	*	N/A	*	*	N/A		*	*
KEYNOTE-087	*	N/A	*	*	N/A	*		*

4.7 Clinical effectiveness results of the relevant trials

4.7.1 KEYNOTE-087

The following information is reported from the KEYNOTE-087 efficacy update report based on a database cut-off date 21st March 2017. The March 2017 data cut corresponds to one year after the last subject was initiated on study treatment. The median follow up time as of 21st March 2017 was 15.9 months (range 1.0 to 20.9 months) with 31 and 21 patients remaining on treatment in cohorts 1 and 2, respectively [46].

The primary objective of best ORR, based on BICR using IWG criteria in the ASaT population, was 75.4% and 66.7% in Cohort 1 and 2, respectively[46] (Table 14).

As per the secondary objectives: The complete remission rate was 27.5% for Cohort 1 and 24.7% for cohort 2 (Table 14). Among all responders the median time to response by BICR was [redacted] for Cohort 1 and [redacted] for Cohort 2. The median duration of response was [redacted] in Cohort 1 [redacted] and [redacted] in Cohort 2 (Table 15). At the time of data cut-off [redacted] and [redacted] of patients with a response in Cohorts 1 and 2 had an ongoing response, respectively (Table 16) [46].

Median PFS in the ASaT population assessed by BICR was 16.7 months (95% CI 11.2, not reached) and 11.1 months (95%CI 7.6-13.7) in Cohorts 1 and 2, respectively. The PFS rate(s) at month 3, 6, 9 and 12 are reported in Table 17 for Cohorts 1 and 2 [46].

Median OS at the time of data cut-off was [redacted] across the total study population or within individual cohort(s). The OS rate at 6 and 12 months was [redacted] and [redacted] in Cohort 1, and [redacted]% and [redacted]% in Cohort 2, respectively (Table 18) [46].

Data for exploratory endpoints were not included within the March 2017 updated efficacy report. However, of particular relevance to this submission are the patient reported outcome (PRO) data relevant to the economic model in Section 5.0. KEYNOTE-087 reported a patients change in health related quality of life (HRQoL) using the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life (QoL) Questionnaire C30 (QLQ-C30) and European Quality of Life Five Dimensions Questionnaire (EuroQoL EQ-5D). Using the updated efficacy report from September 2016 these data were reported using the PRO-specific ASaT for all Cohorts combined (1, 2 and 3); this was all patients who received at least one dose of study medication and completed at least one PRO instrument [47].

The treatment effect on PRO score change from baseline was evaluated at Week 12 using constrained longitudinal data analysis. Week 12 was selected to minimise the loss of data due to death or disease progression while allowing comparisons in scores while patients were still on treatment. Compliance rates for both the EORTC QLQ-C30 and EQ-5D were over 91% at baseline, and over 97% at Week 12 [47]. Completion rates remained at or above 90% at each time point after baseline, until Week 24, when they dropped as patients discontinued the study due to disease progression, physician decision, AEs, or death [47].

Results from the PRO analyses indicated a net improvement in the EORTC QLQ-C30 global health status/QoL score from baseline to Week 12, across all response groups. There was an overall improvement of 8.5 points (SE; 1.6) compared to baseline [47]. The improvement was greatest among those with CR/PR (+10.4 points), followed by those with SD (+7.3 points), then PD (+3.5 points) [47]. Consistency of findings were seen in the EQ-5D measures, with a change in VAS score from baseline to Week 12 that may be considered clinically important in those who responded (10.9+ points for CR/PR), as compared to those who did not respond (5.4+ points for SD patients and 2.6+ points for PD patients) [47]. Together, these results suggest that health-related QoL were improved in this RRcHL population.

Table 13. Summary of follow up duration ASaT population Keynote-087 [46]

	Keynote-87 (Cohort 1-3)
Follow-up duration (months)†	
Median (Range)	15.9 (1.0-20.9)
Mean (SD)	15.9 (2.5)
† Follow-up duration is defined as the time from first dose to the date of death or the database cut-off date if the subject was still alive.	

Table 14. Summary of best overall response based on BICR [46]

Response evaluation	Cohort 1 (N=69), ASAT population				Cohort 2 (N=81), ASAT population			
	n	%	95% CI*	P-value†	n	%	95% CI*	P-value†
Complete Remission (CR)	19	27.5	(17.5-39.6)		20	24.7	(15.8, 35.5)	
Partial Remission (PR)	33	47.8	(35.6, 60.2)		34	42.0	(31.1, 53.5)	
Objective Response (CR+PR)	52	75.4	(63.5, 84.9)	P<0.001	54	66.7	(55.3, 76.8)	P<0.001
Stable Disease (SD)	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
Progressive disease (PD)	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	
No Assessment	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	

*Based on binomial exact confidence interval method

†One-sided p-value based on exact binomial distribution for testing H0 p≤0.20 versus H1: p>0.20

All subjects as treated population

Table 15. Summary of time to response and response duration based on BICR as per IWG in subjects with a response from Cohort 1 and Cohort 2 [46]

	Cohort 1	Cohort 2
Number of subjects with a response*	[redacted]	[redacted]
Time to response (months)*		
Mean (SD)	[redacted]	[redacted]
Median (Range)	[redacted]	[redacted]
Response Duration (months) †		
Median (Range)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Number of subjects with a response ≥3 months (%)†	[redacted]	[redacted]
Number of subjects with a response ≥6 months (%)†	[redacted]	[redacted]
Number of subjects with a response ≥9 months (%)†	[redacted]	[redacted]
Number of subjects with a response ≥12 months (%)†	[redacted]	[redacted]

*Analyses on time to response and response duration are based on subjects with best overall response as complete remission or partial remission only

†From product-limit (Kaplan-Meier) method for censored data

+ indicates there is no progressive disease by the time of last disease assessment

All subjects as treated population

Table 16. Summary of response outcomes based in BICR per IWG in subjects with response [46]

	Cohort 1 N=69	Cohort 2 N=81
Number of Subjects with Response†	[redacted]	[redacted]
Censored Subjects (%)	[redacted]	[redacted]
Subjects who progressed or died after 2 or more missed visits (%)	[redacted]	[redacted]
Subjects started new anti-cancer treatment (%)	[redacted]	[redacted]
Subjects with stem cell transplant (%)	[redacted]	[redacted]
Subjects who were lost to follow-up (%)	[redacted]	[redacted]
Subjects who had no disease assessments in 30 weeks (%)	[redacted]	[redacted]
Ongoing response‡ (%)	[redacted]	[redacted]
Range of DOR (months)	[redacted]	[redacted]
Ongoing response ≥ 3 months	[redacted]	[redacted]
Ongoing response ≥ 6 months	[redacted]	[redacted]
Ongoing response ≥ 9 months	[redacted]	[redacted]
Ongoing response ≥ 12 months	[redacted]	[redacted]

†Response: Analyses are based on subjects with a best overall response as complete remission or partial remission.

‡ Ongoing response: Subjects who are censored, alive, have not progressed, have not started a new anti-cancer therapy, are not lost to follow-up and the last non-"NE" imaging assessment is within 210 days (30 weeks) of the data cut-off date.

All subjects as treated population

Table 17. Summary of Progression free Survival (PFS) based on BICR as per IWG from Cohort 1 and Cohort 2 [46]

	Cohort 1 N=69	Cohort 2 N=81
Number (%) of PFS events	[redacted]	[redacted]
Person-months	[redacted]	[redacted]
Event rate/ 100 Person-months	[redacted]	[redacted]
Median PFS (months) *	16.7	11.1
95% CI for Median PFS*	(11.2, not reached)	(7.6, 13.7)
PFS rate at 3 Months in %*	[redacted]	[redacted]
PFS rate at 6 Months in %*	[redacted]	[redacted]
PFS rate at 9 Months in %*	[redacted]	[redacted]
PFS rate at 12 Months in %*	[redacted]	[redacted]
Progression free survival is defined as time from first dose to disease progression, or death, whichever occurs first. *From product limit (Kaplan Meier) method for censored data All subjects as treated population		

Table 18. Summary of Overall Survival cohorts 1 and 2 [46]

	Cohort 1 N=69	Cohort 2 N=81
Death	[redacted]	[redacted]
Median Survival (Months)†	[redacted]	[redacted]
95% CI for Median Survival†	[redacted]	[redacted]
OS rate at 6 Months in % †	[redacted]	[redacted]
OS rate at 9 Months in % †	[redacted]	[redacted]
OS rate at 12 Months in % †	[redacted]	[redacted]
OS rate at 15 Months in % †	[redacted]	[redacted]
OS: Overall survival. † From product-limit (Kaplan-Meier) method for censored data. All subjects as treated population		

4.8 Subgroup analysis/ Post-Hoc analysis

To address the decision problem as described in Section 1.1, the following post-hoc analyses were conducted. The results of these analyses have been used to inform the naïve indirect treatment comparison and the Matched Adjusted Indirect Comparison (MAIC) as reported in Section 4.10.

To support the economic model outlined in Section 5.0, the following information was derived from the Keynote-087 using the March 2017 data cut.

These post-hoc analyses reported the following [45]:

- Overall Response at Week 12
- Progression-Free Survival (PFS) from Week 12 onwards and over the total treatment period by Overall Response at Week 12
- Time to discontinuation of study drug from Week 12 onwards by Overall Response at Week 12
- Overall Survival

4.8.1 Statistical analysis

The ASaT population, as per the primary analysis population of the KEYNOTE-087 protocol, was used for the post-hoc analyses presented below; i.e., patients are included if they received at least one dose of study medication.

Overall Response

For the purposes of this submission, and to support the economic case reported in Section 5.0; the number and percentage of patients are summarised for overall response at Week 12. Tabulations are provided for Cohort 1, Cohort 2 separately and combined. In addition, tabulations by overall response are provided for patients on study drug at Week 12. A patient was considered on study drug at Week 12 if they were on study drug at the time of the assessment of the overall response.

The overall response assessment is based on BICR using IWG criteria (Cheson 2007) [50] and is collected during the treatment period. Patient assessment could report one of the following responses: CR, PR, SD, PD, and not evaluable (NE). Patients with no assessment available for overall response at Week 12 are classified as no assessment (NA).

Progression-Free Survival

Progression free survival from Week 12 onwards is reported by overall response status at Week 12. These analyses take as a start point the assessment date of overall response at Week 12. In addition, PFS analyses are provided over the total treatment period. These analyses take as start point the day of first study drug initiation.

The PFS is computed from a patients start point towards the first documented progression of disease according to IWG criteria or death due to any cause, whichever occurs first, expressed in days. Patients without an event (progression or death) at the time of last tumour assessment are considered right censored at the last disease assessment date. Responses are based on BICR assessment using IWG criteria (Cheson 2007) [50].

Progression-Free Survival curves were estimated using the Kaplan-Meier method and results are presented graphically. The following statistics are presented below: number of patients included in the evaluation, raw percentage of patients with the event of interest, median time to event and its 95 % confidence interval (if median is reached), survival rates at months 3, 6, 9, 12 and 15 from the Kaplan-Meier survival estimate. Rates at month 18 are provided in addition for the PFS analyses over the total treatment period. A result of “not reached” is displayed for the survival rate at a specific month in case all patients had an event or were censored prior to the specific time point (month). The unit of time used in tabulations and figures is months. Analyses were performed from Week 12 onwards and over the total treatment period by overall response score at Week 12 respectively for Cohort 1 and Cohort 2 (separately and combined). Patients without an assessment of overall response at Week 12 were excluded from the PFS analyses by overall response.

Note that disease progression is assessed periodically, and that PD can occur any time in the interval between the last assessment where PD was not documented and the assessment when PD is documented. For the patients who have PD the true date of disease progression is approximated by the date of the first assessment at which PD is objectively documented per IWG criteria by BICR, regardless of discontinuation of study drug. Death is always considered as a confirmed PD event.

The censoring rules for the PFS are identical to those applied in CSR of study KEYNOTE-087.

Overall Survival

Overall survival (OS) was defined as time from first dose intake to death due to any cause, expressed in days. Patients without documented death are considered right censored at the day of last contact. Patients who had a survival update after the data cut-off date of March 2017 are censored at the cut-off date.

Overall survival curves were estimated using the Kaplan-Meier method and results presented graphically. The following statistics are presented: number of patients included in the evaluation, raw percentage of patients with the event of interest, median time to event and its 95% confidence interval (if median is reached), survival rates at months 3, 6, 9, 12, 15 and 18 from the Kaplan-Meier survival estimate. The unit of time used in tabulations and figures is months.

4.8.2 Post-hoc analysis results[45]

At Week 12 the number of patients on study drug in Cohort 1 and 2 was [redacted] and [redacted] respectively. The ORR reported at Week 12, relevant to CUA model described in section 5.0, was [redacted] in Cohort 1 and [redacted] in Cohort 2. Rates of CR and OR are reported in Table 19 and were broadly comparable between Cohorts 1 and 2.

Progression free survival is reported according to overall response status at Week 12 and can be found in Table 21 and Table 22 for Cohort 1 and 2, respectively. The median PFS, as reported among patients with overall response status at Week 12, was [redacted] in Cohort 1 (n=63); note that the median PFS had not been reached for patients stratified according to: CR, PR, or ORR. The PFS rate for patients with an ORR (CR+PR) was [redacted] at three months, and [redacted] at 12 months. Table 21 reports the PFS rate according to those patients with a response at Week 12 onwards for months 3, 6, 9, 12, and 15.

The median PFS for Cohort 2 (n=74), as reported among patients with an overall response status at Week 12 onwards, was [redacted] The median PFS for those patients who had achieved an ORR at Week 12 was [redacted]. The median PFS for patients stratified according to PR and SD at Week 12 was [redacted] respectively. The median PFS patients stratified according to CR was [redacted] (Table 22). The PFS rate for ORR (CR+PR) was [redacted] at three months, and [redacted] at month 12. Table 22 reports the PFS rate according to those patients with a response at Week 12 onwards for months 3, 6, 9, 12, and 15.

The analysis of PFS over the total treatment period is reported in Table 23 and Table 24 for Cohort 1 and 2, respectively. The median PFS in Cohort 1 was 16.7 months, (95% CI 11.2-not reached). The median PFS in Cohort 2 was 11.1 months (95% CI 7.6-13.7). The corresponding Kaplan-Meier curves are shown in Figure 8 and Figure 9.

Overall Response

Table 19. Summary of Overall Response at Week 12 Based on Central Review per IWG

	Cohort 1 (N=69)		Cohort 2 (N=81)	
	n (%)	95% CI[a]	n (%)	95% CI[a]
Complete Remission (CR)	[redacted]	[redacted]	[redacted]	[redacted]
Partial Remission (PR)	[redacted]	[redacted]	[redacted]	[redacted]
Objective Response (CR+PR)	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease (PD)	[redacted]	[redacted]	[redacted]	[redacted]
No Assessment (NA)	[redacted]	[redacted]	[redacted]	[redacted]

a: Based on binomial exact confidence interval method
All subjects as treated analysis population

Table 20. Patients on study drug at week 12 by Overall Response at Week 12 Based on Central Review per IWG

	Cohort 1 (N=69)		Cohort 2 (N=81)	
	n (%)	95% CI[a]	n (%)	95% CI[a]
Number of subjects on study drug at week 12	[redacted]		[redacted]	
Overall Response at 12 weeks				
Complete Remission (CR)	[redacted]	[redacted]	[redacted]	[redacted]
Partial Remission (PR)	[redacted]	[redacted]	[redacted]	[redacted]
Objective Response (CR+PR)	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease (PD)	[redacted]	[redacted]	[redacted]	[redacted]

a: Based on binomial exact confidence interval method
A subject is considered on study drug at week 12 if the subject was on study drug at the time of the assessment of overall response at week 12. Subjects without an assessment of overall response at week 12 are excluded.
All subjects as treated analysis population

Progression free survival

Table 21. Analysis of Progression-Free Survival from Week 12 onwards by Overall Response Based on Central review per IWG at Week 12 Cohort 1

	Cohort 1							
Progression-Free Survival from Week 12 onwards	**N[a]**	**Patients with Event n (%)**	**Median Time[b] in Months [95 %-CI]**	**Rate at Month 3[b] in % [95% CI]**	**Rate at Month 6[b] in % [95% CI]**	**Rate at Month 9[b] in % [95% CI]**	**Rate at Month 12[b] in % [95% CI]**	**Rate at Month 15[b] in % [95% CI]**
Overall	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
By Overall Response at Week 12								
Complete Remission (CR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial Remission (PR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
By Overall Response(CR+PR) at Week 12								
Objective Response (CR+PR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated with assessment of overall response at week 12. Subjects with a Progression-Free Survival event prior to week 12 are excluded from the analyses.

b: From product-limit (Kaplan-Meier) method. The provided rates at a specific month refer to a period starting 12 weeks after the first study drug intake

Progression-free survival is defined as time from overall response assessment at week 12 to disease progression, or death, whichever occurs first.

CI: Confidence Interval.

All subjects as treated analysis population

Figure 6. Kaplan-Meier of Progression-Free Survival from Week 12 onwards by Overall Response Based on Central review per IWG at Week 12 Cohort 1 (ASAT population)

Table 22. Analysis of Progression-Free Survival from Week 12 onwards by Overall Response Based on Central review per IWG at Week 12 Cohort 2

	Cohort 2							
Progression-Free Survival from Week 12 onwards	**N[a]**	**Patients with Event n (%)**	**Median Time[b] in Months [95 %-CI]**	**Rate at Month 3[b] in % [95% CI]**	**Rate at Month 6[b] in % [95% CI]**	**Rate at Month 9[b] in % [95% CI]**	**Rate at Month 12[b] in % [95% CI]**	**Rate at Month 15[b] in % [95% CI]**
Overall	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
By Overall Response at Week 12								
Complete Remission (CR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial Remission (PR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
By Overall Response(CR+PR) at Week 12								
Objective Response (CR+PR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated with assessment of overall response at week 12. Subjects with a Progression-Free Survival event prior to week 12 are excluded from the analyses

b: From product-limit (Kaplan-Meier) method. The provided rates at a specific month refer to a period starting 12 weeks after the first study drug intake

Progression-free survival is defined as time from overall response assessment at week 12 to disease progression, or death, whichever occurs first.

CI: Confidence Interval.

All subjects as treated analysis population

Figure 7. Kaplan-Meier of Progression-Free Survival from Week 12 onwards by Overall Response Based on Central review per IWG at Week 12 Cohort 2 (ASAT population)

Table 23. Analysis of Progression-Free Survival Based on Central review per IWG over total treatment period - Cohort 1

	Cohort 1								
	N[a]	Patients with Event n (%)	Median Time[c] in Months [95 %-CI]	Rate at Month 3[b] in % [95% CI]	Rate at Month 6[b] in % [95% CI]	Rate at Month 9[b] in % [95% CI]	Rate at Month 12[b] in % [95% CI]	Rate at Month 15[b] in % [95% CI]	Rate at Month 18[b] in % [95% CI]
Progression-Free Survival (IRC Primary Analysis)	69	[redacted]	16.7 [11.2;-]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated, cohort 1
b: From product-limit (Kaplan-Meier) method
CI: Confidence Interval.
All subjects as treated analysis population

Table 24. Analysis of Progression-Free Survival Based on Central review per IWG over total treatment period - Cohort 2

	Cohort 2								
	N[a]	Patients with Event n (%)	Median Time[c] in Months [95 %-CI]	Rate at Month 3[b] in % [95% CI]	Rate at Month 6[b] in % [95% CI]	Rate at Month 9[b] in % [95% CI]	Rate at Month 12[b] in % [95% CI]	Rate at Month 15[b] in % [95% CI]	Rate at Month 18[b] in % [95% CI]
Progression-Free Survival (IRC Primary Analysis)	81	[redacted]	11.1 [7.6;13.7]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated, cohort 2
b: From product-limit (Kaplan-Meier) method
CI: Confidence Interval.
All subject as treated analysis population

Figure 8. Kaplan-Meier of Progression-Free Survival Based on Central review per IWG Cohort 1, total treatment period (ASAT population)

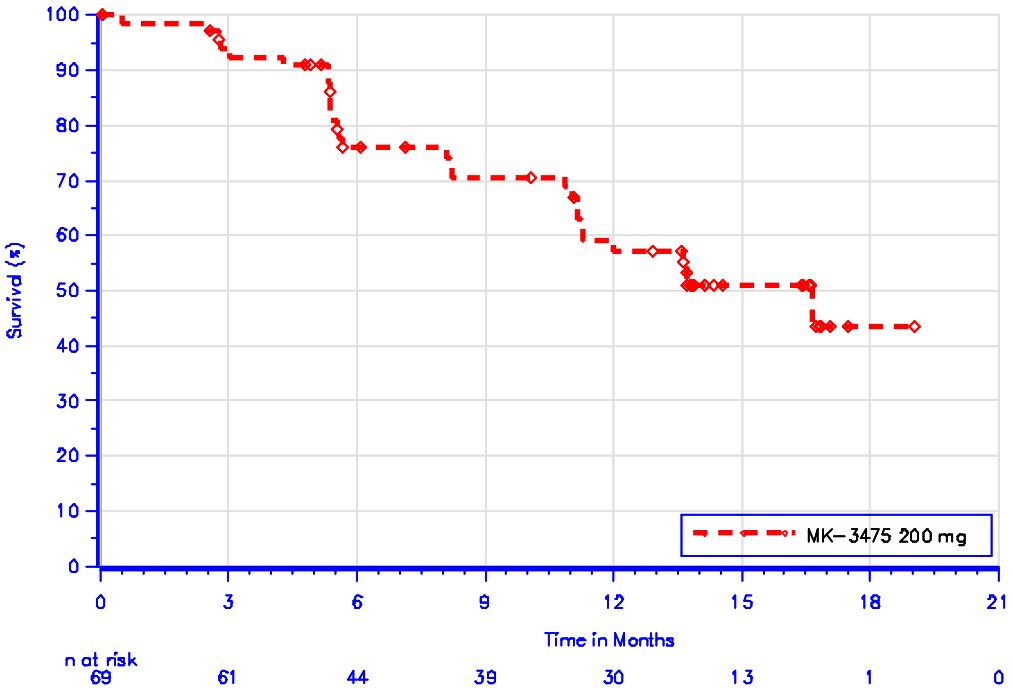
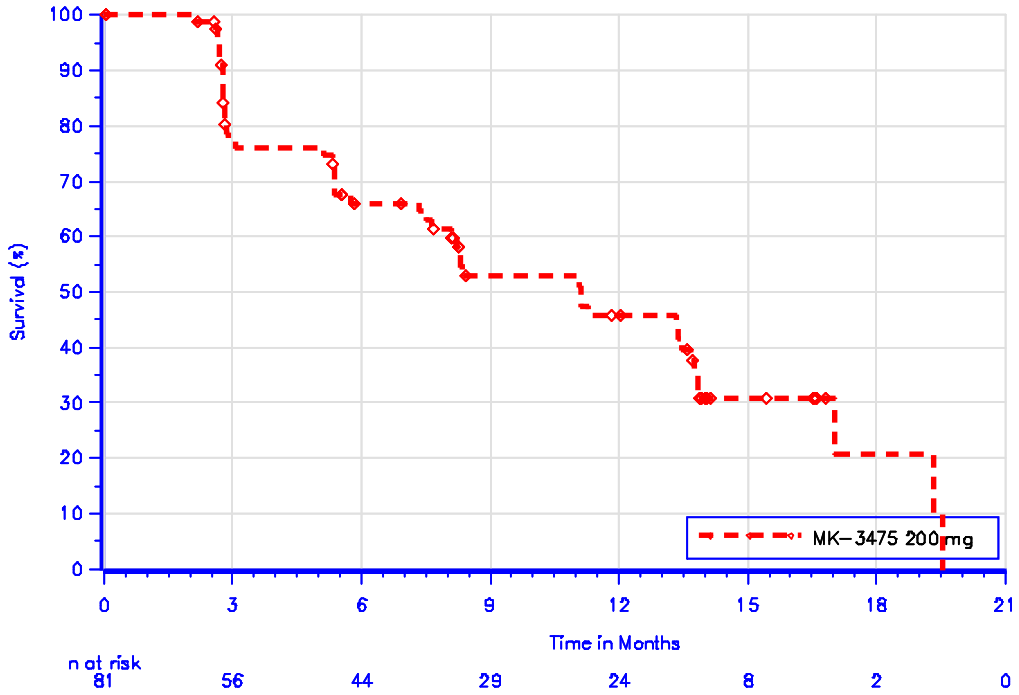


Figure 9. Kaplan-Meier of Progression-Free Survival Based on Central review per IWG Cohort 2, total treatment period (ASAT population)



Overall Survival

Table 25. Analysis of Overall Survival Based on total treatment period (ASAT population) - Cohort 1

	Cohort 1								
	N[a]	Patients with Event n (%)	Median Time[c] in Months [95 %-CI]	Rate at Month 3[b] in % [95% CI]	Rate at Month 6[b] in % [95% CI]	Rate at Month 9[b] in % [95% CI]	Rate at Month 12[b] in % [95% CI]	Rate at Month 15[b] in % [95% CI]	Rate at Month 18[b] in % [95% CI]
Overall Survival	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated, cohort 1
b: From product-limit (Kaplan-Meier) method
CI: Confidence Interval.
All subjects as treated analysis population

Table 26. Analysis of Overall Survival Based on total treatment period (ASAT population) - Cohort 2

	Cohort 2								
	N[a]	Patients with Event n (%)	Median Time[c] in Months [95 %-CI]	Rate at Month 3[b] in % [95% CI]	Rate at Month 6[b] in % [95% CI]	Rate at Month 9[b] in % [95% CI]	Rate at Month 12[b] in % [95% CI]	Rate at Month 15[b] in % [95% CI]	Rate at Month 18[b] in % [95% CI]
Overall Survival	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

a: Number of patients: all subjects as treated, cohort 2
b: From product-limit (Kaplan-Meier) method
CI: Confidence Interval.
All subjects as treated analysis population

Figure 10. Kaplan-Meier of Overall Survival based on total treatment period - Cohort 1 (ASAT population)

Figure 11. Kaplan-Meier of Overall Survival based on total treatment period - Cohort 2 (ASAT population)

4.9 Meta-analysis

The results of a naïve indirect treatment comparison and MAIC are reported in section 4.10.

4.10 Indirect and mixed treatment comparisons

4.10.1 Search strategy

As previously described (Section 4.1), a SLR was conducted to identify studies that could inform the comparative effectiveness of pembrolizumab for the treatment of patients that were considered; 1) to have failed to achieve a response or progressed after autoSCT and have relapsed after treatment with or failed to respond to BV post autoSCT; 2) unable to achieve a complete response (CR) or partial response (PR) to salvage chemotherapy and did not receive autoSCT, but have relapsed after treatment with or failed to respond to BV.

Details relating to the search strategy are reported in Section 4.1.2, and full search strategies can be found in Appendix 2.

As discussed in Section 4.2.1 a single retrospective observational study (Cheah et al 2016) [44] was identified and considered relevant to the decision problem. The study selection process is highlighted in Figure 3 and Figure 4.

4.10.2 Details of treatments

The decision problem is presented in Section 1.1. The treatment(s) considered within the comparative analysis relates to a pooled SoC and is considered representative of UK clinical practice.

4.10.3 Criteria used in trial selection

The inclusion criteria and the study selection process are described in section 4.1.3 (see PICOS eligibility criteria as per Table 7).

4.10.4 Summary of trials included

As previously described there was a paucity of data for the populations considered. As per section 4.1.2 a single retrospective observational study was identified (Cheah et al 2016).

Trial ID	Agent	Route	Dose	Day(s) given	Cycle length (days)	Maximum number of cycles/ duration of treatment
Cheah et al, 2016	Mixed agents	--	--	--	--	--
KEYNOTE-087	Pembrolizumab	IV	200 mg	1	21	24 months

Cheah et al. 2016 [44]

This was a retrospective observational study conducted at the MD Anderson Centre in the USA between June 2007 and January 2015. The study was designed to identify patients with cHL treated with BV who were either refractory to treatment or experienced disease relapse.

Patients with confirmed cHL were considered eligible for inclusion if they had; i) received treatment with BV for relapsed HL, or ii) subsequent disease progression at any time after treatment with BV, and had gone on to receive treatment including: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or alloSCT, or other treatment.

The included population (n=97) was predominantly male (53%) with a median age of 28 years. Patients reported an ECOG score of 0 (84%) or 1 (16%), and had previously received a median of three (range 0-9) prior lines of therapy. All patients received BV, the main reason for discontinuation was disease progression (n=76, 78%); and ten (10%) actively discontinued in order to receive a stem cell transplant. Patient baseline characteristics can be found in Appendix 6.

The results show patients treated with “investigator choice” treatment options achieved a post-BV progression ORR of 45% (CR 15%). After a median observation period of 25 months (range 1-76) from the point of post-BV progression; 65 patients treated with “investigators choice” had either progressed or died with a median PFS of 3.5 months (Appendix 7).

The authors commented that due to the heterogeneity among patients and the treatment options received post-BV progression, it was not possible to identify prognostic factors associated with PFS. The authors concluded that this study demonstrates “the persistent challenge of achieving durable disease control in patients with relapsed/refractory cHL following failure of BV therapy”[44].

KEYNOTE-087

Full details relating to KEYNOTE-087 can be found in Section 4.3 to 4.8.

4.10.5 Trials identified in the search strategy

Not applicable.

4.10.6 Rationale for choice of outcome measure chosen

As described in the decision problem Section 1.1 the outcomes of interest relate to response rate ORR (CR+PR) and survival (PFS).

Both OS and PFS are clinically relevant outcomes that were referenced in the final scope for this appraisal and the decision problem. However, due to a lack of events during the follow-up period it has not been possible to consider OS within the long term model structure in those who do not receive an alloSCT. Therefore, PFS has been extrapolated in those who do not receive an alloSCT within the economic model to demonstrate the benefit of pembrolizumab within the populations considered. In addition the response status has also been reported, and is considered relevant within clinical practice as it denotes the proportion of patients who are able to undergo alloSCT. Long term survival following this is derived from literature

No formal method of data analysis was proposed for AEs or HRQoL, as these data are not available within the comparator study Cheah et al 2016. Within the economic analysis, rates of grade 3+ AE greater than 0% incidence from KEYNOTE-087 were derived for SoC from literature sources and applied (further detail in section 5.3.5). HRQoL for SoC were assumed the same as the pembrolizumab arm adjusted for the proportions of patient from the SoC arm in each response status (further detail in section 5.4.7).

4.10.7 Populations in the included trials

Full details of the KEYNOTE-087 trial population are reported in Section 4.5. The baseline characteristics of Cheah et al. 2016 are reported in Appendix 6.

Note information from Cheah et al., 2016 relates to the time of documented progression following therapy with BV unless otherwise stated. The average age of patients across the two studies was similar, with medians of 32 (range 18-84) in Cheah et al., 2016 and 34 (range 19 to 76) in KEYNOTE-87. The median ages in each cohort separately were also comparable, though the median in Cohort 2 (40 years) was higher than Cohort 1 (34 years). The proportion of patients over 45 was lower in Cheah et al., 2016 (14%) than KEYNOTE-087 (34%), with a higher proportion in Cohort 2 (42%) than Cohort 1 (25%). The distribution of males and

females was also comparable across the two studies, but information on gender was only available from Cheah et al., 2016 before commencement of BV therapy.

The proportions of patients with ECOG performance scores of ≥1 were also comparable between the two studies (59% in Cheah et al., 2016 and 52% in KEYNOTE-087), though the proportion in Cohort 2 (46%) was lower than Cohort 1 (58%). A much higher proportion of patients in KEYNOTE-087 (32% in both Cohorts 1 and 2) than Cheah et al., 2016 (8%). KEYNOTE-087 had a higher proportion with albumin <40 g/l (49% versus 28%) and white blood cell count >15 x 109 /l (13% versus 5%), but a lower proportion with lymphocytes <0.6 x 109 /l (17% versus 41%). The proportion of patients with haemoglobin <105 g/l was comparable across the two studies (31% in KEYNOTE and 35% in Cheah et al., 2016). KEYNOTE-087 had a higher proportion of patients with a maximum tumour diameter >4 cm (45% versus 26%) and a higher proportion of patients with extranodal disease (48% versus 35%). For the latter, a higher proportion was observed in Cohort 1 (57%) than Cohort 2 (41%)

Finally, patients in Cheah et al., 2016 had received a median of 3 (range 0-9) prior lines of therapy before commencing treatment with BV compared to a median of 4 (range 1-12), including treatment with BV, in Cohorts 1 and 2 from KEYNOTE-87. As with gender, this information was recorded before commencement of BV therapy.

4.10.8 Apparent or potential differences in the patient population between the trials

Data relevant to the decision problem is limited. However, NICE confirmed in the FAD TA462 [57], that the Cheah et al. 2016[44] population can be considered generalisable to the UK and is adequate for decision making.

4.10.9-4.10.11 Methods, outcomes, baseline characteristics, risk of bias of each trial

- Study methods for KEYNOTE-087 and Cheah et al 2017 have been described above.
- Appendix 6 for study baseline patient characteristics.
- Appendix 7 for a summary of study results.
- Please see Table 12 for study risk of bias using the NOS; Appendix 4.

4.10.12 Methods of analysis and presentation of results

A feasibility assessment was conducted that focused on two areas: the compatibility of included studies and the data published on potential confounders i.e. the extent to which adjustment could be made to ensure exchangeability [58, 59]. Compatibility was assessed by comparing study design characteristics such as inclusion and exclusion criteria, study endpoints and methods for outcomes assessments.

For each treatment arm of each published study included in the analysis, the reported Kaplan-Meier (KM) curves were digitized (DigitizeIt; http://www.digitizeit.de/) and the number of patients at risk over time was extracted. The algorithm proposed by Guyot et al., 2012 [60] was applied to simulate IPD (i.e. survival and censoring times) for each treatment arm

Naïve indirect comparison[61]

Progression free survival

A Cox proportional hazards regression model was identified for pembrolizumab (index intervention) survival based on the IPD from KEYNOTE-087 and for chemotherapy agents based on the IPD generated from the published KM curves as reported in Cheah et al 2016:

$$\ln(h_{it}) = \beta_{0t} + \sum_{c=1}^{1} \beta_c^x (x_{ci}) \qquad \text{(Equation 1)}$$

where h_{it} reflected the underlying hazard rate at time point t for subject i, β_{0t} was the baseline log-hazard at time t, x_{ci} was the covariate value for covariate c for subject i related to treatment (i.e. pembrolizumab versus chemotherapy), and β_c^x reflected the impact of treatment covariate c on the log hazard.

This model was fitted to obtain a naïve unadjusted log hazard ratio (HR) for pembrolizumab versus chemotherapy for two scenarios:

1. From study initiation to most recent observation;
2. From study initiation to week 12.

Objective response rates

Odds ratios (ORs) were estimated for response rates observed in KEYNOTE-087 versus rates observed in Cheah et al. 2016 using contingency tables and a chi-squared test for difference. Comparisons were conducted using data from three separate time periods from KEYNOTE-087:

1. Response at 12 weeks;

Where comparator studies only presented proportions based on best response, these were used to compare to the 12 Week estimates from KEYNOTE-087. Comparisons were further stratified according to level of response (objective, complete, and partial), see Appendix 10 and Appendix 11.

Matched adjusted indirect treatment comparison (MAIC)[61]

Progression free survival

Weights were estimated for participants from KEYNOTE-087 so that their weighted mean baseline characteristics matched those observed in each of the comparator studies featured in the pairwise comparisons. Weights for the KEYNOTE-087 were derived using the inverse odds of being in their current group (i.e. pembrolizumab) versus the aggregate data group (i.e. chemotherapy) as per the NICE Decision Support Unit (DSU) guidance[62].

In the first instance, matching was conducted using all variables for which data were available in both KEYNOTE-087 and Cheah et al. 2016. In cases where the algorithm used to estimate the weights did not converge using the full set of baseline characteristics, variables were removed in stepwise fashion in a predetermined order until convergence was achieved.

Weights from the propensity model were then applied to a Cox regression model with the same structure as equation (Equation 1 above) to obtain population-adjusted HRs for the same two scenarios:

1. From study initiation to most recent observation;
2. From study initiation to week 12.

Objective response rates

As described above, the same approach was used to estimate weights for each separate comparison. Weighted contingency tables and chi-squared test for difference were then used to estimate odds ratios; see Appendix 10 and Appendix 11 for results.

4.10.13 Programming language[61]

Please see Appendix 9

4.10.14-4.10.16 Results of analysis and results of statistical assessment of heterogeneity

Data on outcomes from pembrolizumab came from a single-arm study broken down into three separate cohorts according to their prior treatment with autoSCT and BV, while data on SoC was taken from an observational study conducted in the USA (Cheah et al. 2016). Single-arm trials and observational studies present a challenge for traditional approaches used to indirectly compare treatments not studies in head-to-head trials. In standard NMA, the sole source of error in estimates of relative treatment effects is statistical sampling error, assuming no differences in the distribution of effect modifiers across studies in included in the network. In naïve unadjusted comparisons conducted using data from single-arm clinical trials, this sampling error will also be present along with the systematic error, or bias, that comes from differences in the distribution of both prognostic factors and effect modifiers between study arms. The goal of methods such as MAIC is to reduce the size of this systematic error in relative treatment effects.

Estimates of relative treatment effects using either methodology suggest that pembrolizumab offers significant improvements in PFS and ORRs compared to agents which currently form SoC for RRcHL. Focussing on the naïve unadjusted comparisons, both PFS in the entire study period and the ORR based on best overall response in KEYNOTE-087 and Cheah et al., 2016 were most improved among patients in Cohort 1. After adjusting for differences in patient characteristics between those in KEYNOTE-087 and the Cheah et al., 2016 study, Cohort 1 still saw the biggest improvements in PFS relative to SOC, with the HR falling to [redacted] from [redacted] versus a drop from [redacted] to [redacted] for Cohort 2[61]. However, for ORR, improvements in Cohort 2 surpassed those for Cohort 1 as the OR increased for the former from [redacted] to [redacted] compared to an increase from [redacted] to [redacted] for the latter[61]. Pembrolizumab retained an advantage over SoC in the analyses comparing response at 12 weeks in KEYNOTE-087 and best overall response in Cheah et al., 2016, though the relative effect sizes tended to be smaller[61]. Smaller differences were also observed between the results from the naïve comparisons and MAICs when data from the two cohorts were combined as opposed to being analysed separately[61].

The MAICs conducted within this study are an example of an “unanchored” population-adjusted comparison, and as such are subject to a number of important limitations. Reliable prediction of absolute outcomes is required in order for unanchored comparisons to be valid. For this requirement to be met, all potential prognostic factors and effect modifiers need to be adjusted for within the propensity score weighting model. This is unlikely to be the case in

these comparisons as covariates included within the model were restricted to those reported in both KEYNOTE-087 and Cheah et al., 2016. Furthermore, in some circumstances covariates for which data were available had to be dropped in order for the model to converge[61]. The NICE DSU recommends when conducting unanchored comparisons, information should be provided on the level of bias that is likely to be introduced as a result of any covariates that are unaccounted for[62]. However, due to a lack of studies in the patient population relevant to this analysis this was not possible. As a result, we cannot comment on the degree of systematic error within the MAIC estimates. In addition, there is also uncertainty regarding the prognostic value of some of the variables included within the model. The Cheah et al., 2016 study did not identify any prognostic factors associated with PFS, though this is understandable given the small sample size in the study. This further complicates the estimation of the degree of error in both the naïve unadjusted comparisons and the MAIC. Another limitation of the MAIC approach is that relative treatment effects can only be estimated for the target population in the comparator trial i.e. Cheah et al., 2016. There is uncertainty over the degree of overlap between the characteristics of patients in the Cheah et al., 2016 and those seen in UK clinical practice. Despite this, Cheah et al., 2016 was still seen as the most suitable basis for comparison, highlighting the paucity of evidence in this area[61].

Comparison of Progression free survival[61]

The results of the naïve comparisons and MAICs of PFS with pembrolizumab versus SoC for the entire study scenario are presented in Table 27. The naïve comparisons resulted in HRs of [redacted] and [redacted] for cohort 1, cohort 2, and cohorts 1 and 2 combined, respectively[61]. All variables were included in the MAIC using data from cohorts 1 and 2 combined, while median prior lines dropped out from the comparisons using each cohort separately. In the comparison using cohort 1 the HR fell to [redacted], while in the other two it slightly increased to [redacted] for cohort 2 and [redacted] for the cohorts combined[61]. A similar pattern was observed in the 12-week scenario, although the HRs in the naïve comparisons were higher for cohort 1 and lower for cohort 2 (see Table 28).

Table 27: Summary of comparisons of progression-free survival for pembrolizumab versus SoC for the entire study scenario[61]

Cohort	Comparison	Sample size/effective sample size, n	Pembrolizumab		Hazard ratio (95% CI)
			Events, n	Censored, n	
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; MAIC, matching-adjusted indirect comparison.

Table 28: Summary of comparisons of progression-free survival for pembrolizumab versus SoC for the 12-week scenario[61]

Cohort	Comparison	Sample size/effective sample size, n	Pembrolizumab		Hazard ratio (95% CI)
			Events, n	Censored, n	
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; MAIC, matching-adjusted indirect comparison.

Comparison of Response rates[61]

Objective Response (ORR)

The results of the naïve comparisons and MAICs of objective response with pembrolizumab versus SoC are presented in Table 29, and Table 30. Using data from KEYNOTE-087 on best overall response, the naïve comparisons resulted in ORs of [redacted] and [redacted] for cohort 1, cohort 2, and cohorts 1 and 2 combined, respectively[61]. All variables were included in the MAIC using data from cohorts 1 and 2 combined, while median prior lines dropped out from the comparisons using each cohort separately. The ORs increased substantially in the MAICs of cohorts [redacted] and 2 [redacted] separately, with a small drop in the OR for the two cohorts combined [redacted] [61].

A similar pattern was observed when conducting the comparisons using data from KEYNOTE-087 on response at Week 12. For the 12-week scenario, the OR in the MAIC compared to the naïve comparison increased from [redacted] to [redacted] for cohort 1, but only four variables were included in the model (ECOG, B symptoms, age, albumin)[61]. Only prior lines of therapy was not included in the model for cohort 2 and the OR increased from [redacted] in the naïve comparison to [redacted][61]. All variables were included in the model for the two cohorts combined and this led to an increase in the OR from [redacted] in the naïve comparison to [redacted] [61].

Table 29: Summary of comparisons of objective response rates (best overall response) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

Table 30: Summary of comparisons of objective response rates (12-weeks) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

Complete Response (CR)[61]

Using data from KEYNOTE-087 on patients achieving a complete response (CR) as their best overall response, the naïve comparisons resulted in ORs of [redacted] and [redacted] for cohort 1, cohort 2, and cohorts 1 and 2 combined, respectively (Table 31)[61]. All variables were included in the MAIC using data from cohorts 1 and 2 combined, while median prior lines dropped out from the comparisons using each cohort separately. The ORs increased again in the MAICs of cohorts 1 [redacted] and 2 [redacted] separately, with a smaller increase in the OR for the two cohorts combined [redacted][61].

For the 12-week scenario, the OR in the MAIC compared to the naïve comparison increased from [redacted] to [redacted] for cohort 1 (Table 32), but only four variables were included in the model (ECOG, B symptoms, age, albumin[61]). Only prior lines of therapy was not included in the model for cohort 2 and the OR increased from [redacted] in the naïve comparison to [redacted][61]. All variables were included in the model for the two cohorts combined and this led to an increase in the OR from [redacted] to [redacted][61].

Table 31: Summary of comparisons of complete response (best overall response) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	CR with pembrolizumab	CR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; CR, complete response; SOC, standard of care

Table 32: Summary of comparisons of complete response (12-weeks) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	CR with pembrolizumab	CR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; CR, complete response; SOC, standard of care

Partial Response (PR)[61]

Using data from KEYNOTE-087 on patients achieving a PR as their best overall response, the naïve comparisons resulted in ORs of [redacted] and [redacted] for cohort 1, cohort 2, and cohorts 1 and 2 combined, respectively (see Table 33)[61]. All variables were included in the MAIC using data from cohorts 1 and 2 combined, while median prior lines dropped out from the comparisons using each cohort separately. The ORs decreased in the MAICs of cohorts 1 [redacted] and 2 [redacted] separately, and the two cohorts [redacted][61].

For the 12-week scenario, the OR in the MAIC compared to the naïve comparison increased from [redacted] to [redacted] for cohort 1 (Table 34), but only four variables were included in the model (ECOG, B symptoms, age, albumin)[61]. Only prior lines of therapy was not included in the model for cohort 2 and the OR decreased from [redacted] in the naïve comparison to [redacted][61]. All variables were included in the model for the two cohorts combined and this led to a decrease in the OR from [redacted] to [redacted].

Table 33: Summary of comparisons of partial response (best overall response) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	PR with pembrolizumab	PR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; PR, partial response; SOC, standard of care

Table 34: Summary of comparisons of partial response (12-weeks) for pembrolizumab versus SoC [61]

Cohort	Comparison	Sample size/effective sample size, n	PR with pembrolizumab	PR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; PR, partial response; SOC, standard of care

Full results can be found in Appendix 10 and Appendix 11[61].

4.10.17 Justification for the choice of random or fixed effect model

Not applicable, please see section 4.10.12

4.10.18 and 4.10.19 Heterogeneity between results of pairwise comparisons and inconsistencies between direct and indirect evidence

Not applicable

4.11 Non-randomised and non-controlled evidence

A comprehensive overview of the KEYNOTE-087 clinical trial has been reported in Section 4.7.

As described in Sections 4.1 and 4.10, there was a paucity of data relating to treatment(s) relevant to the decision problem. In an attempt to identify alternate data sources MSD engaged with the Hematological Malignancy Network (HMRN) based at York University, UK. The HMRN developed a proposal for MSD designed to address three objectives: 1) To describe and characterise treatment pathways (first and subsequent lines) for a population-based cohort of patients with cHL; 2) To examine PFS and OS by demographics, prognostic factors and by treatment; and 3) To examine time to response, response rates, OS and PFS by regimen and treatment line (first and subsequent lines) [63]. However, in the first instance the HMRN group examined available patient numbers; i.e. patients who had received prior therapy with BV.

During initial discussions with the HMRN, MSD was told that out of 700 cHL patients in the HMRN database, for whom data had been collected between 2004 and 2014, only 25 patients had received BV, and that only five patients had undergone a subsequent alloSCT. Therefore, the decision was made by MSD UK, supported by the HMRN that data were immature and non-feasible for the analyses required vs. the current literature sources (Cheah et al, 2016). This has been recently supported by NICE TA462, which explained that upon consideration of the evidence presented Cheah et al study was the best available evidence for standard of care and considered it appropriate for its decision-making [57]. Furthermore the committee noted that the study population partially matched the population of interest and that the company had explored UK SoC data from the HMRN and surveyed clinicians actively treating RRcHL in the UK [57]. The committee considered that both the HMRN network data and the clinician survey supported the Cheah et al 2016 publication as reflective of UK practice, but it recognised that the data were limited [57].

The manufacturer of nivolumab marked the aforementioned HMRN and Clinician survey data as Academic in Confidence (AIC), and therefore cannot be referenced at this time. However, as MSD were in attendance to this committee meeting, the results presented support the approach taken within this submission. In an attempt to validate the standard of care (SoC) evidence reported by Cheah et al. 2016[44], MSD commissioned a clinician survey to support understanding of UK clinical practice[64].

4.11.1 UK Clinician Survey

MSD commissioned a third party vendor to conduct a bespoke piece of market research to determine:

1) UK clinical practice for the treatment of patients with RRcHL

2) The treatment pathway and eligibility of patients with RRcHL following current standard of care

3) The validity of two pieces of literature; i) Cheah et al. 2016 in relation to the expectation of outcomes in UK clinical practice utilising SoC, and ii) Lafferty et al. 2016 in relation to the rates of alloSCT for patients who have received SoC in the relapsed/ refractory setting.

Clinician survey methodology

A questionnaire was developed in collaboration with MSD UK and medeConnect Ltd. This was designed to address the research questions described above. The questionnaire was accessible via the website Doctors.net.uk. This enabled practicing clinicians within the UK to take part in a ~20 minute survey and upon completion receive 7,000 eSR points (equivalent ~£30)[65].

Clinician survey results

A total of 16 clinicians completed the survey and form the evidence base for the results below. This includes practicing clinicians from England (n=12), Wales (n=1), and Scotland (n=3). All clinicians had experience working with Lymphoma malignancies and self-reported as either haematologists (44%) or haematological oncologists (56%). A minority reported experience with PD-1 therapy (4/16), or access to investigational agents (4/16).

To provide context for the outcome data reported within this survey, clinicians provided estimated patient numbers from their own clinical practice. Across both indications, relevant to this submission, patient numbers were low with an average annual estimate of 23 patients. When considering the two patient groups separately the average number of patients seen who; 1) had failed autoSCT and BV was four patients annually, and 2) who were ineligible for autoSCT and had failed BV was three patients annually.

Screening questions show that clinicians had experience using BV and transplanting patients with alloSCT, and thus represent a relevant stakeholder group for the validation of the

literature and for providing insight into UK clinical practice. As can be seen in Table 35 for patients treated with standard of care the mortality in Cohort 1 and 2 is 21% and 30%, respectively. The ORR (CR+PR) for Cohort 1 and Cohort 2 is 35% and 31%, respectively.

Clinicians reported that only a minority of patients would proceed to alloSCT within Cohort 2 (Table 36), which corresponds with the response estimates provided in Table 35. Within Cohort 1 57% and 44% of patients achieving CR or PR would be expected to proceed to an alloSCT; that is compared with 17% and 13% in Cohort 2, respectively. Prior to alloSCT clinicians reported that patients who failed autoSCT and subsequent BV therapy, on average received SoC for approx. 12.5 weeks (median 12 weeks, range 3-24 weeks); similarly those patients described as autoSCT ineligible and who had failed subsequent BV also received SoC for approx. 12.8 weeks (median 12 weeks, range 1-24 weeks). Clinicians provided verbatim free text comments including “SoC therapy is continued if effective until allogeneic donor is ready to minimise time off treatment” and “Generally guided by the allogeneic transplant centre. Delays may be due to donor availability or TBI slot availability”, which might suggest that SoC at this point in the treatment pathway serves as a bridge to alloSCT; this treatment approach was also supported by clinicians who attended an advisory board meeting held by MSD on 13th March 2017 [40].

All survey participants provided validation for the Cheah et al. 2016 and Lafferty et al. 2016 literature. Clinicians were in agreement with the median PFS of 3.5 months, and OS 25.2 months as reported by Cheah et al. 2016; however, three clinicians suggested an alternative of 12 months for median OS based on their practice. The OS (69%) and PFS (54%) at one year, as reported within the UK conference abstract by Lafferty et al. 2016, was supported by the majority of clinicians; however one clinician reported 50% OS at on year [64].

Discussion

The findings of this survey support the treatment algorithm described within this submission. Clinicians reported that the number of patients reaching this later line of treatment in the relapsed/ refractory disease course is small, and that experience is often restricted to specialist centres. Verbatim feedback included “Individual clinicians will have fairly small numbers of these patients so percentages are estimates” attesting to the rarity of this condition.

The poor outcome(s) of patients with RRcHL demonstrate the need for alternative treatment therapies. Furthermore, the outcomes of patients who do not receive/ are not considered eligible for alloSCT are markedly worse, and both patients groups may benefit from a treatment option, that in the first instance is able to provide an ORR, so as to be considered

suitable for alloSCT. Of particular interest is the validation of Cheah et al 2016; the findings of this non-UK study were accepted by the majority of respondents. However, it should be noted that many of the clinicians reported no access to investigational agents, which are included within the Cheah et al. 2016 publication; thus these results could be considered a "upside" estimate of the potential treatment outcome(s) expected within the UK.

Table 35. Average response observed in UK clinical practice for patients treated with SoC

Responses	On SoC having failed both auto SCT and subsequent BV therapy (Cohort 1 equivalent)	On SoC after being identified as ineligible for auto SCT and then failed BV therapy (Cohort 2 equivalent)
CR	14% (range 0-40%, median=10%)	12% (range 0-30%, median=12.5%)
PR	21% (range 0-40%, median=20%)	19% (range 0-40%, median=20%)
SD	20% (range 10-60%, median=20%)	14% (range 0-30%, median=15%)
PD	24% (range 10-40%, median=25%)	25% (range 0-50%, median=25%)
Death	21% (range 0-70%, median=18.5%)	30% (range 0-95%, median=20%)
TOTAL	**100%**	**100%**
CR, complete remission; PD, progressive disease PR, partial response; SD, stable disease		

Table 36. Proceed to alloSCT based on response as observed in UK clinical practice for patients treated with SoC

Responses	On SoC having failed both auto SCT and subsequent BV therapy (Cohort 1 equivalent)	On SoC after being identified as ineligible for auto SCT and then failed BV therapy (Cohort 2 equivalent)
CR	57% (range 10-80%, median=60%)	17% (range 0-60%, median=5%)
PR	44% (range 0-80%, median=40%)	13% (range 0-60%, median=0%)
SD	18% (range 0-50%, median=15%)	12% (range 0-50%, median=0%)
PD	12% (range 0-50%, median=0%)	11% (range 0-50%, median=0%)
CR, complete remission; PD, progressive disease PR, partial response; SD, stable disease		

4.12 Adverse reactions

4.12.1 KEYNOTE-087 Adverse reactions

The data presented below are from KEYNOTE-087 utilising a data cut-of the 25th September 2016 [66]. The results are presented for each of the Cohorts 1, 2, and 3 (Cohorts 1 and 2 directly relate to the patient population of this submission) and the overall population (Cohorts 1, 2, and 3).

Safety and tolerability were assessed by clinical and statistical review of all relevant parameters including AEs and laboratory test abnormalities during the treatment period up to the date cut-off date 25 September 2016. As per Section 4.4, The ASaT population consists of all enrolled patients who received at least one dose of study treatment.

Extent of exposure[66]

The duration of exposure was measured from the date of the first dose to the date of last dose. The median duration of exposure (median time on therapy) for the ASaT population in Cohort 1 and Cohort 2 was [redacted], respectively. Further information relating to number of administrations is reported in (Table 37).

Table 37. Summary of drug exposure by cohort – ASaT population [66]

	COHORT 1	COHORT 2	COHORT 3	Total
Subjects in population	**69**	**81**	**60**	**210**
Number of Days on Therapy (days)				
n	69	81	60	210
Mean	[redacted]	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]	[redacted]
Median	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Number of Administrations				
n	69	81	60	210
Mean	[redacted]	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]	[redacted]
Median	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
(Database Cutoff Date: 25SEP2016).				

Table 38. Clinical trial exposure to pembrolizumab by duration – ASaT population [66]

	Cohort 1		Cohort 2		Cohort 3		Total	
	(N=69)		(N=81)		(N=60)		(N=210)	
Duration of Exposure	n	Person-years	n	Person-years	n	Person-years	n	Person-years
> 0 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 1 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 3 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 6 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 12 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Each subject is counted once on each applicable duration category row. Duration of Exposure is calculated as (last dose date - first dose date +1)/365.25*12 (months). Database Cutoff Date: 25SEP2016								

Adverse Events (AEs)/ Grade 3-5 AEs[66]

Table 39 displays an overview of the numbers and percentages of patients in the ASaT population who had an AE up to 30 days and serious AEs (SAEs) up to 90 days after the last dose of study medication. Between-cohort differences were not expected, because patients in all the 3 cohorts received the same treatment and were similar with regard to disease status/prior transplant history.

In general, pembrolizumab was well tolerated by patients with RRcHL. Most AEs were of low-grade as evidenced by the relatively low rate of patients with AEs categorised as Grade 3, 4, or 5. The rate of AEs was not unexpected for this heavily treated patient population. Across cohorts, [redacted] patients experienced at least one AE (Table 39). The most common AEs (incidence >10% in one or more Cohort) included pyrexia [redacted] cough [redacted] and fatigue [redacted] (Table 40).

[redacted] patients experienced an AE categorised as Grade 3, 4, or 5, the majority of which were Grade 3 (Table 41). The most common Grade-3-5 AE was anaemia, which occurred in [redacted] patients.

In total [redacted] [redacted] of patients of Cohort 1 and 2, respectively discontinued treatment due to AEs (Table 39). Thus, pembrolizumab was generally well tolerated among RRcHL patients as evidenced by the low incidence of AEs that resulted in treatment discontinuation (Table 42).

Table 39. Summary of AEs by cohort – ASaT population [66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse event	[redacted]							
with drug-related† adverse events	[redacted]							
with toxicity grade 3-5 adverse events	[redacted]							
with toxicity grade 3-5 drug-related adverse events	[redacted]							
with non-serious adverse events	[redacted]							
with serious adverse events	[redacted]							
with serious drug-related adverse events	[redacted]							
who died	[redacted]							
who died due to a drug-	[redacted]							

related adverse event								
discontinued‡ due to an adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
discontinued due to a drug-related adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
discontinued due to a serious adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
discontinued due to a serious drug-related adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

† Determined by the investigator to be related to the drug.
‡ Study medication withdrawn.
Grades are based on NCI CTCAE version 4.0.
Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.
(Database Cutoff Date: 25SEP2016).

Table 40. KEYNOTE-087 subjects with AEs by decreasing incidence by Cohort (Incidence ≥10% in one or more Cohort) – AsaT population [66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
with no adverse events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cough	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypothyroidism	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rash	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dyspnea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oropharyngeal pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rhinitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Productive cough	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Every subject is counted a single time for each applicable specific adverse event.

A system organ class appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding.
Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.
(Database Cutoff Date: 25SEP2016).

Table 41. KEYNOTE-087 subjects with Grade 3-5 AEs by Cohort - Incidence >0% in one or more Cohort – AsaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Blood and lymphatic system disorders	[redacted]							
Anaemia	[redacted]							
Febrile neutropenia	[redacted]							
Leukopenia	[redacted]							
Lymphopenia	[redacted]							
Neutropenia	[redacted]							
Thrombocytopenia	[redacted]							
Cardiac disorders	[redacted]							
Myocarditis	[redacted]							
Pericarditis	[redacted]							
Stress cardiomyopathy	[redacted]							
Endocrine disorders	[redacted]							
Hypothyroidism	[redacted]							
Gastrointestinal disorders	[redacted]							
Abdominal pain	[redacted]							
Abdominal pain lower	[redacted]							
Colitis	[redacted]							
Diarrhoea	[redacted]							
Gastrointestinal pain	[redacted]							
Stomatitis	[redacted]							
General disorders and administration site conditions	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Fatigue	[redacted]							
Oedema peripheral	[redacted]							
Pyrexia	[redacted]							
Immune system disorders	[redacted]							
Cytokine release syndrome	[redacted]							
Graft versus host disease	[redacted]							
Infections and infestations	[redacted]							
Bacteraemia	[redacted]							
Bronchitis	[redacted]							
Bronchopulmonary aspergillosis	[redacted]							
Clostridium difficile colitis	[redacted]							
Device related infection	[redacted]							
Escherichia bacteraemia	[redacted]							
Herpes simplex	[redacted]							
Herpes zoster	[redacted]							
Lower respiratory tract infection	[redacted]							
Myelitis	[redacted]							
Pneumonia	[redacted]							
Respiratory tract infection	[redacted]							
Salmonellosis	[redacted]							
Septic shock	[redacted]							
Injury, poisoning and procedural complications	[redacted]							
Foot fracture	[redacted]							
Hip fracture	[redacted]							
Investigations	[redacted]							
Alanine aminotransferase increased	[redacted]							
Amylase increased	[redacted]							
Blood alkaline phosphatase increased	[redacted]							
Blood creatinine increased	[redacted]							
Investigations	[redacted]							
Lipase increased	[redacted]							
Platelet count	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
decreased								
Weight decreased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Metabolism and nutrition disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hyperglycaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hyperuricaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hypoalbuminaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hyponatraemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Arthritis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bone pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Muscle spasms	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Muscular weakness	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myositis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Osteonecrosis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rheumatoid arthritis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cancer pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myelodysplastic syndrome	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Epilepsy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Migraine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Transient ischaemic attack	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Schizophrenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Renal and urinary disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute kidney injury	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Reproductive system and breast disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Amenorrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Respiratory, thoracic and mediastinal disorders	[redacted]							
Cough	[redacted]							
Dyspnea	[redacted]							
Pulmonary embolism	[redacted]							
Skin and subcutaneous tissue disorders	[redacted]							
Dermatitis psoriasiform	[redacted]							
Lichenoid keratosis	[redacted]							
Skin ulcer	[redacted]							
Vascular disorders	[redacted]							
Aortic stenosis	[redacted]							
Hypertension	[redacted]							

Every subject is counted a single time for each applicable row and column.

A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.

Grades are based on NCI CTCAE version 4.0.

Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.

MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.

(Database Cutoff Date: 25SEP2016).

Table 42. Subjects With Adverse Events Resulting in Discontinuation by Cohort (Incidence > 0% in One or More Cohorts) - ASaT Population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Cardiac disorders	[redacted]							
Myocarditis	[redacted]							
Immune system disorders	[redacted]							
Cytokine release syndrome	[redacted]							
Infections and infestations	[redacted]							
Myelitis	[redacted]							
Injury, poisoning and procedural	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
complications								
Infusion related reaction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myositis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myelodysplastic syndrome	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Skin and subcutaneous tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Skin ulcer	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Every subject is counted a single time for each applicable row and column.

A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.

Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.

MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.

(Database Cut off Date: 25 SEP 2016).

Drug related AEs[66]

Adverse events considered by the Investigator to be "possibly," "probably," or "definitely" related to the study treatment are combined into the category of drug-related AEs. Table 43 displays the number and percentage of patients with drug-related AEs (incidence ≥5% in one or more Cohorts) by decreasing incidence (based on the total incidence) in the ASaT population. The number of patients who experienced a drug related AE in Cohort 1 and 2 was as follows: Cohort 1, [redacted] patients [redacted] in Cohort 2, [redacted] patients [redacted]

The most commonly reported drug-related AEs (reported in ≥5% of patients in one or more of the Cohorts) were: hypothyroidism, pyrexia, fatigue, rash, diarrhoea, headache, cough, nausea, neutropenia, infusion related reaction, arthralgia, muscle spasms, vomiting, dyspnea, upper respiratory tract infection, and pneumonitis (Table 43). Most drug-related AEs were low grade (Grade 1 or 2). This is evidenced by the low number of patients reporting discontinuation due to drug related AEs as summarised in Table 44.

Table 43. Subjects With Drug-Related Adverse Events decreasing by incidence by Cohort (Incidence ≥ 5% in one or more Cohorts) – ASaT population [66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Hypothyroidism	[redacted]							
Pyrexia	[redacted]							
Fatigue	[redacted]							
Rash	[redacted]							
Diarrhoea	[redacted]							
Headache	[redacted]							
Cough	[redacted]							
Nausea	[redacted]							
Neutropenia	[redacted]							
Infusion related reaction	[redacted]							
Arthralgia	[redacted]							
Muscle spasms	[redacted]							
Vomiting	[redacted]							
Dyspnea	[redacted]							
Upper respiratory tract infection	[redacted]							
Pneumonitis	[redacted]							

Every subject is counted a single time for each applicable specific adverse event.
A system organ class appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding.
Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.
(Database Cutoff Date: 25SEP2016).

Table 44. Subjects With Drug-Related Adverse Events Resulting in Treatment Discontinuation (Incidence > 0% in One or More Cohorts) – AsaT population [66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Cardiac disorders	[redacted]							
Myocarditis	[redacted]							
Immune system disorders	[redacted]							
Cytokine release syndrome	[redacted]							
Infections and infestations	[redacted]							
Myelitis	[redacted]							
Injury, poisoning and procedural complications	[redacted]							
Infusion related reaction	[redacted]							
Musculoskeletal and connective tissue disorders	[redacted]							
Myositis	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Pneumonitis	[redacted]							

Every subject is counted a single time for each applicable row and column.

A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.

Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.

(Database Cutoff Date: 25SEP2016).

Grade 3 to 5 Drug related AEs[66]

Table 45 displays the number of patients with drug-related Grade 3 to 5 AEs (incidence >0% in one or more Cohort), and shows that the majority of patients [redacted] did not experience drug-related Grade 3 to 5 AEs.

The most commonly reported drug-related Grade 3 to 5 AE (reported in >0% of patients in one of the Cohorts) was neutropenia at [redacted] in Cohort 1 (Table 45).

Table 45. Subjects With Grade 3-5 Adverse Events (Incidence > 0% in One or More Cohorts) – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Blood and lymphatic system disorders	[redacted]							
Neutropenia	[redacted]							
Thrombocytopenia	[redacted]							
Cardiac disorders	[redacted]							
Myocarditis	[redacted]							
Pericarditis	[redacted]							
Endocrine disorders	[redacted]							
Hypothyroidism	[redacted]							
Gastrointestinal disorders	[redacted]							
Colitis	[redacted]							
Diarrhoea	[redacted]							
Gastrointestinal pain	[redacted]							
General disorders and administration site conditions	[redacted]							
Fatigue	[redacted]							
Oedema peripheral	[redacted]							
Pyrexia	[redacted]							
Immune system disorders	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Cytokine release syndrome	[redacted]							
Infections and infestations	[redacted]							
Herpes simplex	[redacted]							
Herpes zoster	[redacted]							
Lower respiratory tract infection	[redacted]							
Myelitis	[redacted]							
Investigations	[redacted]							
Amylase increased	[redacted]							
Lipase increased	[redacted]							
Weight decreased	[redacted]							
Metabolism and nutrition disorders	[redacted]							
Decreased appetite	[redacted]							
Musculoskeletal and connective tissue disorders	[redacted]							
Arthralgia	[redacted]							
Arthritis	[redacted]							
Bone pain	[redacted]							
Myositis	[redacted]							
Rheumatoid arthritis	[redacted]							
Nervous system disorders	[redacted]							
Epilepsy	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Cough	[redacted]							
Dyspnea	[redacted]							
Skin and subcutaneous tissue disorders	[redacted]							
Dermatitis psoriasiform	[redacted]							
Lichenoid keratosis	[redacted]							

Every subject is counted a single time for each applicable row and column.
A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.
Grades are based on NCI CTCAE version 4.0.

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included. (Database Cutoff Date: 25SEP2016).								

Serious AEs (SAEs)[66]

Table 46 shows the incidence of SAEs regardless of causality. A SAE was defined as any AE that occurred during the use of pembrolizumab that resulted in: death, was life threatening, resulted in persistent or significant disability/incapacity, resulted in, or prolonged, an existing in-patient hospitalisation, was a congenital anomaly/birth defect, or was considered as another important medical event. In addition the following events, specified by the sponsor, were also considered; cancer, or associated with an overdose.

Overall [redacted] patients in Cohort 1 and [redacted] patients in Cohort 2 experienced a SAE up to 90 days after the last dose of pembrolizumab. The most common SAEs are summarised in Table 46 below. Discontinuation due to SAEs occurred in [redacted] and [redacted] patients in Cohort 1 and Cohort 2, respectively (Table 39).

Table 46. Subjects With Serious Adverse Events Up to 90 Days After Last Dose (Incidence > 0% in One or More Cohorts) – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Blood and lymphatic system disorders	[redacted]							
Anaemia	[redacted]							
Cardiac disorders	[redacted]							
Myocardial infarction	[redacted]							
Myocarditis	[redacted]							
Pericarditis	[redacted]							
Stress cardiomyopathy	[redacted]							
General disorders and administration	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
site conditions								
Hyperthermia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Immune system disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cytokine release syndrome	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Graft versus host disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bronchitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clostridium difficile colitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Device related infection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Escherichia bacteraemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Herpes zoster	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Influenza	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lower respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myelitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Respiratory syncytial virus infection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Salmonellosis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Septic shock	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hip fracture	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infusion related reaction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Myositis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Osteonecrosis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
(incl cysts and polyps)								
Myelodysplastic syndrome	[redacted]							
Squamous cell carcinoma	[redacted]							
Squamous cell carcinoma of skin	[redacted]							
Nervous system disorders	[redacted]							
Epilepsy	[redacted]							
Headache	[redacted]							
Psychiatric disorders	[redacted]							
Schizophrenia	[redacted]							
Renal and urinary disorders	[redacted]							
Acute kidney injury	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Dyspnea	[redacted]							
Pneumonitis	[redacted]							
Pneumothorax	[redacted]							
Pulmonary embolism	[redacted]							
Skin and subcutaneous tissue disorders	[redacted]							
Skin lesion	[redacted]							
Skin ulcer	[redacted]							

Every subject is counted a single time for each applicable row and column.
A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.
(Database Cutoff Date: 25SEP2016).

Drug related Serious AEs (SAEs)[66]

Table 47 shows that the incidence of drug-related SAEs. As reported in Table 39 the number of patients that discontinued due to a drug-related SAE was small, with only [redacted] each in Cohort 1, and Cohort 2.

Table 47. Subjects with drug-related Serious Adverse Events Up to 90 Days After Last Dose (Incidence > 0% in One or More Cohorts) – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Cardiac disorders	[redacted]							
Myocarditis	[redacted]							
Pericarditis	[redacted]							
Immune system disorders	[redacted]							
Cytokine release syndrome	[redacted]							
Infections and infestations	[redacted]							
Herpes simplex	[redacted]							
Herpes zoster	[redacted]							
Lower respiratory tract infection	[redacted]							
Myelitis	[redacted]							
Injury, poisoning and procedural complications	[redacted]							
Infusion related reaction	[redacted]							
Musculoskeletal and connective tissue disorders	[redacted]							
Myositis	[redacted]							
Nervous system disorders	[redacted]							
Epilepsy	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Dyspnea	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Pneumonitis	[redacted]							

Every subject is counted a single time for each applicable row and column.
A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.
(Database Cutoff Date: 25SEP2016).

Summary of deaths[66]

Overall [redacted] died across the entire study population (Table 39); this was [redacted] in Cohort 2, and [redacted] in Cohort 3. Although both deaths were related to AEs, neither death was considered drug related (Table 48).

Table 48. Subjects With Grade 5 Adverse Events (Incidence > 0% in One or More Cohorts) – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse events	[redacted]							
Immune system disorders	[redacted]							
Graft versus host disease	[redacted]							
Infections and infestations	[redacted]							
Septic shock	[redacted]							

Every subject is counted a single time for each applicable row and column.
A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns meets the incidence criterion in the report title, after rounding.
Grades are based on NCI CTCAE version 4.0.
Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.
MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded.
(Database Cutoff Date: 25SEP2016).

Adverse Events of Special Interest (AEOSI)[66]

Overall, AEOSI occurred in [redacted] patients in Cohort 1 and 2, respectively (Table 49). Of these, [redacted] patients in Cohort 1 and [redacted] patients in Cohort 2 had drug-related AEOSI (Table 50).The median time to onset of first AEOSI [redacted] days in Cohort 1, and [redacted] in Cohort 2[66].

Of the 210 patients treated with pembrolizumab, ten were reported as having received alloSCT at some point after stopping treatment with pembrolizumab (data cut-off September 2016). A total of [redacted] patients [redacted] reported GVHD, of which [redacted] was fatal and the [redacted] were considered recovered/ resolved [67].

Table 49. Adverse Event Summary for AEOSI by Cohort – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
with no adverse event	[redacted]							
with drug-related† adverse events	[redacted]							
with toxicity grade 3-5 adverse events	[redacted]							
with toxicity grade 3-5 drug-related adverse events	[redacted]							
with non-serious adverse events	[redacted]							
with serious adverse events	[redacted]							
with serious drug-related adverse events	[redacted]							
who died	[redacted]							
who died due to a drug-related adverse event	[redacted]							
discontinued‡ due to an adverse event	[redacted]							
discontinued due to a drug-related adverse event	[redacted]							
discontinued due to a serious adverse event	[redacted]							
discontinued due to a serious drug-related adverse event	[redacted]							

† Determined by the investigator to be related to the drug.

‡ Study medication withdrawn.

Grades are based on NCI CTCAE version 4.0.

Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.

(Database Cutoff Date: 25SEP2016).

Table 50. Subjects With Adverse Events by Maximum Toxicity Grade (Incidence >0% in One or More Cohorts) AEOSI – ASaT population[66]

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Subjects in population	69		81		60		210	
with one or more adverse events	[redacted]							
Grade 1	[redacted]							
Grade 2	[redacted]							
Grade 3	[redacted]							
with no adverse events	[redacted]							
Endocrine disorders	[redacted]							
Hyperthyroidism	[redacted]							
Grade 1	[redacted]							
Grade 2	[redacted]							
Hypothyroidism	[redacted]							
Grade 1	[redacted]							
Grade 2	[redacted]							
Grade 3	[redacted]							
Eye disorders	[redacted]							
Iridocyclitis	[redacted]							
Grade 2	[redacted]							
Iritis	[redacted]							
Grade 2	[redacted]							
Gastrointestinal disorders	[redacted]							
Colitis	[redacted]							
Grade 2	[redacted]							
Grade 3	[redacted]							
Enterocolitis	[redacted]							
Grade 1	[redacted]							
Immune system disorders	[redacted]							
Cytokine release syndrome	[redacted]							
Grade 1	[redacted]							
Grade 3	[redacted]							
Drug hypersensitivity	[redacted]							
Grade 2	[redacted]							
Hypersensitivity	[redacted]							
Grade 1	[redacted]							
Grade 2	[redacted]							
Injury, poisoning and procedural complications	[redacted]							
Infusion related reaction	[redacted]							
Grade 1	[redacted]							
Grade 2	[redacted]							
Musculoskeletal and connective tissue disorders	[redacted]							
Myositis	[redacted]							

	COHORT 1		COHORT 2		COHORT 3		Total	
	n	(%)	n	(%)	n	(%)	n	(%)
Grade 2	[redacted]							
Grade 3	[redacted]							
Respiratory, thoracic and mediastinal disorders	[redacted]							
Pneumonitis	[redacted]							
Grade 2	[redacted]							
Skin and subcutaneous tissue disorders	[redacted]							
Dermatitis psoriasiform	[redacted]							
Grade 3	[redacted]							

Every subject is counted a single time for each applicable specific adverse event. A subject with multiple adverse events within a system organ class is counted a single time for that system organ class.

A system organ class or specific adverse event appears on this report only if its incidence in one or more of the columns is greater than or equal to the incidence specified in the report title, after rounding.

Only the highest reported grade of a given adverse event is counted for the individual subject.

Grades are based on NCI CTCAE version 4.0.

Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included.

(Database Cutoff Date: 25SEP2016).

4.12.2 Studies that report additional adverse reactions to those reported in 4.12.1

The search strategy as reported in Section 4.1 did not identify any relevant articles.

For completeness the safety results of a phase Ib study of pembrolizumab in patients with relapsed or refractory disease, who had relapsed after, or were considered ineligible for, or refused autoSCT is summarised below (KEYNOTE-013; NCT01953692)[39]. The dosing of pembrolizumab within this trial does not support the licensed EMA recommendation, and was therefore excluded from the decision problem.

Overall pembrolizumab was associated with a favourable safety profile. AEs of any grade and attribution were reported in 30 of the 31 patients (97%). Overall, 68% of patients experienced one or more AEs that were deemed related to study treatment. Two patients discontinued treatment because of an AE (grade 2 pneumonitis and grade 3 nephrotic syndrome), and both of these patients received steroids for treatment of the AE. There were no grade 4 treatment-related AEs and no deaths related to study treatment. No instances of treatment-related hepatitis, hypophysitis, or uveitis were reported[39].

4.13 Interpretation of clinical effectiveness and safety evidence

4.13.1 Statement of principal findings from the clinical evidence highlighting the clinical benefits and harms of the technology.

As described in Section 4.0 the efficacy of pembrolizumab for patients with RRcHL is clinically meaningful across both patient groups; i.e. those that have failed autoSCT and subsequent BV (Cohort 1), and for those who are considered ineligible for autoSCT and have failed subsequent BV (Cohort 2).

At a median follow-up of 15.9 months the ORR was 75.4% and 66.7% for Cohort 1 and 2, respectively, with many patients reporting CR (27.5% and 24.7%)[45]. These response rates have impacted the incidence of progression and survival; with very few patients (Cohort 1, n=3; Cohort 2, n=5) dying. Results from the naive indirect comparison and MAIC show that pembrolizumab has significantly improved ORR versus the mixed treatment SoC.

Progression free survival, although immature, reached a median of 16.7 months (95% CI 11.2, not reached) in Cohort 1 and 11.1 months (95% CI 7.6, 13.7) in Cohort 2[45]. Similarly, the results of the naïve indirect comparison show a statistically significant reduction in the number of events among patients treated with pembrolizumab versus mixed treatment SoC. Further, using the Kaplan-Meier method, the OS rates at six through to 15 months are in excess of [redacted]; median OS [redacted] Cohort or within the overall study population.

The safety profile of pembrolizumab can be considered acceptable in the context of alternative therapies such as standard chemotherapy regimens. The data presented from KEYNOTE-087 show that the majority of AE experience were low grade, and did not result in study discontinuation[66]. Overall, there were low levels of patient discontinuation due to SAEs or AEOSI. Furthermore, mortality rates were low and [redacted] considered to be study drug related[66].

Finally, improvements in QoL measurements were observed from baseline using the disease-specific patient QoL measure EORTC-QLQ-C30 and the generic measure EQ-5D, demonstrating clinically significant benefits on both scales.

4.13.2 Discussion of the strengths and limitation of the clinical evidence base for the technology.

Internal validity

KEYNOTE-087 is a phase II single arm non-randomised multi-centre study of pembrolizumab 200mg Q3W. This trial design reflects the limited number of patients eligible and also the absence of a formalised treatment pathway. Until recently BV was not recommended for use by NICE, which is required within the treatment pathway before the use of pembrolizumab.

Patients enrolled within the KEYNOTE-087 trial, in terms of baseline characteristics, can be considered broadly representative of the UK population. The enrolment criteria of KEYNOTE-087 required patients to have received BV, which following the publication of TA446 [32] is now aligned with the UK treatment pathway.

The efficacy endpoints considered within the trial and the comparative clinical effectiveness analysis are clinically relevant and directly referenced in the final scope for this appraisal. The endpoints selected are consistent with those used in studies of other therapeutic agents in the population of RRcHL.

HRQoL was an exploratory endpoint of the KEYNOTE-087 study with changes from baseline in patients treated with pembrolizumab used both the preferred measure of EQ-5D according to the NICE reference case, in addition to the cancer specific EORTC-QLQC30.

Although this was a single arm, non-comparative trial assessments and results presented were conducted using BICR, in order to minimise bias.

External validity

KEYNOTE-087 is a global study conducted in 47 centres, of which 23 were in Europe including three UK site enrolling 14 patients.

Baseline characteristics of patients enrolled in KEYNOTE-087 were as expected for patients with RRcHL with the predominant subtype (~80%) nodular sclerosis. The majority of patients were male, aged greater than 65 years, White, and reported a median of four prior lines of therapy. All patients had received prior BV, and prior to that had either failed autoSCT or were considered ineligible for autoSCT.

The observed safety profile of pembrolizumab in KEYNOTE-087 was consistent with that seen previously with pembrolizumab for the treatment of other types of tumours [68-74].

4.13.3 Application of NICE end-of-life criteria to pembrolizumab use in relapsed/ refractory Classica Hodgkin Lymphoma

The evidence presented in this submission highlight the paucity of data relevant to this small patient group. Furthermore, the results of the clinician survey (Section 4.11.1), and the recent stakeholder feedback, as per the committee meeting papers of ID972, suggest there is a substantial unmet need for patients with RRcHL who have failed/ considered ineligible for autoSCT and subsequent BV treatment[75]. The case for end-of-life criteria is reported in Table 51.

Recently, the NICE appraisal committee for nivolumab (TA462) concluded that within this indication (relevant to the current decision problem) it was plausible that the criteria for short life expectancy could apply[57]. Therefore, EoL criteria were factored into its decision making[57].

Table 51. End-of-life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	In summary, the literature does not support a valid estimate of OS for patients with RRcHL as expected within UK clinical practice. However, estimates provided below provide some reassurance that OS ranges from 17.1 months to 19 months. As reported in Section 4.2, results of the SLR show that there is a paucity of UK specific data relevant for the patient populations considered within this submission. There is a general consensus that treatment options available at this later line of therapy, i.e. among those patients with RRcHL is limited and associated with poor outcomes. A recent article by Bair et al. 2017 suggests that the OS from disease progression among 87 patients with RRcHL post-ASCT could reach up to 26.1 months (95% CI: 20.4–45.9 months)[9]. Within the same patient population, analyses showed that when excluding novel agents the OS reduced to 17.1 [9]. However, these estimates do not represent the severity of the patient population within this submission who are more advanced. Therefore, within the context of this submission, those patients who relapse/ are refractory/ or considered ineligible for treatment post-autoSCT may experience lower rates of survival. These outcome are supported by a UK clinician survey (n=16). Clinician reported that only a minority of patients with RRcHL who have either failed/ considered ineligible for autoSCT and have failed subsequent BV experience ORR, 35% and 31%, respectively. Within the same two patients groups clinicians reported that current SoC followed by alloSCT provided a median OS of 18.9 months and 14.2 months, respectively (MSD data on file[64]). Recently, the NICE considered (ID 972) the relevance of Cheah et al 2016, which reported OS estimates of around two years. The appraisal committee agreed that although these data were not exactly generalisable to the UK setting, they were suitable for decision making[57]. However, this estimate of OS was skewed by the inclusion of investigational agents (47.4 months), when investigation agents are removed the median estimate of OS reduces to around 19 months[44]. Again, these estimates are broadly comparable to the findings reported above.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	As of March 21st 2017 [redacted] for Cohorts 1 and 2. However, the small number of deaths reported during the current follow-up period (15.9 months) indicates a substantially longer median survival than that offered by current therapies. The OS rate at 15 months in cohort 1 and 2 was reported using Kaplan-Meier estimates at [redacted], respectively[45, 46].

.

4.14 Ongoing studies

KEYNOTE-087 is an on-going, single arm, non-comparative, phase II study of pembrolizumab in patients with RRcHL who have failed BV having either failed autoSCT or were considered ineligible for autoSCT. All available data from this trial have been included within this submission.

KEYNOTE-204 (NCT02684292) is an on-going phase III, randomised, non-blinded, active control study of pembrolizumab versus BV in patients with RRcHL. However, this trial does not represent the indication/ license considered within this submission.

After searching the International Clinical Trials Registry platform for ongoing clinical trials, there were 112 hits captured (Search strategy and results can be found in Appendix 2. A total of 85 records were removed for being a duplicates, 13 for not being ongoing, and 13 were excluded based on population, resulting in 1 ongoing trial (NCT03077828) of interest.

This single-arm open-label phase II study is recruiting patients with RRcHL (estimated enrolment of 40) to be treated with pembrolizumab + carboplatin + etoposide + ifosfamide. The trial started in April 2017 and the estimated study completion date is February 2020. To be included, patients must have relapsed/ refractory disease, with at least one line of prior chemotherapy, but two or less prior lines of treatment. The only limitation to prior treatment is the use of immune checkpoint inhibitors. As prior treatment with BV is not explicitly mentioned in the inclusion criteria, this study may be relevant to the decision problem.

5 Cost effectiveness

5.1 Published cost-effectiveness studies

5.1.1 Strategies used to retrieve cost-effectiveness studies relevant to decision-making in England

In line with the NICE guide to methods of technology appraisal [76], an SLR was conducted to identify cost-effectiveness studies from the published literature between 2001 and 12th July 2017. The target population in this submission is patients with relapsed or refractory classic Hodgkin's Lymphoma (RRcHL) which formed the basis for the search in order to identify all relevant data that could inform the development and population of the model.

The first stage in the review was to identify all relevant economic evidence for the comparator treatments by implementing comprehensive searches. The following research questions were posed in accordance with the decision problem:

- What is the cost-effectiveness of comparator therapies to pembrolizumab in treating patients with RRcHL?
- What is the health-related quality of life (in terms of utilities) associated with RRcHL?
- What are the resource requirements and costs associated with the treatment of RRcHL?

A comprehensive literature search was carried out using the following electronic databases and is presented in Appendix 12. Details of the search strategies conducted for the health related quality of life and utilities and costs are also provided in Appendix 12.

- MEDLINE and MEDLINE In-process (using Embase.com)
- EconLit
- EMBASE (using Embase.com)
- The Cochrane Library, including NHS EED and HTA databases

Manual searches were also performed in the American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO) and International Society for Pharmacoeconomics and Outcomes Research (ISPOR). The manual searches were limited to the most recent 2 years.

In addition to the formal literature search and manual searches, the National Institute for Health and Care Excellence (NICE) website was searched to identify relevant information from previous submissions not otherwise captured.

All retrieved studies were reviewed and quality checked by an independent researcher and assessed against the eligibility criteria presented in Table 52. These selection criteria are detailed below for the cost-effectiveness search. The other two literature searches relative to the costs and health related quality of life and utilities are provided in Appendix 12 and detailed in section 5.4.1 and 5.5.2.

Table 52: Inclusion and exclusion criteria for cost-effectiveness studies

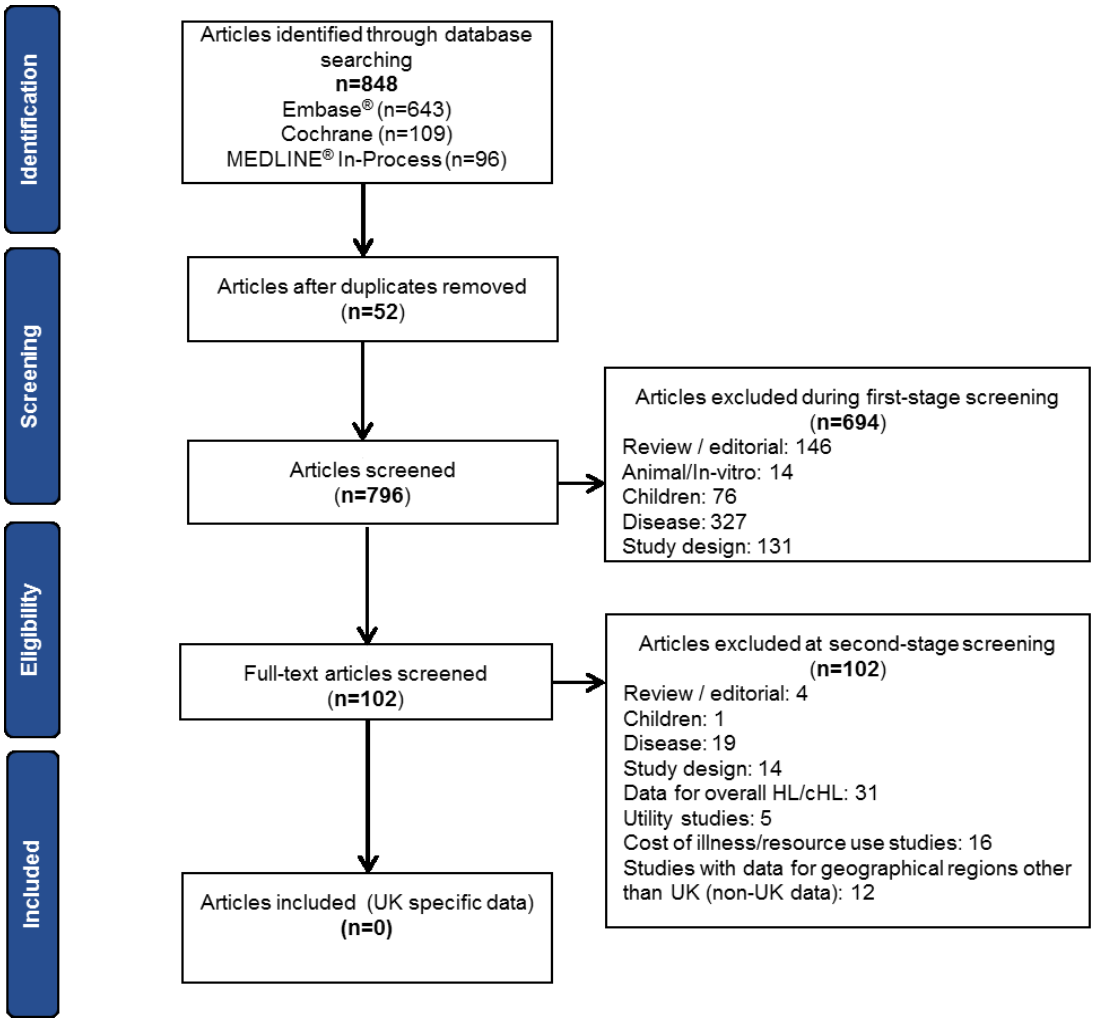
Criteria	Inclusion	Exclusion	Rationale
Population	• Adult (age ≥18 years) patients with relapsed/refractory cHL, irrespective of age or gender	• Patients under the age of 18 • Disease other than relapsed/refractory cHL	• The relevant patient population of interest to the review
Intervention/ Comparator	• No restriction on inclusion of studies based on interventions or comparators • All pharmacological interventions to be captured	• Studies assessing non-drug treatments (e.g. surgery, radiotherapy)	• To have a holistic overview of the available literature in relapsed or refractory cHL settings.
Outcomes	• Studies including a comparison of benefits and costs between the intervention and comparator arms. • Results should be expressed in incremental costs, ICER, QALYs, LYG, or any other measure of effectiveness reported together with costs	• Cost and resource use only • Utility data only	• To identify relevant cost-evaluation outcomes
Study type	• Cost effectiveness analysis • Cost utility analysis • Cost benefit analysis • Cost minimisation analysis • Budget impact models • Cost consequence studies	**Other study designs:** • Epidemiology studies • Clinical studies • Pharmacokinetic/Pharmacodynamic (Animal/*in-vitro*) study • General quality of life studies	• To identify relevant cost-evaluation studies
Publication type	• Economic evaluation studies	• Letters, editorials, notes, and reviews (systematic or otherwise)	• To identify primary study articles

Time-frame	• Evidence published from 2001 onwards	• Evidence published prior to 2001	• To ensure recent economic models are included and limit the number of studies identified to those most relevant to the decision problem
Language	• Studies with full text available in English language	• Studies published in non-English language	• To ensure the studies can be correctly understood and interpreted
Subgroup data of interest	• Studies including ≥80% of study population that qualifies the disease criteria	• Studies which enrol a mixed population of relapsed or refractory cHL and other types of Lymphomas, but not providing subgroup data for population of interest • Studies enrolling a mixed patient population of children and adults, but not providing sub-group data for adult population	• To ensure that review included data specific to population of interest
Data specific to relapsed/refractory HL	• Studies reporting data for relapsed/refractory HL or cHL	• Studies reporting data for early stage HL/cHL	• To ensure that review included data specific to population of interest • This data could be used as proxy, in case of limited evidence of relapsed/refractory population
Country	• Studies reporting UK specific data	• Studies reporting data for geographical regions, other than UK (non-UK data)	• To identify UK specific data

5.1.2 Brief description of identified cost-effectiveness studies

Of a total of 848 potentially relevant papers or abstracts were identified, no cost-effectiveness studies in patients with RRcHL were found that met all the inclusion criteria. Thus, a summary list of published cost-effectiveness studies has not been compiled. The PRISMA flow diagram is presented in Figure 12.

Figure 12: PRISMA diagram – Economic evaluation review



5.1.3 Complete quality assessment for each relevant cost-effectiveness study identified

This is not applicable as no cost-effectiveness study meeting all the inclusion criteria was identified, indicating a de novo cost-effectiveness model is required to assess the cost-effectiveness of pembrolizumab compared with the relevant comparators.

5.2 De novo analysis

A UK clinician advisory board conducted in February 2017 indicated that UK clinical practice in this patient population is with the aim to reach a CR, PR or SD and then to transplant using allogeneic stem cells. A UK clinician survey (n=16) also supports this with 75% reporting they have transplanted using allogeneic stem cells [64]. Further information on the clinician survey can be found in section 4.11. Clinicians agreed that this could be particularly important in autoSCT ineligible patients as currently they may feel obliged to try autoSCT to ensure they were not excluded from treatments and that due to the lack of current treatment options in this patient population to allow patients to achieve an adequate response, clinicians felt that pembrolizumab would most likely be used as a bridging treatment to allow alloSCT [40]. Clinicians also suggested that they would use treatments such as pembrolizumab for the maximum permitted time period in those who could not withstand an alloSCT provided they could tolerate this.

A de novo economic analysis was performed to assess the incremental cost-effectiveness of pembrolizumab monotherapy versus standard of care (SoC) within its marketing authorisation for relapsed refractory classical Hodgkin Lymphoma (RRcHL). A de novo analysis was required because of the absence of published cost-effectiveness studies for pembrolizumab in RRcHL.

In line with the recent NICE appraisal for nivolumab in RRcHL (TA462) [75], it is expected that pembrolizumab monotherapy will be used as a “bridge” to alloSCT, where the aims of treatment are to control the disease, and if possible, elicit a disease response to enable alloSCT. To estimate the lifetime cost-effectiveness of “bridging” therapy in RRcHL, a de novo cohort based decision analytical model was developed with states based on response, uptake of alloSCT, and survival.

The de novo model captures:

- The initial aim of treatment in the form of eliciting disease response, measured in terms of the overall response at week 12
- Uptake of alloSCT conditional on response status
- The quality of life (QoL) implications of tumour response
- Survival, cost and QoL implications of alloSCT
- Survival, cost and QoL implications of continuation of pembrolizumab, or SoC in those unable to receive alloSCT

Further detail on each aspect is provided in later sections of the submission.

5.2.1 Patient population

The marketing authorisation for pembrolizumab monotherapy in RRcHL is for the treatment of adult patients with RRcHL who have failed autoSCT and BV, or who are autoSCT ineligible and have failed BV.

In line with the final scope [77] of the appraisal, the eligible population are defined as two distinct populations:

- Cohort 1: RRcHL who have failed autoSCT and BV
- Cohort 2: RRcHL who are autoSCT ineligible and have failed BV

The economic analysis focuses on the use of pembrolizumab in each of the two cohorts listed above.

The cost-effectiveness analysis was modelled on the characteristics of cohort 1 and cohort 2 and the main body of clinical evidence derived from KEYNOTE-087, which included the two populations of RRcHL patients aforementioned[48] compared with the real world evidence study, Cheah et al [44].

A summary of the characteristics of cohort 1 and 2 of KEYNOTE-087 (combined and as individual cohorts) and the Cheah et al population is provided in Table 53.

Table 53. Baseline characteristics of patients included in the model

Characteristic		KEYNOTE-087, Cohort 1&2[48]	KEYNOTE-087, Cohort 1[48]	KEYNOTE-087, Cohort 2[48]	Cheah et al. (2016) [44]
Treatment		Pembrolizumab 200mg			Mix of therapies including chemotherapy, and investigational agents
Number of patients		150	69	81	89[α]
Age (median)		37.5	34.0	40.0	32
Female (%)		71 (47.3%)	33 (47.8%)	38 (46.9%)	46 (47%)
ECOG	0	73 (48.7%)	29 (42.0%)	44 (54.3%)	33 (41%)
	1	76 (50.7)%	39 (56.5%)	37 (45.7%)	44 (54%)
	2	1 (0.7%)	1 (1.4%)	0 (0.0%)	3 (4%)
Baseline B symptoms		48 (32.0%)	22 (31.9%)	26 (32.1%)	7 (8%)
Bulky Lymphadenopathy		16 (10.7%)	5 (7.2%)	11 (13.6%)	15 (37%)
Bone marrow involvement		8 (5.3%)	3 (4.3%)	5 (6.2%)	NR
Disease status - relapse		70 (46.7%)	46 (66.7%)	24 (29.6%)	NR

Disease status – refractory	80 (53.3%)	23 (33.3%)	57 (70.4%)	NR
Previous BV therapy	150 (100%)	69 (100.0%)	81 (100.0%)	89 (100%)
Prior autoSCT	69 (46.0%)	69 (100.0%)	0 (0.0%)	66 of 97(68%)
Prior radiation	52 (34.7%)	31 (44.9%)	21 (25.9%)	NR
Median no. of prior line of therapy	4	4	4	4
*Calculated; [α] not all characteristics were available from this sample. BV: Brentuximab vedotin; ECOG: Eastern Cooperative Oncology Group; N/A: Not applicable; NR: Not reported; SCT: Stem cell transplant				

The populations in KEYNOTE-087 and Cheah et al differ in age, ECOG performance status, baseline B symptoms, use of Bulky Lymphadenopathy, and prior autoSCT use.

The study population in Cheah et al comprised a mix of patients who had received prior autoSCT (68%) and those who had not. The corresponding rates of prior autoSCT in cohort 1 and cohort 2 of KEYNOTE-087 were 100% and 0% respectively. The mix of prior autoSCT in the combined cohort 1 and 2 was 46%.

The generalisability of the Cheah et al population to UK practice was considered in the committee deliberations for TA462[75] (nivolumab in RRcHL), where the study was judged to not reflect UK practice, in part, because of expert testimony that subsequent rates of alloSCT in the UK would exceed those in the US. It was however noted that Cheah et al was the best available evidence for standard of care, and the most appropriate dataset for SoC for use in an indirect comparison.

5.2.2 Model structure

A cohort based decision analytical model was developed in Microsoft Excel® using standard Excel® functions and visual basic for applications consisting of two related decision models:

- A short-term decision-tree model to predict response and alloSCT uptake of the population during the first 12 weeks of treatment.
- A set of Markov state transition models to predict the lifetime survival of patients from Week 12 to death, conditional on alloSCT uptake or continued use of pembrolizumab or SoC.

When combined, these models provide an estimate of the lifetime costs and effectiveness of either a “bridging” treatment to alloSCT in RRcHL or continued treatment with pembrolizumab

or SoC when this is not possible which is in line with the clinical pathway of care and health states experienced by the patient population described in section 3.3.

The model submitted for this appraisal uses a state transition approach, which is fundamentally different to the partitioned survival approach commonly used in advanced oncology. Recently, the decision support unit at NICE critically appraised the use of partitioned survival models within the technology appraisal process [78]. A fundamental limitation of this approach highlighted by the DSU is the structural assumption that overall survival, a key driver of QALY gains in advanced oncology, is modelled independently of an underlying disease model. Specifically, the partitioned survival approach relies on the extrapolation of within-trial mortality rates without an explicit link to the mechanism of drug effect, such as delayed progression or response leading to alloSCT use. The DSU argues that ignoring information on the treatment effect mechanism and focusing on observed time-trends in survival within the trial period may result in inappropriate extrapolations. The company further argues that inappropriate extrapolations may be more likely in cases that include a complex "downstream" disease pathway that includes treatment with curative intent (e.g. alloSCT). This is because within-trial mortality trends are unlikely to represent "curative" trends expected from those who transplant. The lack therefore of an explicit way of linking treatment effect mechanism to long-term survival in partitioned survival analysis precluded its use in this appraisal. A method that links response to alloSCT use and subsequent outcomes was hence preferred, and developed using a transition state approach as recommended by the DSU. Further detail on the clinical justification of the model structure is provided below.

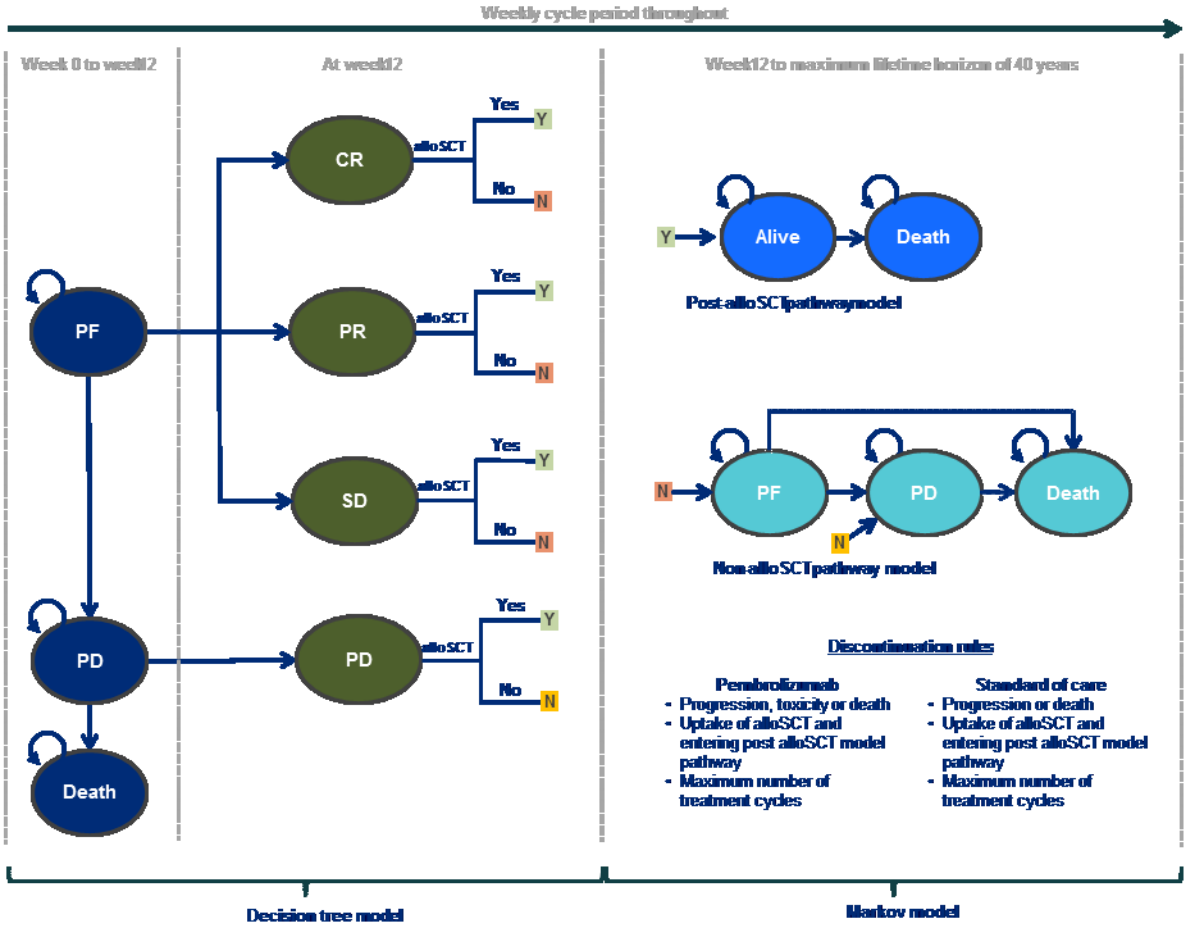
In line with the NICE reference case, cost-effectiveness was assessed in terms of the cost per Quality Adjusted-Life Years (QALY) gained. Both costs and health outcomes were discounted at a rate of 3.5% per annum.

A weekly cycle length was used to accurately predict the number of patients treated with pembrolizumab based on its recommended posology of infusions every 3 weeks, and with SoC based on treatment intervals that vary from every 2 to 4 weeks. Costs and health outcomes were calculated using lifetable mid-cycle estimates, with the exception of:

- one-off costs for subsequent treatment and AEs, which were applied at the start of the model time horizon
- the costs for drug acquisition and administration were calculated using the number of patients occupying the progression-free (PF) state at the start of each relevant cycle, to reflect that therapy is given at fixed and discrete time points (e.g. every 3 weeks).

A lifetime horizon of 40 years was used in the base case analysis Figure 13 shows the model structure.

Figure 13. Overview of model structure including decision-tree and the lifetime Markov state transition models



Short-term decision tree model for weeks 0 - 12

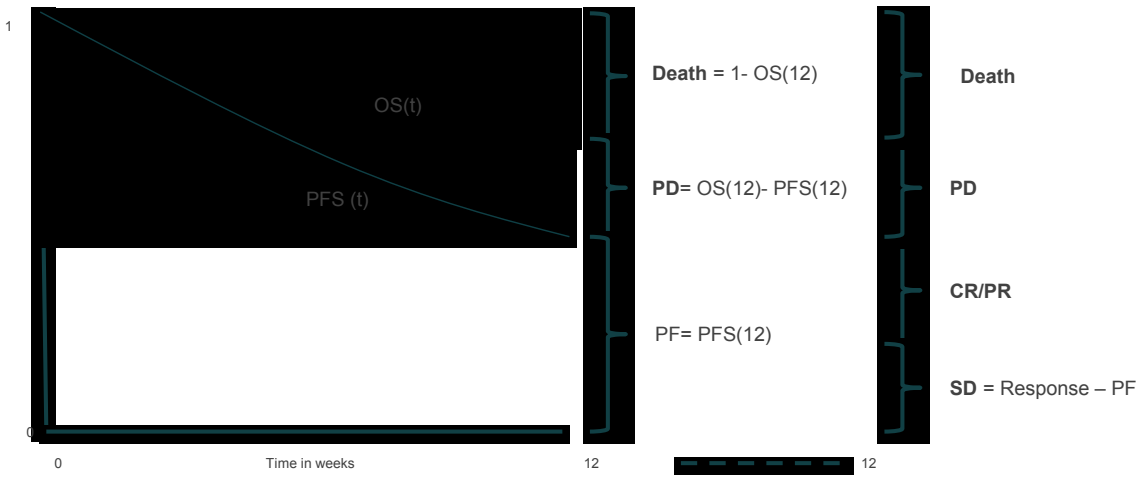
The short-term decision tree component of the model consists of two chance nodes that represent the outcomes of the initial 12 weeks of treatment.

The first chance node represents the outcomes of treatment in terms of response (complete or partial), stable disease (SD), progressed disease (PD) or death after the first 12 weeks. The proportion of patients occupying each state prior to the first chance node (PF, PD and death) is modelled via a partitioned survival (or area under the curve) technique using data on the progression free survival (PFS), overall survival (OS) and response with each treatment.

At model entry, the cohort is assigned to the PF state. Between weeks 0 and 12, the proportion of patients that occupy the PD state is modelled based on the weekly cumulative survival

probabilities of OS minus PFS, and the proportions that occupy the death state as one minus OS. At week 12, the proportion that are PF, calculated from the cumulative survival probabilities for PFS, are partitioned into those with complete response (CR) or partial response (PR) and those with SD (PF minus response) based on the response rates with treatment at week 12. This calculation provides the estimated proportions in each of the states at the end of the decision period, as illustrated in Figure 14.

Figure 14: Example calculation of the proportion in the death, PD, stable disease (SD) and response states over the first 12 weeks of treatment



The probabilities of PFS, OS, and response for pembrolizumab were estimated directly from the KEYNOTE-087 study. The corresponding probabilities for SoC are based on the pembrolizumab probabilities, adjusted for the inferior outcomes of SoC. This included adjustment to PFS and response based on a naïve indirect comparison of Cheah[44] versus KEYNOTE-087 detailed in section 4.10. PFS adjustments were based on the hazard ratio (HR) effect size, while response rates were adjusted via odds ratios (ORs). A PFS HR from week 12 to end of follow-up could not be estimated given the low number of events post week 12 observed in Cheah. Therefore, weeks 0 to 12 were not used to estimate the effect of treatment as it would double count patients if the week 0 to end of follow-up HR was applied after the week 0 to 12 HR. Thereby, treatment with pembrolizumab is expected to improve response and reduce the rates of PD when compared to SoC. OS was assumed conservatively to be equal across the treatments given the uncertainty associated with estimating a HR from the immature KEYNOTE-087 data.

The second chance node represents the uptake of alloSCT at week 12 of the analysis, which is modelled on a series of probabilities that vary depending on the underlying response status of the cohort. It is assumed that all alloSCT occur at week 12 in line with; i) the median time

to alloSCT in KEYNOTE-87 (mean of [redacted] weeks based on [redacted] doses [79]); ii) the first response assessment in KEYNOTE-087; iii) the results of the clinician survey (12 weeks median duration of SoC prior to alloSCT). As noted in TA462[75] (nivolumab in RRcHL), alloSCT is offered to relatively fit patients whose disease achieves a partial or complete response to therapy, and as such, those with improved responses rates are expected to have a higher chance of undergoing alloSCT. Therefore, in the submitted model set proportions of patients that achieved either, CR, PR or SD received an alloSCT. The rate of alloSCT of the cohort at week 12 is an important determinant of long-term prognosis given its curative potential in RRcHL and was improved at week 12 versus week 24 in KEYNOTE-087. A UK clinician survey [64] detailed in Section 4.11 and advisory board [40], also suggested that patients would be transplanted as soon as they showed a CR or PR and that in SoC the mean length of time before a transplant would also be 12 weeks.

Long-term Markov state transition model for Week 12 to death.

At the end of the decision tree period, the modelled cohort is split into those who go on to alloSCT and those unable to receive alloSCT based on the response to treatment in the first 12 weeks and alloSCT uptake rates described previously. The long-term survival of the cohort is then modelled through two independent Markov state transition models that predict the long-term outcomes of alloSCT (post-alloSCT pathway) and the outcomes of continued treatment with pembrolizumab or SoC in those unable to undergo alloSCT (non-alloSCT pathway).

The non-alloSCT pathway model consists of three states representing PF, PD, and death. Patients that do not undergo alloSCT at the end of the decision-tree component (e.g. at week 12) are automatically re-distributed to the PF (if CR, PR or SD at week 12) or PD states (if in PD at week 12). Patients who enter the PF state are assumed to continue on their existing therapy until PD, toxicity or death. The health state utility assigned to the PF state is based on a weighted average of the utilities from the CR, PR or SD states at the end of week 12. The same utility is applied to PD patients regardless of prior therapy.

The non-alloSCT pathway model consists of three states representing PF, PD, and death. In the Patients that occupy the PF state are at risk of progression or death, while patients that occupy the PD state are at risk of death. The transition probabilities for PF to PD are derived from parametric survival models fitted to individual patient data for PFS during the post-week 12 period of KEYNOTE-087 for pembrolizumab. The corresponding transition probabilities for standard of care are based on the pembrolizumab models, adjusted for the inferior outcomes of standard of care. This adjustment is performed using a HR estimated from a naïve indirect comparison of PFS for standard of care in Cheah versus pembrolizumab in KEYNOTE-087

detailed in section 4.10. The transition probabilities for PF to death and PD to death are modelled based on external data, and are conservatively assumed at the same rate between treatment groups (section 5.3.1).

The post-alloSCT pathway model consists of two states; alive and dead. Patients that undergo alloSCT at week 12 automatically discontinue their existing treatment, and enter the alive state of the Markov model. During each subsequent cycle, the cohort is at risk of death from any cause. The risk of death after alloSCT was derived from parametric survival models fitted to approximated patient-level data from a follow-up study of alloSCT outcomes in heavily pre-treated cHL patients in the UK [80]. The outcomes of alloSCT are conservatively assumed to be the same regardless of prior therapy, e.g. whether pembrolizumab or SoC was given previously. The health state utility assigned to the post-alloSCT alive state was assumed to vary between the first 100 days and post-100 day periods to account for the impact of the procedure and associated recovery on quality of life. The health state utility assigned to the post day-100 period is based on a weighted average of the KEYNOTE-087 response specific utilities from the CR, PR or SD after alloSCT. The lack of a PD state in this part of the model should not underestimate the cost or QALY implications of alloSCT as the cost applied in the model covers the total costs post-alloSCT up to six years post-follow-up which is expected to include the costs of the procedure, recovery and potentially, further therapy. In addition, the utilities post 100 days are based on response post-alloSCT, which include PD, combined with the utility by response from the KEYNOTE-087. Hence, the results take into account some of the implications of PD post-alloSCT on QALYs. In addition, it is life expectancy post-alloSCT that is a key driver of QALY gain, and not quality of life per se. With this simplified structure, we can better capture life expectancy by avoiding the modelling of transitions between PF, PD and death. Details of utility and costs applied to each health state are detailed in sections 5.4 and 5.5.

Clinical justification for health state structure

The conceptual structure of the economic model is based on an assumed relationship between response status, uptake of alloSCT use and the final clinical benefits of treatment. These relationships are supported by clinical expert testimony given in TA462 [75] (nivolumab in RRcHL) that alloSCT is offered to patients who achieve response to therapy, and that alloSCT is potentially curative in around 60% of patients who receive it. In addition, clinician surveys conducted for TA462 and this appraisal have highlighted that a small proportion of patients that achieve SD may also be eligible for alloSCT. In TA462, it was acknowledged that nivolumab, a PD-1 inhibitor with a similar mechanism of action to pembrolizumab, may act as salvage therapy to enable alloSCT, and hence through higher rates of response (in the first

12 weeks of treatment) increase the overall eligible population for alloSCT and, therefore, its uptake. Similarly, with pembrolizumab, it is expected that higher rates of response will yield an overall increase in the uptake of alloSCT, leading to significant clinical benefits to patients with this condition due to the chance for cure with alloSCT. The model developed for this analysis, therefore, focused on the link between response and subsequent alloSCT in order to quantify these benefits.

As the main goal of alloSCT is to cure the patient of disease, the post-alloSCT pathway model does not include the modelling of PFS post-alloSCT and hence does not consider the potential impact of post-alloSCT PD on outcomes. The omission of the PD state from the model simplifies the calculation of post-alloSCT survival, and hence the cure rates for alloSCT, as they can be derived directly from post-alloSCT survival data without the complications of modelling transitions between intermediary states such as PF and PD. In addition, the role of PFS in determining the QoL of patients who undergo alloSCT is unclear, given that longitudinal studies suggest that the time since alloSCT plays an important role in determining overall QoL, with an early deficit immediately after transplantation that is followed by a return to pre-transplantation levels by day 100, and stabilization or continuation of this improvement from day 100 up to 3 years of follow-up [81]. This trend is captured in the model through the application of different health state utilities (HSU) in the pre and post 100 day periods of alloSCT.

In the no-alloSCT pathway model, it was necessary to include a PD state to link PFS on therapy to OS, given that OS data in KEYNOTE-087 was judged too immature to provide robust extrapolations of survival with pembrolizumab, and because the total duration and hence costs of drug therapy is conditional on PFS.

The health state structure of the model follows that of a number of previous evaluations that have considered “bridging” therapy to SCT;

The recent mock NICE appraisal of CAR T therapy in acute lymphoblastic leukemia which utilised a decision tree for the initial period to identify patients’ remission, minimal residual disease and transplantation status, which subsequently determined entry into any of four state transition models to assess long-term outcomes.

A previous NICE appraisal of bortezomib for induction therapy in multiple myeloma prior to autoSCT (TA311 [82]); used a decision tree to separate patients into one of three health state (CR, PR, non-responder) following induction therapy. Patients were assumed to then receive autoSCT conditional on the post-induction response. Long-term outcomes were conditional

on response to post induction therapy and not the use of autoSCT which is where this model differs from the de novo structure presented here.

5.2.3 Key features of the de novo analysis

A summary of the key features of the economic analysis and of previous appraisals in RRcHL is provided in Table 54.

Table 54: Features of the de novo analysis

Factor	Chosen values	Justification
Time horizon	40 years	A lifetime time horizon was considered in TA462. This was to ensure all important differences in costs and outcomes were reflected (NICE reference case) [76]. Despite >2% of the cohort being alive after 40 years, extending the time horizon further was associated with additional uncertainty. Scenario analysis was conducted with 50 year time horizons to test the sensitivity of this assumption.
Cycle length	1 week	Weekly cycles were the common denominator between the periods between treatment cycles for pembrolizumab and SoC treatments.
Half-cycle correction	Yes	In line with previous submissions and to mitigate bias
Were health effects measured in QALYs; if not, what was used?	Yes	NICE reference case [76]
Discount of 3.5% for utilities and costs	Yes	NICE reference case [76]
Perspective (NHS/PSS)	Yes	NICE reference case[76] Please note that the costs to the NHS were included, but PSS costs have not been considered due to the unavailability of data to incorporate this into the model.
PSS, personal social services; QALYs, quality-adjusted life years		

5.2.4 Intervention technology and comparators

The intervention (i.e. pembrolizumab) was applied in the model as per the anticipated licensed dosing regimen (i.e. administered intravenously at a fixed dose of 200mg over 30 minutes every 3 weeks [Q3W]). The license states that pembrolizumab is to be administered until disease progression or unacceptable toxicity. In this economic analysis, pembrolizumab was given up to a maximum of 24 months as per the stopping rule within KEYNOTE-087 [48].

Pembrolizumab is licensed in adults with RRcHL who have already received autoSCT and BV or are ineligible for autoSCT. The NICE scope specifies the following treatment regimens as relevant comparators [77]:

- Standard of care (made up of chemotherapy and bendamustine)
- Best supportive care

In the specific context of relapsed or refractory HL, with low patient numbers and short survival the clinical pathway for HL patients is subject to considerable uncertainty and heterogeneity, particularly in the post-autoSCT, post-BV setting. Further, data describing treatment in the post autoSCT, post-BV setting is likely to apply investigational therapies rather than established clinical practice.

In light of this uncertainty and the lack of data surrounding comparator composition, the approach has been to use assumptions based on independent sources, such as the published literature, British HL guidelines or previous NICE appraisals in the field of HL or NHL. These assumptions were then assessed for clinical plausibility, and alternative assumptions were assessed in scenario analyses.

In line with this approach, and following an SLR to obtain the most relevant comparator information (further described in section 5.3.1) the base case analysis assumes that SoC is equivalent to the therapies described within the Cheah 2016 [44]real world data. Patients in this study had previously received BV (100%) and autoSCT (71%) and so can be said to adequately represent the post-autoSCT, post-BV HL population. Given the lack of evidence in this area, even more so in autoSCT ineligible post BV patients, Cheah 2016 was assumed applicable in this patient population for the economic analysis also. Cheah et al 2016 was conducted in the USA and so we have attempted to validate the outcomes and comparator regimens included in this study via a UK clinician survey [64, 65].

Results of a survey of 16 UK consultant haem oncologists/ haematologists involved in treating cHL showed that when presented with the outcomes and baseline characteristics of the Cheah

et al 2016 study there was broad agreement with outcomes (CR, PR, ORR, PFS and OS) of between 69-88% when considering this in the context of their own UK practice. Hence Cheah et al 2016 was deemed an appropriate data source to inform the UK SoC arm of the model. Further information about the clinician survey and its results can be found in section 4.11.

Treatments administered within Cheah et al 2016 [44] and the outcomes from these therapies are presented in Table 55. In order to provide the most robust base case analysis, these therapies are assumed to comprise SoC, with the following assumptions and amendments to reflect UK clinical practice and enable calculation of costs and utilities and re-weighted composition of SoC is detailed in the cost and resource use section of this report (Table 88):

- The "Other" category does not provide enough detailed information to allocate costs and utilities, consequently the composition of SoC has been weighted excluding these therapies.
- Second autoSCT is not considered to be a relevant comparator in this patient population, as clinical advisors have explained that patients with RRcHL would rarely receive this [40]. Therefore, composition of SoC has been weighted excluding this therapy.
- BV retreatment after its NICE recommended place in the care pathway is not explicitly recommended by NICE. Therefore, composition of SoC has been weighted excluding this therapy.
- Use of the "Gemcitabine", "Other alkylator" and "platinum based" regimens have been pooled to inform the proportion of patients receiving chemotherapy; composition of chemotherapy in UK clinical practice has been assumed based on equal usage of regimens specified by the BCSH guidelines [3].

Table 55: Cheah 2016 [44]: therapies administered and outcomes

Treatment	*n*	Eval	CR (%)	PR (%)	ORR (%)	mPFS (m)	mOS (m)
Investigational agent	28	28	4(14)	3(11)	7(25)	2.4	47.7
Gemcitabine	15	12	4(27)	4(27)	8(53)	2.1	NR
Bendamustine	12	11	2(17)	4(33)	6(50)	3.7	34.0
Other alkylator	6	4	1(17)	1(17)	2(33)	5.0	9.5
BV retreatment	6	4	0(0)	2(33)	2(33)	3.5	10.4
Platinum based	4	4	0(0)	1(25)	1(25)	0.9	25.2
AutoSCT	3	3	1(33)	1(33)	1(33)	-	11.9
Other	5	1	0(0)	0(0)	0(0)	-	24.9
Total	79	67(85)	12(15)	15(19)	27(34)	3.5	25.2

AutoSCT: autologous stem cell transplant; BV: brentuximab; CR: complete response; mOS: median overall survival; mPFS: median progression free survival; ORR: objective response rate; PR: partial response

Based on BCSH guidelines and clinician opinion, it is believed that use of BSC is minimal at this stage in the treatment pathway, as eligible patients are likely to receive therapy where feasible. As such, BSC has been applied within the model as a subsequent therapy in the base case analysis, with the composition derived from a recent NHL NICE Technology Appraisal (TA306) [83].

In order to provide cost-effectiveness evidence with direct relevance to the NICE scope, scenario analyses have been provided assessing the impact of chemotherapy (as specified within the NICE scope [77]) and BSC as a comparator.

5.2.5 Intervention technology and comparators

In KEYNOTE-087, patients were to continue pembrolizumab until disease progression as determined by the investigator, unacceptable toxicity or a maximum of 24 months of uninterrupted treatment with pembrolizumab [48]. In the cost-effectiveness model, the survival estimates of OS and PFS are based on KEYNOTE-087 data, thus reflecting the within-trial maximum treatment duration.

Based on clinical expert opinion, it was assumed that up to a maximum of 6 cycles were administered to reflect the UK clinical practice for the treatment regimens included under this SoC comparator.

5.3 Clinical parameters and variables

5.3.1 Describe how the clinical data were incorporated into the model

As discussed in section 4, no direct comparative evidence is available for pembrolizumab versus SoC. In order to identify data describing SoC, a SLR was conducted and using this data, a naïve indirect comparison of pembrolizumab versus SoC has been used to inform the base case analysis. A matched adjusted indirect comparison of pembrolizumab versus SoC has also been conducted. As described in section 4, the results of MAIC should be interpreted with caution; this is in the context of the complexities associated with population-adjustments in the context of single arm, and retrospective observational data. Therefore, in an attempt to minimise data loss the naïve indirect comparison was used in the base case and the outcomes of the MAIC analysis investigated in scenario analysis. The primary data source for the SoC arm in the naïve indirect comparison, and therefore in the economic model was Cheah 2016 real world data [44] from which the regimes included were validated by UK clinicians [64]. Patients in this study had previously received BV (100%) and autoSCT (71%) and so can be said to adequately represent the post-autoSCT, post-BV HL population in the base case scenario. As already discussed, an assumption has been made, given the paucity of data in this area, that patients who have been ineligible for an autoSCT (cohort 2 of KEYNOTE 087) would experience the same SoC outcomes. This assumption was validated with UK clinicians that specialise in this area.

Evidence to describe the efficacy of BSC in this population has not been identified; scenario analyses describing BSC as a comparator have been based on the efficacy of SoC, in order to provide a highly conservative analysis of the benefits of pembrolizumab versus BSC. Composition of BSC has been derived from a recent NICE appraisal in the NHL population.

Given the limited availability of literature in the late stage of this disease, the baseline profile of the modelled population was obtained from KEYNOTE-087.

Data from KEYNOTE-087 was used to estimate patient baseline characteristics of interest to the economic analysis. These are age, gender (used for the calculation of general population mortality rates), and body surface area (used for calculating acquisitions costs for surface area dependent therapies). A summary of the values used in the model are presented in Table 56. These values have been varied in deterministic sensitivity analysis (DSA) to assess the impact of the uncertainty of these parameters.

Table 56: Patient characteristics

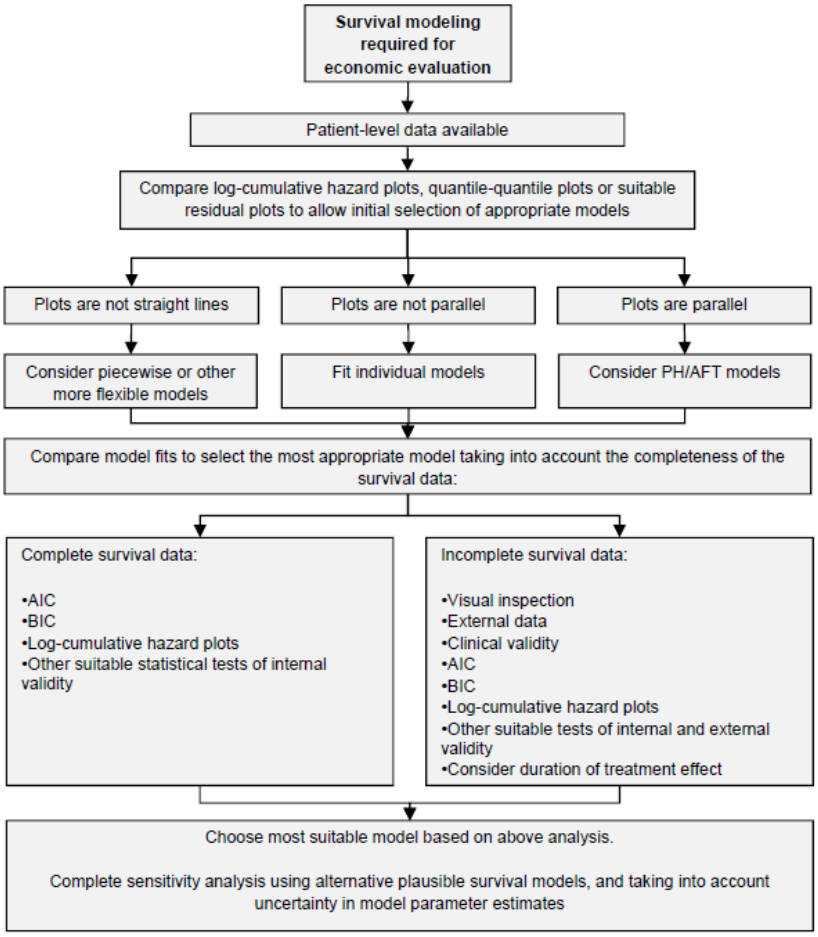
Parameter	Mean	SE
Age	39.9	--
Female (%)	47.3%	--
Body surface area (m²)	1.77	0.024

Parametric survival analysis

Pembrolizumab is the reference treatment for the economic analysis to ensure that the parametric models accurately reflect outcomes for the indicated populations (representative of the cohorts 1 and 2 within the KEYNOTE-087 study).

Survival analyses were conducted using approaches outlined by the decision support unit at NICE. An overview is presented in Figure 15[84].

Figure 15: Process chart for survival modelling for economic evaluations



AFT: Accelerated failure time; AIC: Akaike information criterion; BIC: Bayesian information criterion; PH: Proportional hazards Source:[84]

The first step in the algorithm is the assessment of the proportional hazards assumption judged via the plotting of the log-cumulative hazard function and associated residual plots. However, a large number of progression events occurred during the first 12 weeks of the SoC study [44]. Therefore, it was not possible to estimate a HR between the two treatments after 12 weeks. This was due to the small number of patients at risk and low number of events after 12 weeks which was associated with substantial levels of uncertainty. Thereby, comparative efficacy was only assessed using a constant HR across both periods (pre- and post- 12 weeks). Thus, it was necessary to assume that proportional hazards held and acknowledge this as a limitation of the analysis.

The second step involves the fitting of a series of parametric survival distributions to the patient-level data from KEYNOTE-087 PFS and OS pre week 12 and PFS and time on treatment (ToT) post week 12. In the post alloSCT pathway parametric distributions were fitted to OS data digitized from Lafferty [80]. In all analyses, the following survival distributions were considered:

- Exponential
- Weibull (accelerated failure time)
- Log normal
- Log-logistic
- Gompertz
- Generalized gamma

A summary of the parameterization of the conventional survival distributions is shown in Table 57. All analyses were performed using the FlexSurv package in R.

None of the models considered here included covariates for baseline patient characteristics (e.g. stratified analyses).

Following the NICE recommendations, the “best fitting” model is selected based on internal goodness of fit assessed using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), visual inspection of the fit of the model to the Kaplan-Meier curves and based on an assessment of the clinical plausibility of long-term survival projections.

Of note, the log-logistic, log-normal and generalized gamma are examples of a distinct class of model where the effect of treatment, if included as a covariate in the model fitting, is simulated as an acceleration or deceleration factor on the expected timing of an event. Therefore, when a HR is applied to the baseline distribution (not as a covariate) the generated

comparator curve does not necessarily follow the same distributional form as the reference arm; which is contrary to recommendations that the parametric models should be the same between treatments [84]. This methodological inconsistency did not yield clinically implausible distributions and has not precluded the combining of these distributions with HRs in previous technology appraisals. Hence, we applied HRs to non-proportional hazard models, although one should acknowledged this limitation of the analysis

Table 57: Overview of parametric functions evaluated in the survival analysis, including their mathematical formulation

Distribution name	Survival function or probability density function	Characteristic
Exponential	$s(t) = \exp(-\lambda x)$	Constant hazard function; proportional hazards model
Weibull	$S(t) = \exp\left(-\left(\frac{x}{b}\right)^{a}\right)$	Hazard function can increase or decrease monotonically over time; proportional hazards (or accelerated failure time)
Gompertz	$s(t) = \exp\left(-\frac{b}{a}(\exp(ax) - 1)\right)$	Hazard function can increase or decrease monotonically over time; proportional hazards
Log normal	$s(t) = 1 - \Phi\left(\frac{\log(x) - \mu}{\sigma}\right)$	Hazard function increases initially to a maximum, before decreasing over time
Log-logistic	$s(t) = 1 - \frac{1}{1 + \left(\frac{x}{\alpha}\right)^{-\beta}}$	Hazard function can be non-monotonic with respect to time; accelerated failure time.
Generalized gamma	$f(x) = \lvert Q\rvert (Q^{-2})^{Q^{-2}} \frac{1}{\sigma x \Gamma(Q^{-2})} \exp\left(Q^{-2}(Qw - e^{Qw})\right)$ $x = \exp(\mu + \sigma w)$	Flexible three-parameter model, and can be generalized to the Weibull, exponential and lognormal distributions

Progression-free survival (week 0 to 12) - Pembrolizumab

For the analysis of week 0 to 12, parametric models were fitted using all observed data from study initiation given that only a small number of events occurred in the first 12 weeks. PFS for SoC was estimated by applying the HR described in section (as described in section 4) to the base case pembrolizumab model ([redacted]).

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of PFS at week 12 are shown for each distribution in Table 58.

Table 58: Summary of the goodness of fit qualities of the survival models (cohort 1; PFS prior 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087
AIC	329.4	326.7	328.6	326.2	327.1	328.3	--
Rank	6	2	5	1	3	4	--
BIC	331.6	331.2	333.1	330.7	331.6	335.0	--
Rank	4	2	5	1	3	6	--
% at week 12	89.54%	94.71%	92.75%	95.20%	95.12%	95.13%	[redacted]

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 16.

Figure 16. PFS (BIRC) extrapolations (cohort 1; PFS 12 weeks)

Both the exponential and Gompertz provided poor visual fits to the observed data and predicted substantially lower rates of patients progression-free after 12 weeks compared to the observed data. Of the remaining four distributions there was little difference in statistical fit (AIC/BIC), visual fit and predicted proportion of patients at week 12. Therefore, the log-logistic was applied in the base case model as it had the best statistical fit (lowest AIC/BIC) and predicted the most comparable rate of patients progression-free compared to the observed data at week 12.

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of PFS at 12 weeks are shown for each distribution in Table 59.

Table 59. Summary of the goodness of fit qualities of the survival models (cohort 2; PFS prior 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087
AIC	482.1	474.6	477.3	474.9	471.0	465.0	--
Rank	6	3	5	4	2	1	--
BIC	484.5	479.4	482.1	479.6	475.8	472.2	--
Rank	6	3	5	4	2	1	--
% at week 12	82.76%	90.96%	88.37%	91.42%	92.33%	92.79%	[redacted]

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 17.

Figure 17: PFS (BIRC) extrapolations (cohort 2; prior 12 weeks)

The exponential substantially underestimated the proportion of patients progress-free at week 12 compared to the observed data and had a poor visual fit so was excluded from consideration. The generalised gamma provided a significantly better fit both statistically and visually compared to all other distributions, therefore it was applied in the base case model. However, it overestimated the proportion progression-free at week 12 compared to the observed data; therefore, the Weibull was considered during scenario analysis as it predicted a lower proportion of patients progression-free at week 12 and had the third best statistical fit (AIC/BIC).

Overall survival (week 0 to 12) – Pembrolizumab

The analysis of OS from week 0 to 12, involved fitting parametric models to all the observed data from study initiation given the low numbers of events. This allowed some events to be predicted during the first 12 weeks. Given the immaturity of the KEYNOTE-087 OS data it was assumed conservatively that SoC was equivalent to pembrolizumab as any HR estimated from an indirect comparison would have been subject to significant uncertainty.

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of OS at 12 weeks are shown for each distribution in Table 60.

Table 60: Summary of the goodness of fit qualities of the survival models (cohort 1; OS prior 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087
AIC	51.9	53.9	53.9	53.9	53.9	55.8	--
Rank	1	2	4	3	5	6	--
BIC	54.1	58.3	58.4	58.3	58.4	62.5	--
Rank	1	2	4	3	5	6	--
% at week 12	99.24%	99.07%	99.25%	99.07%	99.00%	99.07%	[redacted]

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 18.

Figure 18: OS cohort 1 week 0 to 12 extrapolations

There is no meaningful difference in the statistical fit (AIC/BIC), visual fit or predicted patients alive at week 12. Therefore, as only one event occurred in the first 12 weeks in cohort 1, the log-normal was chosen as it predicted the highest rate of mortality at week 12.

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of OS at 12 weeks are shown for each distribution in Table 61.

Table 61: Summary of the goodness of fit qualities of the survival models (cohort 2; OS prior 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	KEYNOTE-087
AIC	80.8	80.1	81.2	80.0	79.4	--
Rank	4	3	5	2	1	--
BIC	83.2	84.9	86.0	84.8	84.2	--
Rank	4	3	5	2	1	--
% at week 12	98.83%	99.85%	99.52%	99.86%	99.95%	[redacted]

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 19.

Figure 19: OS cohort 2 week 0 to 12 extrapolations

There is no meaningful difference in the statistical fit (AIC/BIC) or visual fit or predicted patients alive at week 12. Therefore, as no death occurred in the cohort 2 during the first 12 weeks, the exponential was chosen as it predicted the highest rate of mortality at week 12.

Response rates

Response rates were applied at week 12 in the model to apportion patients that were progression-free into CR, PR or SD. The proportions of patients with either CR/PR were estimated directly from observed data (presented in this section) with the remaining progression-free patients that had not achieved responses were assumed to occupy the SD node.

Pembrolizumab

The cohort specific response rates from KEYNOTE-087 [46] are presented in Table 62.

Table 62: KEYNOTE-087 number of complete and partial responders

Response	Cohort 1	Cohort 2
	n	N
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]

Standard of Care

The comparative response of SoC was estimated via ORs estimated from a naïve indirect comparison. The associated ORs estimated for cohort 1 and 2 are presented below in Table 63.

Table 63: Odds ratios for response

Response	Cohort 1		Cohort 2	
	Mean	SE	Mean	SE
CR	[redacted]	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]	[redacted]

AlloSCT rates conditional on response

KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2 ([redacted]). However the rates of alloSCT in UK patients from KEYNOTE-087 were greater than this, probably largely due to alloSCT being seen as clinical practice within the UK compared to other countries. There were [redacted] patients in cohort 1 and 2 being transplanted with allogeneic stem cells respectively. It should be noted that alloSCT events were not censored from the survival analysis of KEYNOTE-087, this was due to the fact that it would not be possible to censor the SoC arm data. This is accepted as a limitation however this should be limited given that some patients

received alloSCT in the Cheah et al 2016 publication. With limited UK patient numbers it is difficult to draw robust conclusions from the KEYNOTE-087 rates of alloSCT and hence the proportions of patients that received alloSCT conditional on response (CR/PR/SD) were estimated from clinician surveys. MSD conducted a survey [64] of sixteen clinicians from the UK who were asked the proportion of patients they would expect to proceed to alloSCT conditional on response to treatment. The results of the MSD survey were combined with the results of an alternative clinician survey completed by Bristol-Myers Squibb and presented marked as academic in confidence in the recent cHL submission (TA462) [85] from which mean rate was calculated and applied in the base case model. The estimated rates of alloSCT are presented in Table 64. It should be noted that the MSD clinician survey did return some responses which suggested alloSCT in the PD state. Following further discussion with UK clinicians on this topic, alloSCT has not been applied in PD as this is not thought to be standard UK clinical practice in this area.

Table 64: AlloSCT rates conditional on response

	MSD Mean[64]	Alternative Mean	Overall Mean	SE
CR	56.79%	[redacted]	[redacted]	[redacted]
PR	43.93%	[redacted]	[redacted]	[redacted]
SD	18.36%	[redacted]	[redacted]	[redacted]

The rates of alloSCT were judged to be appropriate given that they were:

- Higher than values from a French study[86] (22.2% CR; 14.1% PR; 5.56% SD) which were considered too low for UK clinical practice by the previous committee (TA462)[75]
- Lower than rates reported in Cheah [44] (66% responder received alloSCT) which was deemed too high for UK clinical practice by the previous committee (TA462).
- Broadly in line with the KEYNOTE-087 UK patient alloSCT rates.
- The rates used in this submission were validated with a UK clinician specialising in this area who also suggested that they would even expect these rates to be higher in clinical practice with PR rates also being as high as a CR. However in order to show a

conservative analysis, the base case has utilised the mean calculated rates shown above.

- Comments in the final appraisal document for TA462 [57] also suggest that the true proportion of alloSCT is likely to be somewhere around the values reported here based on the information above.

The values for cohort 2 from the MSD clinician survey could not be validated as no external data was available and the previous cHL appraisal did not consider this population. Based on discussions with UK clinicians, it was suggested that due to the unmet need in this population, they would be likely to attempt an alloSCT if the patient showed a sufficient response and that the real alloSCT rates in cohort 2 would likely even be higher than that in cohort 1. After discussion with UK clinicians, and in order to show a conservative analysis, the same alloSCT rate values have been used in the base case across both populations.

Progression-free survival (post 12 weeks) – Pembrolizumab

The analysis of PFS from week 12, involved fitting parametric models to the observed data in KEYNOTE-087 from week 12. Given that over half the PFS events in Cheah occurred within the first 12 weeks of the study this created significant uncertainty within the SoC arm, therefore, a HR using post-12 week data from KEYNOTE-087 and Cheah was not fitted. It was assumed that the treatment effect was constant across the trial period (the HR from start of treatment (section 4.10) was applied to the post-12 week data from KEYNOTE-087 described in this section ([redacted]).

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of PFS over time post 12 weeks are shown for each distribution in Table 65.

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 20.

Table 65: Summary of the goodness of fit qualities of the survival models (cohort 1; PFS post 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort
AIC	285.37	287.31	287.22	287.90	291.01	289.17	--
Rank	1	3	2	4	6	5	--
BIC	287.50	291.56	291.47	292.15	295.26	295.55	--
Rank	1	3	2	4	5	6	--
Median (months)	13.34	13.34	13.11	13.80	15.41	13.11	[redacted]
Mean (months)	19.68	18.77	16.30	39.22	50.23	16.34	[redacted]
% at 1 year	54.79%	54.53%	54.46%	55.10%	56.57%	54.49%	[redacted]
% at 2 years	29.84%	28.42%	24.86%	34.77%	39.98%	24.66%	--
% at 5 years	8.85%	7.40%	2.53%	18.84%	25.09%	2.79%	--
% at 10 years	0.23%	0.11%	0.00%	7.18%	11.05%	0.00%	--

Figure 20: PFS (BIRC) cohort 1 from week 12 extrapolations

According to AIC/BIC the best performing model was the exponential, followed closely by the gompertz, Weibull and log-logisitc. The worst performing models were the log-normal and generalised gamma. Given the small relative difference in the values, these models are not particularly inferior.

The log-normal and log-logistic models are characterized by a hazard function that initially increases to a maximum before decreasing over time, leading to a gradual shallowing of the predicted PFS curve. This is demonstrated with the high proportion of patients predicted as progression-free after 10 years (11% and 7%, respectively). These models were therefore judged to provide implausible fits given the assumption that proportional hazards held between KEYNOTE-087 and Cheah [44], where a relatively constant decline in PFS was observed with all patients progressing within 18 months.

Of the remaining distributions, the generalised gamma and gompertz provided medians and 1-year PFS rates that differed even further from the observed data. The exponential was applied in the base case over the Weibull as it had a marginally superior statistical fit and a slightly higher 1-year PFS rates.

In conclusion, the preferred distribution for the base case was the exponential, for the following reasons:

1. Best goodness-of-fit values (AIC/BIC), when compared with other distributions
2. Provided the closest median and 1-year PFS rate to the observed data of the plausible distributions
3. Followed a hazard rate over time consistent with that observed within Cheah

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of PFS over time after week 12 are shown for each distribution in Table 66.

Table 66: Summary of the goodness of fit qualities of the survival models (cohort 2; PFS post 12 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort 2)
AIC	352.3	353.8	350.8	360.9	368.0	347.1	--
Rank	3	4	2	5	6	1	--
BIC	354.4	358.0	355.1	365.2	372.2	353.5	--
Rank	2	4	3	5	6	1	--
Median (months)	8.51	8.51	8.97	9.43	9.20	8.05	[redacted]
Mean (months)	12.60	13.76	10.00	38.19	50.58	8.85	[redacted]
% at 1 year	39.07%	40.07%	36.72%	45.06%	45.67%	33.97%	[redacted]
% at 2 years	15.12%	18.00%	2.57%	29.01%	33.09%	0.00%	--
% at 5 years	2.27%	4.04%	0.00%	16.95%	22.21%	0.00%	--
% at 10 years	0.01%	0.06%	0.00%	7.54%	11.54%	0.00%	--

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 21.

Figure 21: PFS (BIRC) cohort 2 from week 12 extrapolations

According to AIC/BIC, the best performing model was generalised gamma, followed closely by gompertz, exponential and Weibull. Given the relative difference in the AIC/BIC values, this indicated that the log-normal and log-logistic models provided an inferior fit to the data when compared with the other models.

The final drops in the Kaplan-Meier curve observed from month 11 were associated with substantial uncertainty given the low number of patients at risk (n=3) and was not considered particularly informative in the process of selecting the most plausible parametric model.

All of the remaining distributions underestimated the median and 1-year PFS rate observed in KEYNOTE-087, particular the generalised gamma, therefore, despite the superior visual fit to the tail of the Kaplan-Meier data it was not considered plausible for the base case analysis.

In the base case analysis, the exponential was applied with the gompertz considered in scenario analysis given the uncertainty in the tail of the Kaplan-Meier curve.

In conclusion, the preferred distribution for the base case was the exponential, for the following reasons:

1. Third/second best goodness-of-fit values (AIC/BIC)

2. Plausible visual fit to the Kaplan-Meier data
3. Followed a hazard rate over time consistent with that observed within Cheah
4. More clinically plausible tail given the uncertainty from 11 months in the Kaplan-Meier curve

Mortality pre-progression

Due to the limited OS data available from KEYNOTE-087, the rate of death in patients who have yet to progress in the non-alloSCT pathway was assumed equal to that of the general population mortality, obtained from UK life tables [87]. The application of a zero rate of death pre-progression would systematically underestimate long-term mortality trends in patients with RRcHL as they would at the very least, be at the same risk of all-cause mortality as an age and gender matched sample of the general population.

The general population life tables for the UK [87] reported the annualised probability of death by age and gender. These mortality probabilities were converted to weekly rates using standard conversion methods. The mortality rates applied in the model were adjusted in line with the proportion of males and females enrolled to KEYNOTE-087. An excerpt of the life tables used in the model is presented in Table 67.

Table 67: Excerpt from UK life tables annual mortality rates

Age (years)	Mortality rate between age x and (x +1): Male	Mortality rate between age x and (x +1): Female
30	0.000676	0.000357
31	0.000712	0.000384
32	0.000808	0.000424
33	0.000806	0.000452
34	0.000883	0.000508
35	0.00096	0.000552

Post-progression survival

At the time of analysis, the number of patients that progressed in KEYNOTE-087 was judged to be too small to support robust analysis of post-progression survival. In the absence of data, external literature sources that reported mortality in patients with RRcHL who had progressed on treatment were identified from the clinical systematic review.

The most appropriate study identified in the included study list was the retrospective study used for the SoC comparison [44]. This study reported median overall survival of 25.2 months across all treatments administered after progression on BV. It should be noted that the OS in SoC is subject to uncertainty as discussed in the recent TA462. The committee reached the conclusion that although Cheah et al 2016 was a US study, it was the most relevant source of evidence for SoC but also that there was a case for end of life criteria to be met suggesting that the actual OS for SoC in the UK is expected to be lower.

Using standard techniques, the median OS reported was converted to a weekly mortality rate assuming a constant hazard rate based on an exponential distribution. The following equations were used:

$$E[median] = \frac{\ln(2)}{\lambda}$$

$$\lambda_{month} = \frac{\ln(2)}{25.2}$$

$$\lambda_{week} = \lambda_{month} / (4.348)$$

The monthly rate of death was converted to a weekly rate assuming there are 365.25 days per year, 30.4375 days per month (365.25/12), and 4.348214 7-day weeks per month (30.4375/7). The weekly rate was converted to a weekly probability (0.63%). The model calculations use the probability of survival to estimate state transitions. This is calculated as one minus the probability of death, e.g. (1-0.63%).

Post-progression survival (PPS) was assumed constant due to both a lack of data to model a time dependent PPS and for simplicity as it removed the need of tracking patients within the Markov model. However, despite this limitation the predicted OS had a good level of face validity when compared to the observed OS from the Cheah study. After the maximum follow-up of Cheah (72 months) approximately 15% of patients were alive, which was correlated to the predicted SoC OS in the model at 72 months of approximately 15% in both cohorts 1 and 2.

The assumption that there was no post-progression survival benefit was conservative given the current OS rates of [redacted]% from KEYNOTE-087 cohort 1 and 2, respectively versus ~56% in Cheah at approximately 20 months. Despite some potential slight imbalances in the populations this extensive difference in the observed survival cannot be dismissed.

Overall survival post-AlloSCT

In line with the previous NICE submission (TA462 [75]) the OS of patients that received alloSCT was taken from a UK study of 13 patients with cHL who received an alloSCT after 3 previous therapies [80]. It was acknowledged in the submission that there was significant uncertainty in the long-term extrapolations given the small sample size and that the median follow-up was only about 28 months. However, the study was also deemed the best available evidence for alloSCT in the UK setting.

It was not possible to reproduce the patient level data from the digitised Lafferty[80] Kaplan Meier provided in the previous submission document (TA462) (Appendix 17), fully or with reasonable approximation, using the Guyot [60] algorithm. This was most likely due to the limited number of events during the follow-up (i.e. five events) and the unknown rate of censoring in the tail of the curve. Therefore, the pseudo-patient level data resulting from the Guyot algorithm was manually adjusted to provide a more accurate representation of the data reported by Lafferty. Following the fitting of parametric survival models to the manually-adjusted data in R, the estimated parameters were similar to those reported in TA462 [75] However, given the Guyot output required manual adjustment it was considered better practice

to use the point estimates and AIC/BIC values from TA462. Nevertheless, in the absence of reported measures of uncertainty around the estimates such as variance-covariance or Cholesky decomposition matrices, the Cholesky decompositions based on the parametric models fitted to the manually adjusted pseudo-patient level data were applied to account for the uncertainty associated to the parameter estimates in the PSA.

The generalised gamma predicted an infinite hazard after approximately 150 months, which lead to both Excel and R not being able to calculate the survival function associated to the generalised gamma model beyond this time. Therefore, to be included within the model, the average increase in cumulative hazard over the 6 months prior to the error was used to project future survival, effectively assuming a constant event hazard from 150 months onwards. This adjustment is expected to have produced an underestimation of the mean survival time, as the associated hazard function was decreasing with time. The generalised gamma could not be incorporated in PSA because of this necessary adjustment; nevertheless, this was not considered an issue as the generalised gamma model was not applied in the base case analysis.

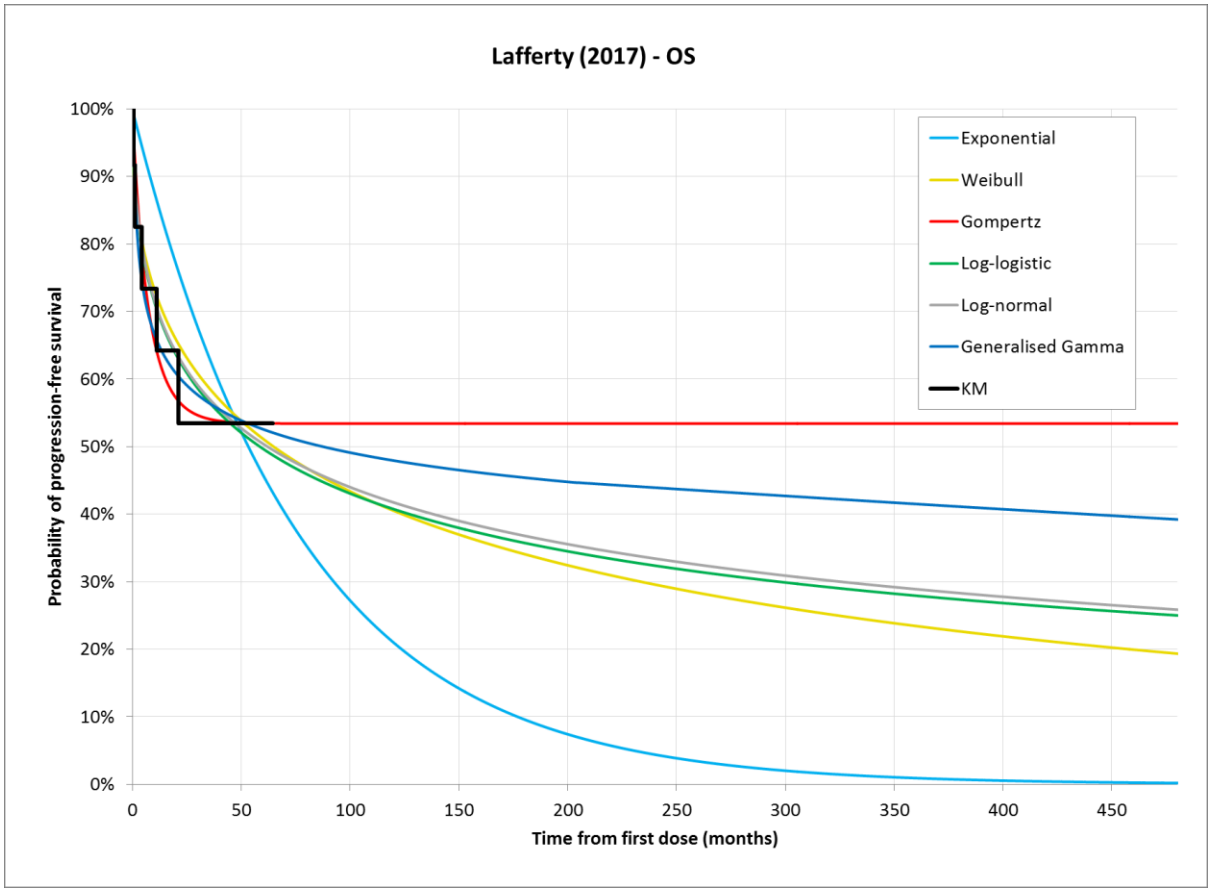
A summary of the goodness of fit statistics and modelled probabilities of OS after alloSCT are shown for each distribution in Table 68.

Table 68: Summary of the goodness of fit qualities of the survival models (OS after alloSCT)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
AIC	55.22	51.12	49.04	50.6	50.1	48.63	--
Rank	6	5	2	4	3	1	--
BIC	55.62	51.91	49.84	51.39	50.9	49.83	--
Rank	6	5	2	4	3	1	--
Median (months)	53.13	64.62	483.42	58.41	61.86	87.39	--
Mean (months)	76.77	163.60	260.05	174.65	179.05	220.77	--
% at 1 year	85.73%	71.68%	63.33%	69.74%	70.01%	65.28%	64.17%
% at 2 years	73.39%	63.78%	55.90%	61.55%	61.93%	59.48%	53.47%
% at 5 years	53.77%	54.50%	53.58%	52.68%	53.33%	54.21%	53.47%
% at 10 years	21.09%	40.56%	53.40%	40.79%	41.77%	47.96%	--
% at 15 years	9.67%	34.13%	53.40%	35.78%	36.83%	45.43%	--

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
% at 20 years	4.43%	29.61%	53.40%	32.40%	33.45%	43.94%	--
% at 30 years	0.93%	23.46%	53.40%	27.94%	28.91%	41.52%	--
% at 40 years	0.20%	19.37%	53.40%	25.02%	25.88%	39.23%	--

Figure 22: OS after alloSCT extrapolations

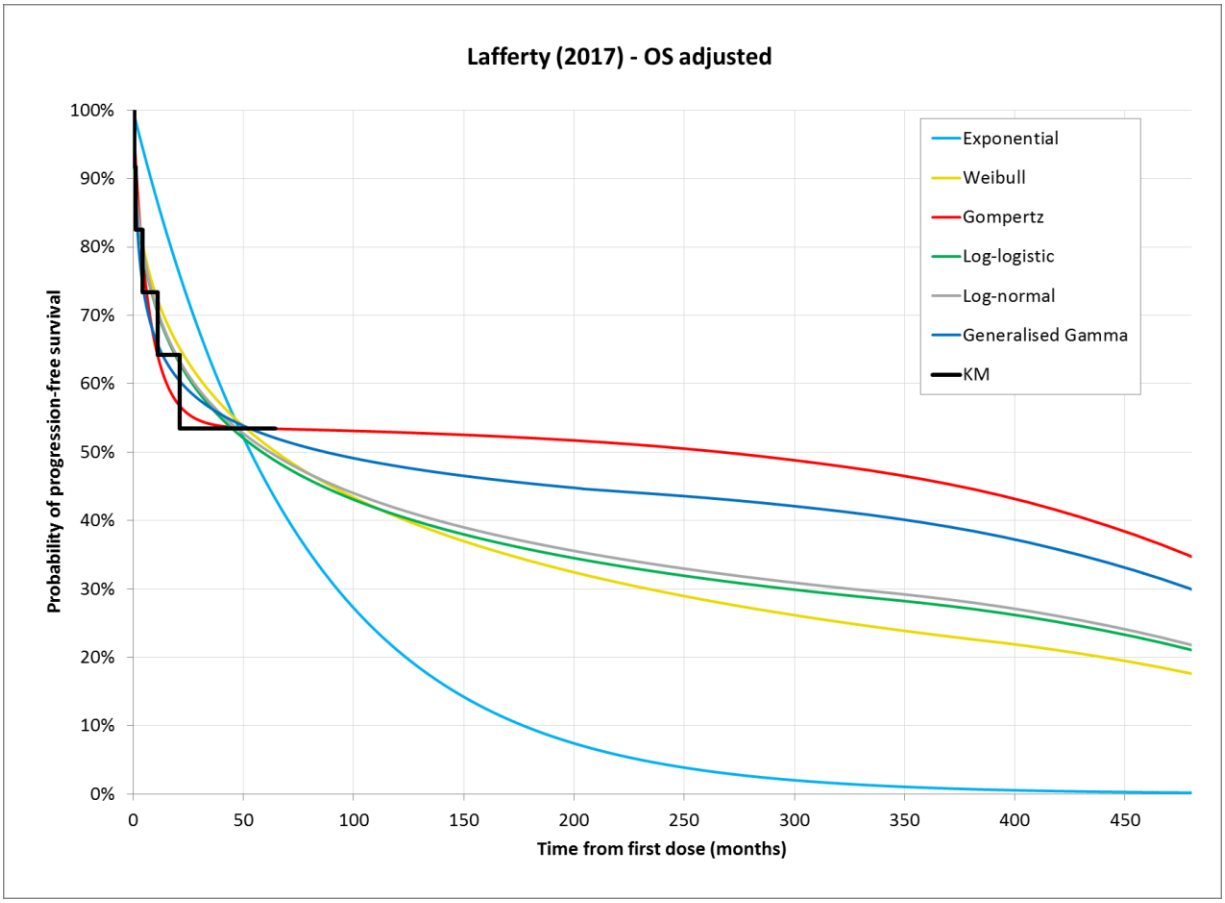


The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 22. The parametric survival analysis confirmed declining hazard over time, leading to plateaus in most parametric models, as reported in Table 68. Therefore, it was appropriate to apply a control to ensure that the mortality rate in any model cycle was greater than or equal to that of the age and gender matched general population all-cause mortality. The modelled probabilities and the fit of the parametric models to the Kaplan-Meier data following this adjustment are presented in Table 69 and Figure 23, respectively.

Table 69: Summary of the goodness of fit qualities of the survival models (OS after alloSCT adjusted for all-cause mortality)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
Median (months)	53.13	64.62	266.78	58.41	61.86	87.39	--
Mean (months)	76.77	163.07	237.71	172.88	177.21	213.93	--
% at 1 year	85.73%	71.68 %	63.33%	69.74 %	70.01 %	65.28%	64.17%
% at 2 years	73.39%	63.78 %	55.90%	61.55 %	61.93 %	59.48%	53.47%
% at 5 years	53.77%	54.50 %	53.58%	52.68 %	53.33 %	54.21%	53.47%
% at 10 years	21.09%	40.56 %	52.90%	40.79 %	41.77 %	47.95%	--
% at 15 years	9.67%	34.13 %	52.08%	35.78 %	36.83 %	45.43%	--
% at 20 years	4.43%	29.61 %	50.80%	32.40 %	33.45 %	43.82%	--
% at 30 years	0.93%	23.46 %	45.95%	27.88 %	28.84 %	39.63%	--
% at 40 years	0.20%	17.64 %	34.77%	21.10 %	21.83 %	29.99%	--

Figure 23: OS after alloSCT adjusted for all-cause mortality extrapolations



According to AIC/BIC measures of goodness of fit (Table 68), the best performing model was the generalised gamma, followed by gompertz, log-normal, log-logistic and Weibull. Given the relative difference in the AIC/BIC values, this indicated that the exponential model provided an inferior fit to the data when compared with the other models, indicating that the constant hazard property of the exponential model was not compatible with the curative nature of alloSCT.

In the previous submission (TA462) the ERG considered that the log-normal and Weibull were more clinically plausible as they did not predict infinite survival (when the all-cause mortality constraint was not applied). The log-normal had a marginally better statistical fit (AIC/BIC) and visual fit compared to the Weibull. However, the conservative option was taken to use the Weibull (lowest mean survival and percentage alive after 40 years) in the base case model. The log-normal was used in a scenario analysis.

Time on treatment post-12 weeks

Of primary importance to the economic model is the rate of treatment discontinuation, which is a key driver of costs and incremental cost-effectiveness. For comparative purposes, the time on treatment (ToT) curve is presented alongside the Kaplan-Meier for PFS from KEYNOTE-087 cohort 1 and cohort 2 in Figure 24 and Figure 25, respectively.

Figure 24: Kaplan-Meier Analysis of ToT and PFS from KEYNOTE-087 cohort 1

Figure 25: Kaplan-Meier Analysis of ToT and PFS from KEYNOTE-087 cohort 2

Figure 24 demonstrates that PFS is a reasonable proxy for treatment duration in cohort 1 with some overlap in the curves, however on average PFS slightly exceeded ToT. In contrast, Figure 25 illustrates that there are differences in time to PFS and ToT for pembrolizumab, with the probability of PFS generally exceeded that of ToT. Thus, on average, patient's discontinued pembrolizumab prior to progression and PFS is not a suitable proxy of ToT in cohort 2. Use of PFS to simulate treatment exposure would lead to an overestimate of pembrolizumab costs in the economic analysis.

The trend towards the discontinuation of pembrolizumab prior to progression may be due to a number of factors, such as tolerability and safety and the impact of the design of KEYNOTE-087, which allowed study investigators to discontinue therapy if complete response had been achieved after at least 6 months of treatment.

The duration of pembrolizumab treatment is modelled via the simulation of ToT data from week 12 onwards in KEYNOTE-087 and extrapolated to a maximum time period of 24 months. The analysis of ToT was conducted as discussed earlier in this section and the details of the "best fitting" model for cohort 1 and cohort 2 is described below. As the model only requires ToT up to month 24, after which point no further pembrolizumab therapy is provided as per the study protocol, the main selection criteria for best fitting model was based on internal goodness of fit and the estimated proportion on treatment after two years.

It was assumed that PFS was a reasonable proxy for ToT for SoC as no treatment discontinuation data was available for SoC..

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of ToT over time are shown for each distribution in Table 70.

Table 70: Summary of the goodness of fit qualities of the survival models

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort 1)
AIC	365.2	367.0	367.0	367.4	366.7	368.6	--
Rank	1	3	4	5	2	6	--
BIC	367.4	371.3	371.3	371.7	371.0	375.0	--
Rank	1	3	4	5	2	6	--
Median (months)	12.19	11.96	12.19	11.96	11.96	11.96	[redacted]
Mean* (months)	13.25	13.19	13.08	13.49	13.50	13.41	[redacted]
% at 1 year	51.49%	51.34%	51.57%	51.21%	51.05%	51.00%	[redacted]
% at 2 years	26.17%	23.94%	21.55%	30.14%	31.49%	29.41%	--
*2 year restricted mean							

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 26.

Figure 26: ToT cohort 1 from week 12 extrapolations

All parametric models provided similar medians, restricted means and statistical and visual fits to the observed data. Therefore, the exponential was applied in the base case model as it provided the best statistical fit and maintained consistency with base case PFS distribution.

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of ToT over time are shown for each distribution in Table 71.

Table 71: Summary of the goodness of fit qualities of the survival models

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort 1)
AIC	523.93	525.4575	525.3865	526.44	532.4001	527.4514	--
Rank	1	3	2	4	6	5	--
BIC	526.2475	530.0925	530.0215	531.075	537.0351	534.4039	--
Rank	1	3	2	4	6	5	--
Median (months)	6.67	6.44	6.21	5.98	5.75	6.44	[redacted]
Mean* (months)	9.02	9.08	9.19	9.52	9.71	9.09	[redacted]
% at 1 year	29.95%	30.61%	30.52%	31.73%	33.54%	30.61%	[redacted]
% at 2 years	8.76%	10.37%	12.56%	16.77%	19.30%	10.60%	--
*2 year restricted mean							

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 27.

Figure 27: ToT cohort 2 from week 12 extrapolations

According to AIC/BIC all parametric models apart from the log-normal had similar statistical fits to the observed data. Given the relative difference in the AIC/BIC values, this indicated that the log-normal model provided an inferior fit to the data when compared with the other models.

Given the similar medians and means predicted by the remaining five distributions the exponential was chosen to maintain consistency with the base case PFS distribution.

5.3.2 Demonstrate how the transition probabilities were calculated from the clinical data

As discussed in the sections above, at the end of the decision tree period, the modelled cohort is split into those who go on to alloSCT and those unable to receive alloSCT based on the response to treatment in the first 12 weeks and alloSCT uptake rates described previously. The long-term survival of the cohort is then modelled through two independent Markov state transition models that predict the long-term outcomes of alloSCT (post-alloSCT pathway) and the outcomes of continued treatment with pembrolizumab or SoC in those unable to undergo alloSCT (non-alloSCT pathway). Transition probabilities were derived from parametric survival models, please refer to the section above titled 'long term markov state transition model for Week 12 to death' for details of this.

5.3.3 If there is evidence that transition probabilities may change over time for the treatment effect, confirm whether this has been included in the evaluation

As described in section 5.3.1, a large number of progression events occurred during the first 12 weeks of the SoC study used in this economic analysis. Therefore, it was not possible to estimate a HR between the two treatments after 12 weeks due to the small number of patients at risk and low number of events after 12 weeks which was associated with substantial levels of uncertainty. Thereby, comparative efficacy was assessed using a constant HR across both periods (pre- and post- 12 weeks). Thus, it was necessary to assume that proportional hazards held and this is acknowledged as a limitation of the analysis.

5.3.4 Inputs from clinical experts

Throughout the submission, MSD has sought to provide the most robust and clinically relevant economic analysis in this patient population. As mentioned, clinical input to this submission has come from a UK clinical advisory board, a UK clinician survey and in addition 1:1 discussion with an expert in this area.

As discussed earlier, the key assumptions used in the economic model regarding alloSCT rates were validated with UK clinicians focussed in this area. These assumptions were considered to be, if anything, a conservative picture of UK clinical practice.

The model structure itself was also suggested by an UK clinical expert to be a strong representation of the UK clinical pathway for RRcHL patients, particularly in an area where there is a paucity of data surrounding current clinical practice.

In addition, the base case curve fitting for the following long-term extrapolations were validated with UK clinical experts and deemed to be clinically relevant:

- PFS from week 12 no alloSCT
- OS following alloSCT
- ToT from week 12 compared to PFS
- ToT from week 12
- OS following no alloSCT

5.3.5 Adverse Events

Grade 3+ AEs were applied in the model as a one-off cost and disutility in the first model cycle. As serious AEs can potentially lead to treatment discontinuation, patients remaining on treatment beyond the first year are assumed to be tolerating treatment well and therefore not experiencing severe AEs. To best reflect clinical practice, the included AEs were taken from a previous Hodgkin's Lymphoma appraisal (TA462) [75]. The list was subsequently validated during a clinician survey[64, 65] and was deemed to capture all relevant AEs. Additionally, given the positive safety profile of pembrolizumab in KEYNTOE-087, no additional AEs were identified for inclusion in the model from KEYNOTE-087 as all other grade 3+ AEs occurred in ≤2 patients.

In the base case analysis, SoC was assumed to consist of chemotherapy, bendamustine and investigational agents. The mix of chemotherapy was assumed equivalent to that used within TA462 in line with BCSH guidelines [3] with mini-BEAM and DexaBEAM excluded following previous committee comments [75]. Table 72 present the AEs incidence rates and sources for chemotherapy regimens included in the model.

The majority of studies included in Table 72 report treatment related AEs however, for some studies it was unclear. Therefore, the conservative assumption was made to include all-cause AEs from KEYNOTE-087 in the model (Table 73).

Investigational agents were assumed to have no AEs. The incidences of AEs for bendamustine are presented in Table 74 along with the SoC weighted average included in the model.

Table 72: Chemotherapy adverse events incidence (number of events)

	ICE	IVE	MINE	IVOx	IGEV	GEM-P	GDP	GVD	ESHAP	ASHAP	DHAP	DHAOx
Sample size	-	62 (1451)	802	-	91 (3131)	21	23	37	22	-	102 (2011)	70
Anaemia	NR	NR	20	NR	17	2	2	6	6	NR	17	4
Diarrhoea	NR	NR	2	NR	NR	0	NR	1	7	NR	NR	NR
Dyspnea	NR	NR	NR	NR	NR	NR	1	4	NR	NR	NR	NR
Fatigue	NR	NR	NR	NR	NR	NR	2	4	NR	NR	NR	NR
Leukopenia	NR	NR	NR	NR	NR	13	NR	6	NR	NR	69	NR
Nausea	NR	NR	6	NR	1	0	0	0	NR	NR	13	0
Neutropenia	NR	15	74	NR	26	15	2	19	7	NR	NR	25
Pyrexia	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Thrombocytopenia	NR	11	30	NR	18	10	3	16	7	NR	71	26
Vomiting	NR	NR	NR	NR	NR	0	3	1	NR	NR	NR	0
Source	88	89	90	91	92	93	94	95	96	97	98	99

1 adverse event incidence were originally reported over this number of treatment course/ cycles. Therefore, some patients may have been double counted if they experienced the adverse events over multiple treatment cycles. The rates were subsequent reweighted using the patient sample size to apply the correct study weighting in the overall chemotherapy safety profile, however the potential double counting issue remained.

2 it was unclear if haematological adverse events had been reported over 207 cycles or 80 patients

Table 73: KEYNOTE-087 adverse events (all-cause grade 3+)

Adverse event	Cohort 1&2 (n=150)		Cohort 1 (n=69)		Cohort 2 (n=89)	
	Number of events	%	Number of events	%	Number of events	%
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dysponea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Leukopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pyrexia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 74: SoC adverse events incidence

Adverse event	Chemotherapy (38.46%)		Bendamustine (18.46%)		SoC	
	n	N	n	N	N	N
Anaemia	163	894	5	36	29	178
Diarrhoea	10	160	0	36	4	68
Dysponea	5	60	0	36	2	30
Fatigue	6	60	1	36	3	30
Leukopenia	155	259	NR	NR	0	0
Nausea	37	745	1	36	8	170
Neutropenia	383	838	3	36	71	163
Pyrexia	0	0	1	36	0	7
Thrombocytopenia	366	1039	7	36	75	202
Vomiting	4	151	0	36	2	65
Source	Table 72		[100]		Weighted average of chemotherapy, bendamustine and investigational agents	

5.4 Measurement and valuation of health effects

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

HRQoL was evaluated in KEYNOTE-087 using two QoL measures; the EORTC-QLQ-C30 and the EQ-5D-3L as a measure of generic QoL. The EQ-5D-3L was collected:

- At treatment cycles 1, 2, 3, 4, 5 (i.e. every 3 weeks) and every 12 weeks thereafter until progression whilst the subject was receiving study treatment
- On treatment discontinuation
- 30 Days post treatment discontinuation

HSU values were generated by mapping the domain scores of the EQ-5D-3L to a single index value using the UK social tariff [101] consistent with the NICE reference case [76]. A post hoc analysis allowed utilities to be calculated using observations from week 12 only [102] (to maintain consistency with the model structure), stratified by response (Table 75).

Table 75: KEYNOTE-087 EQ-5D-3L health utility values by overall response rate at week 12

Response status	Mean	Standard error	N
CR	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]

5.4.2 Mapping

Utilities were evaluated using EQ-5D directly from patients from the KEYNOTE-087 trial, which is consistent with NICE reference case. Therefore, no mapping was conducted.

5.4.3 Health-related quality-of-life studies

The relevant HRQoL data from the published literature were identified through a systematic literature search carried out 12th July 2017 from 2001, for patients with RRcHL regardless of previous therapy. The objective was to identify HRQoL (in terms of utilities) associated with RRcHL, in line with the research question posed in section 5.1.

A comprehensive literature search was carried out using the databases presented in Section 5.1.1. Conference searches were also performed to identify potentially relevant conference abstracts or posters of interest (see Section 5.1.1). These searches were restricted to abstracts published during the last 2 years.

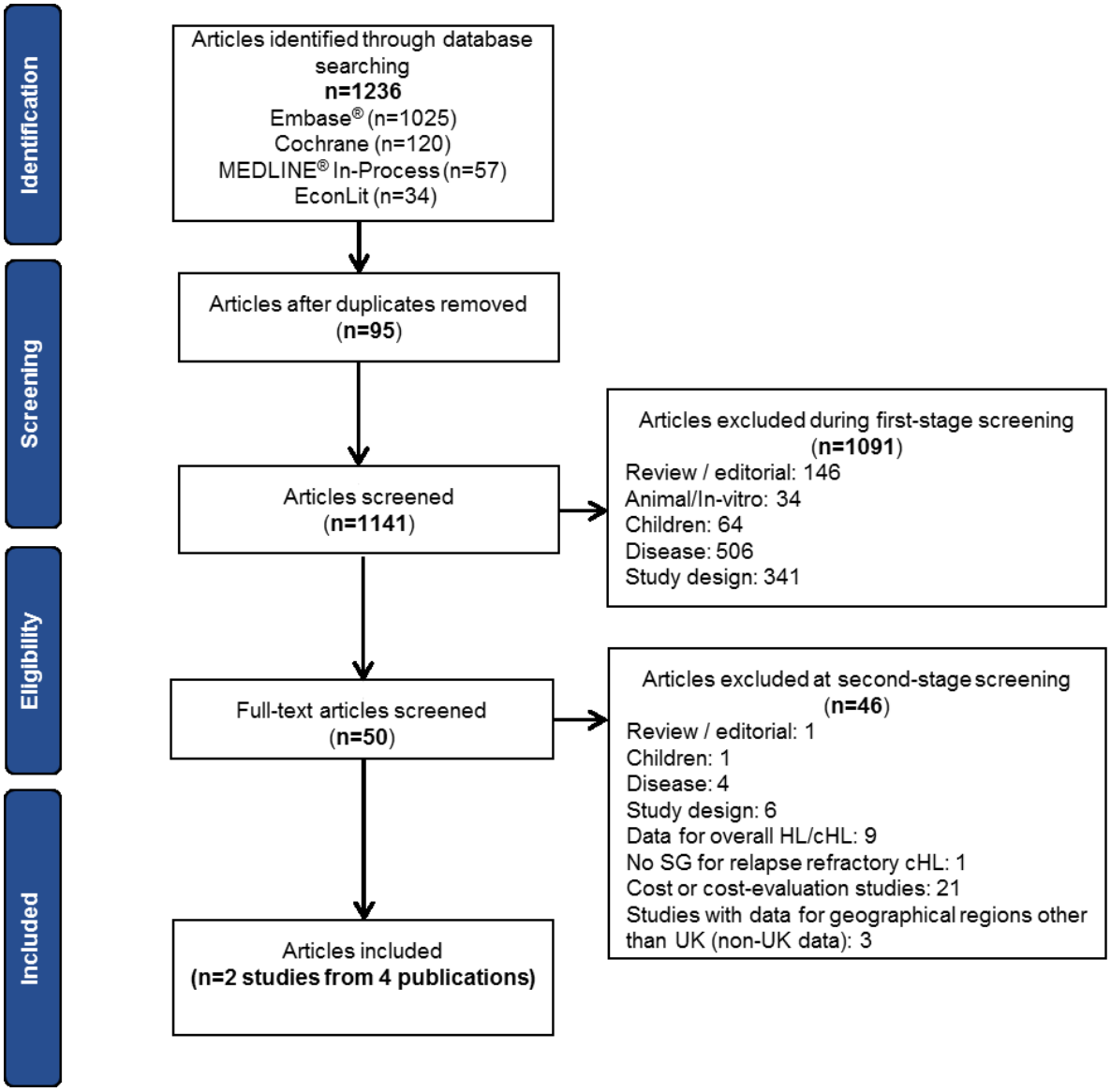
Appendix 12 provides details of the search strategies for HRQoL and utilities and the eligibility criteria set out in the final protocol can be found in Appendix 15.

Systematic searches identified 1,236 separate references. Due to an overlap of evidence across different databases, 95 abstracts were removed as duplicates. Initial screening of the titles and abstracts of the remaining 1,141 citations yielded 50 relevant references, which were evaluated as full-text articles. Of these 50 references, four studies met the inclusion criteria of the review. Finally, having linked the multiple publications from a single study, two studies from four publications were included in the SLR (see PRISMA flow diagram in Figure 28)

Two relevant review studies were identified:

- Swinburn et al, Health utilities in relation to treatment response and AEs in RRcHL and systemic anaplastic large cell Lymphoma [103]
- Ramsay et al, Quality of life results from a phase 3 study of BV consolidation following autoSCT for persons with Hodgkin Lymphoma [104].

Figure 28: PRISMA Diagram: HRQoL and Utility studies



5.4.4 Details of studies in which HRQoL was measured

Swinburn et al. (2015) reports the results of a systematic review and vignette study of HSU in RRcHL. The absence of published literature was the motivation for Swinburn et al to conduct a vignette study to elicit HSU relating to nine health states for RRcHL. These states included complete response, partial response, stable disease with and without B-symptoms (weight loss, fever and night sweats), progressive disease and four states combining complete response with acute graft versus host disease, chronic graft versus host disease and grades I-II and III-IV peripheral sensory neuropathy. The valuation exercise was performed in a representative sample of the UK (n=100), Australian (n=75), Thai (n=75), Taiwanese (n=75), South Korean (n=75), Brazilian (n=101) and Mexican (n=100) general public. Each participant in the study was presented with clinician-validated vignettes describing each of the health states and was asked to value each state using the time trade off method. A summary of selected HSU data from the UK are presented in Table 76.

Table 76: Swinburn 2015 UK utility values

Response status	Mean	SE
Complete response	0.91	0.008
Partial response	0.79	0.017
Stable disease	0.71	0.020
Progressed disease	0.38	0.028

Another study by Ramsey and colleagues assessed the impact of BV as consolidation therapy on quality of life of patients with high risk of relapse after post-ASCT (Ramsey 2016)[104]. The index scores were imputed from baseline through end of treatment at a 3-months interval. Following 6-months of treatment, utility scores of BV were lower than that of placebo. The scores over the treatment period did not show any difference between BV and placebo arm. For both the treatment groups decrease in quality of life was observed indicating that BV did not have a sustained impact of quality of life of patients with relapsed or refractory cHL. Based on the utility data being presented by response status in Swinburn [103], which follows the economic model structure, any literature required utility values have been sourced from here.

5.4.5 Key differences between values derived from literature search and those reported from clinical trials

The majority of the studies and the HTA submission identified do not use EQ-5D data, using mainly EORTC QLQ-C30 questionnaire. The results presented focus either on the impact on HRQoL by treatment group or on specific symptoms of the disease such as pain and fatigue.

Only one of the studies or the HTA submission identified from the SLR estimated utilities for SoC pre and post progression which is in line are in line with the utilities utilised in TA462 [75].

5.4.6 Describe how adverse reactions affect HRQoL

The cost and quality of life burden of treatment related AEs are captured in the model. This is restricted to AEs experienced while on initial therapy and does not include events that may result from further treatment.

- The following criteria were applied for the inclusion of AEs:
- all causes, including those not considered specific to treatment
- grade 3+ AE, according to the Common Terminology Criteria for AEs (CTCAE)
- ≥0% incidence in any study arm

Neither KEYNOTE-087 nor the published literature reported the HSU loss from key grade 3+ AEs considered of interest to the economic analysis. In the absence of data for RRcHL, alternative input sources were identified in oncology (leukaemia, lung, breast, soft tissue carcinoma and pancreatic cancer) and post myocardial infarction. Further detail of the population, valuation method and country of each study is provided in Table 77.

Table 77: Summary of disutility sources

Source	Disease area	Population (sample size)	Method of valuation	Country
Beusterien (2010) [105]	Chronic lymphocytic leukaemia	General public (n=89)	SG	UK
Doyle (2008) [106]	Non-small cell lung cancer	General public (n=101)	SG & VAS	UK
Lloyd (2006) [107]	Breast cancer	General public (n=100)	SG	UK
Nafees (2008) [108]	Small cell lung cancer	General public (n=100)	TTO	UK
Shingler (2013) [109]	Soft tissue sarcoma	General public (n=100)	TTO	UK
Tolley (2013) [110]	Late-stage chronic lymphocytic leukaemia	General public (n=110)	TTO	UK
PEGASUS-TIMI 54 study (TA420) [111]	Post myocardial infarction	Trial population (n=21,162 [n=118,745 completed questionnaires; 0 to 54 months])	EQ-5D-3L (UK value set)	Global

Dyspnea was the only trial based disutility estimate identified. All other AEs were based on general population estimates; therefore where multiple sources were available an average was taken across the studies.

Adverse events durations from TA306 [83] and TA360 [112] were applied over all other sources as they were derived directly using patient level data from a phase III study in relapsed aggressive non-Hodgkin's Lymphoma and a phase III study in patients with locally advanced untreated pancreatic cancer, respectively. All other identified durations from previous submissions were either based on clinical expert opinion or assumptions. When durations were reported in both submissions an average was taken. A summary of the HSU data inputs are provided in Table 78.

Table 78: Adverse event disutilities and durations

Adverse event	Disutility	Source	Used in	Duration (days)	Source
Anaemia	-0.09	Beusterien (2010) [105]	TA462	16.1	TA306
				(12.4+14.5)/2=13.45	TA360
Diarrhoea	-0.08	Beusterien (2010) [105]	TA462	(5.567+5.5)/2 = 5.53	TA360
	-0.0468	Nafees (2008) [108]	TA395		
	-0.103	Lloyd (2006) [107]			
	-0.327	Shingler (2013) [109]			
Dysponea	-0.0481	PEGASUS-TIMI 54 study (TA420) [111]	TA420	12.7	TA306
Fatigue	-0.07346	Nafees (2008) [108]	TA462; TA440; TA411; TA395	31.5	TA306
	-0.262	Shingler (2013) [109]	TA440	(19.885+19.14)/2 = 19.51	TA360
	-0.115	Lloyd (2006) [107]			
Leukopenia	Assumed same as neutropenia			14	TA306
				(10.041+10.4) = 10.22	TA360
Nausea	-0.04802	Nafees (2008) [108]	TA462 TA411 TA395 TA360	6	TA306
	-0.357	Shingler (2013) [109]	--	(11.179+20.933)/2 = 16.06	TA360
	-0.05	Beusterien (2010) [105]			
Neutropenia	-0.08973	Nafees (2008) [108]	TA462	15.1	TA306
	-0.163	Tolley (2013) [110]	TA359	(9.547+9.291)/2 =9.42	TA360
Pyrexia	-0.11	Beusterien (2010) [105]		12.3	TA306
Thrombocytopenia	-0.108	Tolley (2013) [110]	TA462TA359 TA360	23.2	TA306
				(8.057+9.32)/2 = 8.69	TA360
Vomiting	-0.04802	Nafees (2008) [108]	TA462 TA411	2.3	TA306
	-0.357	Shingler (2013) [109]	--	(5.852+10.875)/2 = 8.36	TA360
	-0.05	Beusterien (2010) [105]			
	-0.103	Lloyd (2006) [107]			

The QALY loss associated with AEs in the model is estimated by combining data on the HSU loss (disutility) and mean duration of each AE from the published literature. These were then multiplied by the incidence to give a one-off disutility applied in the first cycle of the model (assuming all AEs occur within 12 months of treatment initiation).

5.4.6.1 Complications of alloSCT

Given that the review was focused on identifying RRcHL specific utility data, a targeted search was undertaken to identify any EQ-5D data collected in patients post alloSCT. One study of interest was identified:

Kurosawsa et al. 2015 [113] conducted a cross-sectional questionnaire study which elicited utility values (via EQ-5D [96% completion rate]) for patients with acute leukaemia (n=524) of which 338 patients were post alloSCT. A utility decrement of 0.15 was calculated for patients experiencing graft versus host disease related symptoms after alloSCT.

Table 79: Kurosawsa 2015 utility values

Description	Mean
AlloSCT (no graft versus host disease)	0.80
AlloSCT (with graft versus host disease)	0.65

5.4.7 Definition of the health states in terms of HRQoL in the cost-effectiveness analysis.

Treatment specific constant health state utility (HSU) values were applied in the PF states (pre- and post-12 weeks). The treatment specific values were calculated from the values presented from response specific values from KEYNOTE-087 multiplied by the response rates from KEYNOTE-087 and Cheah for pembrolizumab and SoC, respectively. This was the preferred approach of the committee in TA462 [75] (response specific values from a single utility source).

Table 80: Calculation of PF HSUVs using KEYNOTE-087 response specific utility values

	Response utility values from KEYNOTE-087	Pembrolizumab (cohort 1)		Pembrolizumab (cohort 2)		SoC	
		Response (%)[KEYNOTE -087]*	Input	Response (%) [KEYNOTE -087]*	Input	Response (%) [Cheah 2016]	Input
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial response	[redacted]	[redacted]		[redacted]		[redacted]	
Stable disease	[redacted]	[redacted]		[redacted]		[redacted]	
*patients that had progressed at week 12 were excluded; those who were not assessed were assumed to have stable disease							

For the PD health states (pre- and post-12 weeks), HSU data were available from KEYNOTE-087 (EQ-5D) and time-trade off values from Swinburn et al [103]. While EQ-5D values are generally preferred to time trade off values, it is noted that the HSU decrement for progression in KEYNOTE-087 is small indicating that progression is not predictive of a meaningful decrement in QoL. However, this is expected given that the values applied within the model are from week 12 of the trial - EQ-5D was only captured up to 30 days post discontinuation; therefore the gradual decline in QoL will not have been captured sufficiently. A progression decrement of [redacted] was calculated from Swinburn et al [103] as SD ([redacted]) minus PD (0.39). This decrement was then applied to the SD values estimated from KEYNOTE-087 (mean HSU [redacted]). This approach was validated with UK clinical experts

5.4.8 Clarification on whether HRQoL is assumed to be constant over time in the cost-effectiveness analysis

Treatment specific constant health state utility (HSU) values were applied in the PF and PD states (pre- and post-12 weeks). The HSU value post 100 days consisted of response specific KEYNOTE-087 (Table 75) values multiplied by the response rates 100 days post alloSCT from Lafferty [80]. (During the first 100 days following alloSCT a decrement was assumed following the findings of the review by Pidala [81] that patients experience a decline in QoL immediately following alloSCT but return to baseline values or improve after 100 days. It was assumed that acute graft versus host disease would be the principal source of the QoL decrement during the first 100 days. Therefore, the decrement was calculated as the difference in utility (estimated from patients EQ-5D) between those with and without GVHD symptoms after alloSCT from Kurosawsa [113] (0.80 without and 0.65 with graft versus host disease symptoms); multiplied by the proportion of patients that had graft versus host disease (61.54%) in the

alloSCT survival data [80]. This decrement was applied to the HSU values 100 days post alloSCT.

Table 81: Calculation of alive HSUV using KEYNOTE-087 response specific utility values

	Response utility values	AlloSCT (post 100 days)		AlloSCT (pre 100 days)	
		Response (%)	Input	GVHD (%)	Input
Complete response	[redacted]	70.0	0.865	61.5	0.773
Partial response	[redacted]	30.0			
Stable disease	[redacted]	0.0			

Consistent with the two previous Hodgkin Lymphoma NICE appraisal (TA462 [75], TA446 [32]), age related utility decrements were applied in all health states. These were applied in each weekly cycle of the model; derived from UK population norms [114] (Table 82) conditional on the model start age (cohort 1: 34.0 cohort 2: 40.0) and gender (cohort 1: 47.8 cohort 2: 46.9).

Table 82: Age related utility decrements [114]

Age	All	Male	Female
<25	0.94	0.94	0.94
25-34	0.93	0.93	0.93
35-44	0.91	0.91	0.91
45-54	0.85	0.84	0.85
55-64	0.80	0.78	0.81
65-74	0.78	0.78	0.78
75+	0.73	0.75	0.71

5.4.9 Description of whether the baseline HRQoL assumed in the cost-effectiveness analysis is different from the utility values used for each of the health states

Not applicable.

5.4.10 Description of how and why health state utility values used in the cost-effectiveness analysis have been adjusted, including the methodologies used

As described in 5.4.8.

5.4.11 Identification of any health effects found in the literature or clinical trials that were excluded from the cost effectiveness analysis

No health effects on patients were excluded from the cost effectiveness analysis. All relevant safety and efficacy has been included as specified in the sections above.

5.4.12 Summary of utility values chosen for the cost-effectiveness analysis

The utility values chosen for the cost-effectiveness model are presented in Table 83.

Table 83: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	Reference in submission	Justification
Pembrolizumab (cohort 1): progression-free	[redacted] [48]	Table 80	Response specific values based on same utility source; consistent with TA462 committee preference
Pembrolizumab (cohort 2): progression-free	[redacted] [48]		
SoC: progression-free	[redacted] [48]		
AlloSCT (<100 days)	0.773 (0.077) [48]	Table 81 Table 81	Finding from Pidala (2009) that patients experience significant decrease in QoL for the first 100 days following alloSCT
AlloSCT (>100 days)	0.865 (0.087) [48]		Response specific values based on same utility source; consistent with TA462 committee preference
Progressed disease	[redacted] [48]	Table 80	Sufficient decrement not captured within KEYNOTE-087 analysis; most relevant value available from the literature applied to KEYNOTE-087 values
Anaemia	-0.0036 (NA) [32]	Table 78	Disutilities associated with grade 3+ treatment related AEs from published literature
Diarrhoea	-0.0021 (NA) [32]		
Dysponea	-0.0017 (NA) [32]		
Fatigue	-0.0105 (NA) [32]		
Leukopenia	-0.0042 (NA) [32]		
NeutropeniaNausea	-0.0046 (NA) [32]		
Neutropenia	-0.0042 (NA) [32]		
Pyrexia	-0.0037 (NA) [32]		
Thrombocytopenia	-0.0047 (NA) [32]		
Vomiting	-0.0020 (NA) [32]		

[1] Assumed 10% of mean, to reflect additional uncertainty in combining response and/or multiple utility values

[2]Disutilities and durations sampled individually in probabilistic analysis

5.4.13 Clinical expert assessment of utility values

The applicability of the selected health state utility values was validated by UK clinical experts. Specifically, discussion over the PD utility utilised in this submission was deemed appropriate as it was considered that a decrement to QoL more than that apparent at 12 weeks from the KEYNOTE-087 data would be seen over the long term. The utility decrement applied to alloSCT in the first 100 days was also deemed appropriate through validation with clinical experts.

5.5 Cost and healthcare resource use identification, measurement and valuation

5.5.1 Parameters used in the cost effectiveness analysis

The full list variables used in the cost effectiveness analysis is presented in Appendix 16.

5.5.2 Resource identification, measurement and valuation studies

The type of costs considered in the economic model included the drug and administration costs related to the intervention and comparator, including the costs related to subsequent therapies (see section 5.5.5), the monitoring and management of the disease (see section 5.5.6), the management of AEs (see section 5.5.7), and the costs related to terminal care (see section 5.5.6).

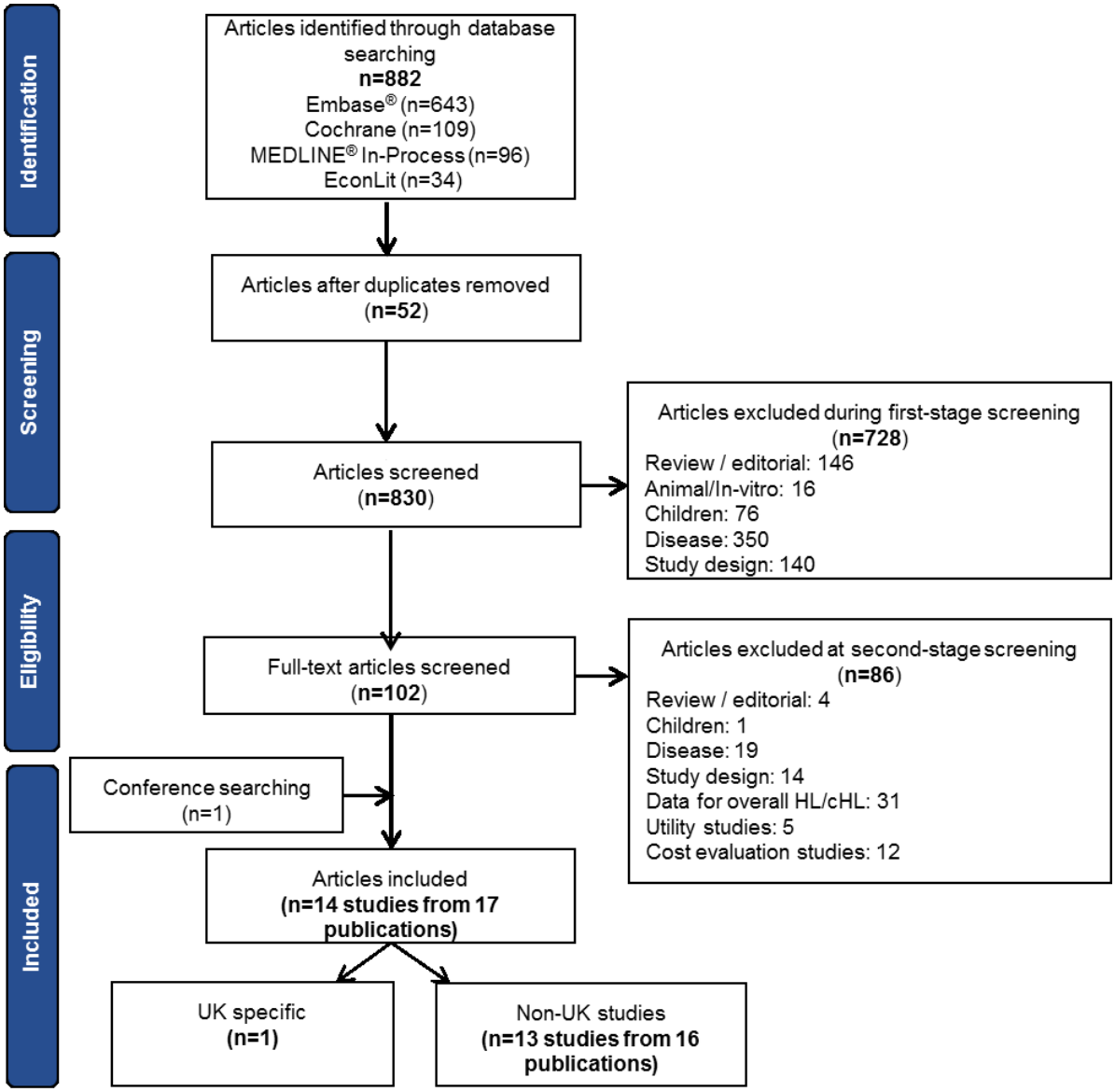
A comprehensive literature search was conducted on 12th July 2017 from 2001 to identify costs and resource use in the treatment and on-going management of RRcHL patients. The search was limited to only include studies published since 2001. While the scope of the searches was broad only studies from UK NHS perspective where finally included in the SLR results.

The searches conducted for resource use data and the selection criteria followed for the identification and inclusion of relevant studies are provided in Appendix 12.

The systematic database searches 882 records for cost and resource use studies. Due to an overlap of evidence across different databases, 52 abstracts were removed as duplicates. Initial screening of the titles and abstracts of the remaining 830 citations yielded 102 relevant references, which were evaluated as full-text articles. Of these 102 references, 16 studies met the inclusion criteria of the review and one study was retrieved through conference searching. Finally, having linked the multiple publications from a single study, one study specific to UK based data and 13 studies from 16 publications from other geographical regions were included

in the SLR. Details of the studies can be found in Appendix 14. The UK study, Radford (2013) [115], was of relapsed HL patients after autoSCT, chemotherapy followed by alloSCT. There was however an update to this publication in 2017 [28] which was excluded from the search as it was a review article. However it was the preferred source of cost and resource use in TA462 and hence the Radford 2017 information have been used in the economic analysis presented here.

Figure 29: PRISMA diagram for included cost and resource use studies



5.5.3 Use of NHS reference costs or payment-by-results (PbR) tariffs

There are no NHS reference costs or payment-by-results (PbR) tariffs specific for costing pembrolizumab. Details about the cost estimation of treatment with pembrolizumab in terms of acquisition and administration are reported below. As previously agreed with NHS England (personal communication, 9th December 2014) for the single technology assessment (STA) submission of pembrolizumab for advanced melanoma [116], the administration cost of pembrolizumab can be reflected through NHS Reference Cost code SB12Z [117], since this corresponds to the administration of a simple therapy (i.e. involving the administration of only one agent without IV anti-emetics), with the infusion lasting less than one hour.

5.5.4 Input from clinical experts

The costing approach detailed here was previously validated with clinical experts in previous HTA submissions of pembrolizumab [116] [118].

5.5.5 Intervention and comparators' costs and resource use

The costs of acquisition, administration and monitoring are assumed to apply for the duration that people remain on therapy in the model. This duration is dependent on a number of factors that vary across treatments. These include drug efficacy in terms of PFS, its tolerability and AE profile, and any restrictions on the maximum number of cycles permitted on treatment.

5.5.5.1 Cost of Drug

The drug acquisition costs per treatment are presented below, with the unit costs for comparator regimens being taken from the latest electronic market information tool [119] (eMit) published in 2017, which provides information about prices for generic drugs based on the average price paid by the NHS over the last four months. Table 84 summarises the costs of pembrolizumab, including acquisition and administration that are applied in the model.

Table 84: Summary costs associated with the technology in the economic model

Items	Intervention	Cost	Standard of Care	Cost and range
Technology cost	100mg vial = £2,630	£2,630	See Table 88 for SoC composition	See Table 88
Cost of technology treatment per cycle	200 mg for cHL by intravenous infusion over 30 minutes every 3 weeks (£2,630 x 2)	£5,260[63, 12066, 12666, 12566, 125]	Usage weighted acquisition cost per cycle	£0 - £1,619 [119 120] *Investigational agent acquisition cost assumed to be free to the health system*
Administration cost	Deliver Simple Parenteral Chemotherapy at First Attendance (SB12Z)	£236.19	Weighted cost (per cycle) using NHS reference costs for delivering complex chemotherapy, including prolonged infusion treatment, at first attendance (SB14Z) and subsequent attendance (SB15Z) as appropriate.	£383 - £1,367 [117] *Investigational agent administration cost assumed to be free to the health system*
Monitoring cost	N/A	£0	N/A	£0
Tests	N/A	£0	N/A	£0

The tables (detailing the cost and healthcare resource use) that follow are separated into sections pertaining to (i) Intervention - pembrolizumab, (ii) Comparator - standard of care (SoC), and (iii) Subsequent therapy – best supportive care (BSC).

Pembrolizumab costs

Table 85 to Table 87 detail the resource-use and costing for pembrolizumab in RRcHL.

Table 85: Dosing and cycle description for pembrolizumab in RRcHL

Dosing	Cycle
200mg on 1 day per cycle	Cycle length of 21 days, to a maximum of 35 cycles (ca.2 years)
Source: MSD	

As per the licence, the model used a fixed pembrolizumab dosage of 200mg by intravenous infusion over 30 minutes every 3 weeks (see the Summary of Product Characteristics [SmPC] in Appendix 1) thus requiring two 100mg vials every 3 weeks costing £5,260 (£2,630 x 2) for the medicinal form per cycle and £236.19 [121] for the administration costs. [redacted]

Table 86: Acquisition cost of pembrolizumab

Component	Strength	Units per pack	Pack cost	Cycle Cost	Cycle cost with CAA	Source
Pembrolizumab 100mg powder for concentrate for solution for infusion vials	100mg	1	£2,630.00	£5,260.00	[redacted]	MSD

Regarding the administration of pembrolizumab (Table 87), as previously agreed with NHS England in NICE submissions of pembrolizumab (TA357 [122] and TA428 [118]) the administration cost of pembrolizumab can be reflected through the NHS Reference Cost code SB12Z [117]. This corresponds to the administration of a simple therapy involving the administration of only one agent (pembrolizumab) without intravenous anti-emetics and the infusion lasting only 30 minutes.

Table 87: Administration cost per cycle of pembrolizumab

NHS Reference (HRG Code)	Administration cost
Deliver Simple Parenteral Chemotherapy at First Attendance (SB12Z)	£236.19
Source: DoH 2016 [117]	

SoC Costs

Costs of SoC are based on the costs required for each of the following components:

- Chemotherapy: assumed to be equal usage of all regimens specified for the treatment of relapsed or refractory HL within BCSH guidelines.
- Bendamustine
- Investigational agents

Table 88 summarises the regimens included that form the SoC comparator arm of the model. The source for chemotherapy regimens are derived both from the "guideline on the management of primary resistant and relapsed classical Hodgkin Lymphoma" as published in the British Journal of Haematology [3] and from a previous NICE technology appraisal TA462 [57]. The proportion of each treatment was assumed from the SoC efficacy data [44], (excluding BV-retreatment, autoSCT and Other) and assuming all patients not treated with bendamustine or investigational agents were distributed equally between the chemotherapy regimens.

Table 88: Composition of SoC

Treatment	Included regimens	Percentage	Source
Chemotherapy*	ASHAP, DHAOx, DHAP, ESHAP, GDP , GEM-P, GVD, ICE, IGEV, IVE, IVOx, MINE	38.5% (3.2% per regimen)	Remaining percentage
Bendamustine	--	18.5%	
Investigational agents	--	43.1%	

*NB - Dexa and Mini-BEAM regimens excluded from chemotheraoy and BV re-treatment removed.
Abbreviations: ASHAP: doxorubicin, methylprednisolone, cytarabine, cisplatin; DHAOx: dexamethasone, cytarabine, oxaliplatin; DHAP: dexamethasone, cytarabine, cisplatin; ESHAP: etoposide, methylprednisolone, cytarabine, cisplatin; GDP: gemcitabine, dexamethasone, cisplatin; GEM-P: gemcitabine, cisplatin, methylprednisolone; GVD: gemcitabine, vinorelbine, liposomal doxorubicin; ICE: ifosfamide, carboplatin, etoposide; IGEV: ifosfamide, gemcitabine, vinorelbine; IVE: ifosfamide, epirubicin, etoposide; IVOx: ifosfamide, etoposide, oxaliplatin; MINE: mitoxantrone, ifosfamide, vinorelbine, etoposide

Table 89 details the dosing and cycle details for the specific regimens and components of SoC.

Table 89: SoC dosing and cycle descriptions

Treatment/Regimen		Dosing	Cycles
ICE	Ifosfamide	5000mg/m² on 1 day per cycle	Cycle length of 14 days, to a maximum of 3 cycles
	Mesna	5000mg/m² on 1 day per cycle	
	Carboplatin	800mg on 1 day per cycle	
	Etoposide	100mg/m² on 3 days per cycle	
IVE	Ifosfamide	3000mg/m² on 3 days per cycle	Cycle length of 21 days, to a maximum of 3 cycles
	Mesna	3000mg/m² on 3 days per cycle	
	Eporubicin	50mg/m² on 1 day per cycle	
	Etoposide	200mg/m² on 3 days per cycle	
MINE	Mitoxantrone	8mg/m² on 1 day per cycle	Cycle length of 28 days, to a maximum of 2 cycles
	Ifosfamide	1330mg/m² on 3 days per cycle	
	Mesna	1330mg/m² on 3 days per cycle	
	Mesna	500mg on 3 days per cycle	
	Etoposide	65mg/m² on 3 days per cycle	
IVOx	Ifosfamide	1500mg/m² on 3 days per cycle	Cycle length of 21 days, to a maximum of 3 cycles
	Mesna	1500mg/m² on 3 days per cycle	
	Carboplatin	150mg/m² on 3 days per cycle	
	Oxaliplatin	130mg/m² on 1 day per cycle	
IGEV	Ifosfamide	2000mg/m² on 4 days per cycle	Cycle length of 21 days, to a maximum of 4 cycles
	Mesna	2600mg/m² on 4 days per cycle	
	Gemcitabine	800mg/m² on 4 days per cycle	
	Vinorelbine	20mg/m² on 1 day per cycle	
	Prednisolone	100mg on 4 days per cycle	
GEM-P	Gemcitabine	1000mg/m² on 3 days per cycle	Cycle length of 28 days, to a maximum of 3 cycles
	Cisplatin	100mg/m² on 1 day per cycle	
	Methyl-prednisolone	1000mg on 5 days per cycle	
GDP	Gemcitabine	1000mg/m² on 2 days per cycle	Cycle length of 21 days, to a maximum of 2 cycles
	Dexamethasone	40mg on 4 days per cycle	
	Cisplatin	75mg on 1 day per cycle	
GVD	Gemcitabine	1000mg/m² on 2 days per cycle	Cycle length of 21 days, to a maximum of 2 cycles
	Vinorelbine	20mg/m² on 2 days per cycle	
	Pegylated liposomal doxorubicin	15mg/m² on 2 days per cycle	

ESHAP	Etoposide	50mg/m² on 4 days per cycle	Cycle length of 28 days, to a maximum of 4 cycles
	Methyl-prednisolone	500mg on 4 days per cycle	
	Cytarabine	2000mg/m² on 1 day per cycle	
	Cisplatin	25mg/m² on 4 days per cycle	
ASHAP	Doxorubicin	10mg/m² on 4 days per cycle	Cycle length of 28 days, to a maximum of 3 cycles
	Methyl-prednisolone	500mg on 5 days per cycle	
	Cytarabine	1500mg/m² on 1 day per cycle	
	Cisplatin	25mg/m² on 4 days per cycle	
DHAP	Dexamethasone	40mg on 4 days per cycle	Cycle length of 21 days, to a maximum of 2 cycles
	Cytarabine	2000mg/m² on 1 day per cycle	
	Cisplatin	100mg/m² on 1 day per cycle	
DHAOx	Dexamethasone	40mg on 4 days per cycle	Cycle length of 21 days, to a maximum of 4 cycles
	Cytrabine	2000mg/m² on 2 days per cycle	
	Oxaliplatin	130mg/m² on 1 day per cycle	
Bendamustine		120mg/m² on 2 days per cycle	Cycle length of 28 days, to a maximum of 6 cycles

The comparator acquisition costs (Table 90) were obtained from estimates of the average price paid for products in the NHS; the drugs and pharmaceutical electronic market information (eMit) [119]. Where the average prices for particular therapies were not available via eMit, prices were obtained from the current British National Formulary (BNF) [120].

Since many of the components in the UK (Table 90) are available in different strengths and pack size, the model contained up to a maximum of four vial/pack size for each component. The model calculated the lowest cost combination of vials to make up the required dosage for the average patients including drug wastage.

Table 90: Unit cost from sourced prices of SoC regimen components in the UK

Component	Strength	Units per pack	Pack cost (£)	Source	Cost per unit (£)
Ifosfamide	1000mg	1	£91.32	BNF	£91.32
	2000mg	1	£179.88		£179.88
Mesna	400mg	15	£201.15	BNF	£13.41
	1000mg	15	£441.15		£29.41
Carboplatin	50mg	1	£3.25	eMIT	£3.25
	150mg	1	£7.49		£7.49

	450mg	1	£20.39		£20.39
	600mg	1	£27.89		£27.89
Etoposide	100mg	1	£2.09	eMIT	£2.09
	500mg	1	£9.10		£9.10
Epirubicin	10mg	1	£2.57	eMIT	£2.57
	50mg	1	£9.02		£9.02
	200mg	1	£24.24		£24.24
Mitoxantrone	20mg	1	£45.40	eMIT	£45.40
Oxaliplatin	50mg	1	£3.40	eMIT	£3.40
	100mg	1	£8.77		£8.77
Gemcitabine	200mg	1	£2.76	eMIT	£2.76
	1000mg	1	£7.96		£7.96
	2000mg	1	£16.52		£16.52
Vinorelbine	10mg	10	£43.47	eMIT	£4.35
	50mg	1	£17.56		£17.56
Prednisolone	1mg	28	£0.26	eMIT	£0.01
	5mg	28	£0.41		£0.01
	25mg	56	£26.19		£0.47
Cisplatin	10mg	1	£1.99	eMIT	£1.99
	50mg	1	£6.48		£6.48
	100mg	1	£8.45		£8.45
Methyl-prednisolone	40mg	10	£13.69	eMIT	£1.37
	125mg	1	£4.79		£4.79
	500mg	1	£3.96		£3.96
	1000mg	1	£7.24		£7.24
Dexamethasone	0.5mg	28	£38.85	eMIT	£1.39
	2mg	50	£28.93		£0.58
	2mg	100	£58.35		£0.58
Pegylated liposomal doxorubicin	20mg	1	£360.23	BNF	£360.23
	50mg	1	£712.49		£712.49
Cytarabine	100mg/1mL	5	£16.86	eMIT	£3.37
	500mg	5	£19.26		£3.85
	1000mg	1	£5.69		£5.69
	2000mg	1	£6.60		£6.60
Doxorubicin	10mg	1	£1.34	eMIT	£1.34
	50mg	1	£4.51	eMIT	£4.51
	200mg	1	£16.98		£16.98

Bendamustine	100mg	1	£27.77	BNF	£27.77
	25mg	1	£6.85		£6.85
	25mg	1	£6.85		£6.85
Source: [119]					

The acquisition cost per administration for each component of the various regimens were multiplied by their respective frequency in a given cycle to derive a per cycle cost for each regimen.

The cost of administration for the various SoC chemotherapy regimens (Table 91) were obtained from NHS reference costs [117]. Specifically codes SB14Z and SB15Z have been applied pertaining to delivering complex chemotherapy at first attendance and delivering subsequent elements of a chemotherapy cycle respectively in line with a previous submission (TA462).

Table 91: Administration cost of SoC per cycle

Regimen	Administration cost	Description
ICE	£711.23	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
IVE	£1,039.33	Delivering complex chemotherapy at first attendance and delivering two subsequent complex chemotherapy elements within the same cycle
MINE	£1,039.33	Delivering complex chemotherapy at first attendance and delivering two subsequent complex chemotherapy elements within the same cycle
IVOx	£1,039.33	Delivering complex chemotherapy at first attendance and delivering two subsequent complex chemotherapy elements within the same cycle
IGEV	£1,367.43	Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle
GEM-P	£711.23	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
GDP	£383.13	Delivering complex chemotherapy at first attendance

GVD	£711.23	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
ESHAP	£1,367.43	Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle
ASHAP	£1,367.43	Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle
DHAP	£383.13	Delivering complex chemotherapy at first attendance
DHAOx	£383.13	Delivering complex chemotherapy at first attendance
Bendamustine	£383.13	Delivering complex chemotherapy at first attendance
Where delivering complex chemotherapy at first attendance is £383.13 (SB14Z) and delivering a subsequent complex chemotherapy element within the same cycle is £328.10 (SB15Z) Source: [117]		

Table 92 presents the acquisition cost per cycle of the various regimens used in SoC alongside the respective cycle length and maximum number of treatment cycles.

Table 92: Acquisition costs per cycle and maximum number of cycles

Regimen	Cost per cycle (£)	Cycle length (days)	Maximum number of cycles
ICE	£1,230.82	14	3
IVE	£2,183.65	21	3
MINE	£1,209.02	28	2
IVOx	£1,132.46	21	3
IGEV	£2,109.48	21	4
GEM-P	£100.86	28	3
GDP	£93.06	21	2
GVD	£1,491.60	21	2
ESHAP	£63.32	28	4
ASHAP	£68.73	28	3
DHAP	£76.39	21	2
DHAOx	£89.69	21	4
Bendamustine	£123.30	28	6
Source: [3 119 120 32]			

The cost of each regimen is applied in the model at the start of each of the respective treatment cycles until the maximum treatment duration (for example the acquisition and administration

cost per cycle of ASHAP was applied at the start of week 0, 4, 8 and 12). The total weighted cost of SoC was calculated using the composition of SoC (presented in Table 88) which was subsequently multiplied by the proportion of patients in the PF state in the SoC arm at the start of any given cycle.

5.5.6 Health-state unit costs and resource use

The published data exploring in detail the resource use associated with patients with previously treated RRcHL is limited. Consequently the main source of resource utilisation used in this submission comes mainly from published NICE TAs.

Subsequent therapy – best supportive care (BSC)

Best supportive care (BSC) is applied as a one-off cost to all patients upon progression (subsequent treatment) in the base case and included as a direct comparator in scenario analysis.

Table 93 summarises the regimens included for BSC. It was necessary to use regimens that pertain to non-Hodgkin's Lymphoma (NHL) since there is a large amount of uncertainty, and very little information available, regarding the details surrounding BSC for cHL in the UK. This same approach was adopted in a previous submission TA462. [75]

Table 93: Composition of BSC

Therapy	Distribution of patients across therapies (%)
Gemcitabine monotherapy (administered over 4 weeks)	8.33
RVIG	16.67
DHAP	11.67
CHOP	1.67
IVAC	3.33
Weekly therapy (PMitCEBO)	8.33
Palliative care	46.67
Clinical trial treatment	3.33

Abbreviations: CHOP: cyclophosphamide, doxorubicin, prednisolone, vincristine; DHAP: dexamethasone, cytarabine, cisplatin; IVAC: cytrabine, etoposide, ifosfamide, mesna; PMitCEBO: bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine; RVIG: gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine
Source: [83, 123]

Table 94, Table 95, Table 96, and Table 97 provided details of the dosages, cycles, acquisition unit costs, administration cost and expected treatment duration for each regimen included in BSC. Palliative care and clinical trials were assumed to have no cost in line with assumption made in TA462.

Table 94: Dosing and cycle descriptions

Regimen		Dosing	Cycles
Gemcitabine (monotherapy) [124]		1000mg/m2 (IV) on 3 days per cycle	Cycle length of 28 days, to a maximum of 6 cycles
RVIG [92]	Ifosfamide	2000mg/m² (IV) on 4 days per cycle	Cycle length of 21 days, to a maximum of 4 cycles
	Mesna	2600mg2000mg/m² (IV) on 4 days per cycle	
	Mesna	400 mg/m² (IV) twice on 4 days per cycle	
	Gemcitabine	800mg/m² (IV) on 4 days per cycle	
	Vinorelbine	20mg/m² (IV) on 1 day per cycle	
	Prednisolone	100mg (PO) on 4 days per cycle	
	Rituximab	375mg/m2 (IV) on 1 day per cycle	
DHAP [98]	Dexamethasone	40mg (PO) on 4 days per cycle	Cycle length of 21 days, to a maximum of 2 cycles
	Cytarabine	2000mg/m² (IV) on 1 day per cycle	
	Cisplatin	100mg/m² (IV) on 1 day per cycle	
CHOP [125]	Cyclophosphamide	750mg/m² (IV) on 1 day per cycle	Cycle length of 21 days, to a maximum of 8 cycles
	Doxorubicin	50mg/m² (IV) on 1 day per cycle	
	Vincristine	2mg (IV) on 1 day per cycle	
	Prednisolone	100mg (PO) on 5 day per cycle	
IVAC [126]	Etoposide	60mg/m² (IV) on 5 days per cycle	Cycle length of 21 days, to a maximum of 6 cycles
	Cytarabine	2000mg/m2 (IV), twice on 2 days per cycle	
	Mesna	300mg/m² (IV) on 5 days per cycle	
	Ifosfamide	1500mg/m² (IV) on 5 days per cycle	
	Mesna	300mg/m² (IV), twice on 5 days per cycle	
PMitCEBO [123]	Bleomycin	10mg/m² (IV), on 1 day per cycle (day 8)	Although PMitCEBO is otherwise known as “weekly therapy”, the cycle length is 14 days with a maximum of 8 cycles (16 weeks)
	Cyclophosphamide	300mg/m² (IV), on 1 day per cycle	
	Etoposide	150mg/m² (IV), on 1 day per cycle	
	Mitoxantrone	300mg7mg/m² (IV), on 1 day per cycle	
	Prednisolone	50mg (PO), each day per cycle	
	Vincristine	1.4mg/m² (IV), on 1 day per cycle (day 8)	

Table 95: Unit costs from sourced prices of BSC regimen components (UK)

Component	Strength	Units per pack	Pack cost (£)	Source	Cost per unit (£)
Cyclophosphamide	500mg	1	£8.87	eMIT	£8.87
	1000mg	1	£7.84		£7.84
	2000mg	1	£29.55		£29.55
Rituximab	100mg	2	£349.25	BNF	£174.63
	500mg	1	£785.84		£785.84
Vincristine	1mg	5	£15.71	eMIT	£3.14
	2mg	5	£33.31		£6.66
	5mg	5	£90.10		£18.02
Bleomycin	15mg (15,000 unit)	10	£190.60	BNF	£19.60
NB- only components not already reported in Table 90 are listed here Source: [119] [120]					

Table 96: Administration cost of subsequent therapy

Regimen	Cost	Description
Gemcitabine	£236.19	Delivering simple parenteral chemotherapy at first attendance
RVIG	£1,367.43	Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle
DHAP	£383.13	Delivering complex chemotherapy at first attendance
CHOP	£383.13	Delivering complex chemotherapy at first attendance
IVAC	£1,696.53	Delivering complex chemotherapy at first attendance and delivering four subsequent complex chemotherapy elements within the same cycle
PMitCEBO	£711.23	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
Where delivering a simple parenteral chemotherapy at the first attendance is £236 (SB12Z), delivering complex chemotherapy at first attendance is £383.13 (SB14Z), and delivering a subsequent complex chemotherapy element within the same cycle is £328.10 (SB15Z) [117]		

Table 97: Duration of treatment

Therapy	Number of cycles
Gemcitabine monotherapy (administered over 4 weeks)	4.0
RVIG	4.5
DHAP	6.0
CHOP	6.0
IVAC	3.5
Weekly therapy (PMitCEBO)	7.0

The acquisition and administration cost per cycles of each component of BSC were multiplied by the expected duration and expected usage to give a one-off cost of £4,848.22.

Terminal care cost

A terminal care cost is applied upon death, to patients on pembrolizumab or SoC, at a total cost of £4,064.64, to reflect the additional intensive disease management in the months leading up to death. The proportion of patients in each care setting and resource usage was

derived from a previous HTA assessment in non-small cell lung cancer [127] and the respective unit costs were updated. A breakdown of the total cost and sources is provided in Table 98.

The cost was not included for alloSCT in the base case to avoid potential double counting as the cost from Radford [28] included the cost of some patients up until death.

Table 98: Terminal care costs

Care setting	Proportion treated in setting	Resource	Resource usage	Unit cost (£)	Description
Home	27.3%	GP home visit	28 hours	£106.18	Cost of out of surgery visit lasting 23.4 minutes (incl. qualification) PSSRU 2012 updated using HCHS index [121, 128]
		Community nurse visit	7 visits	£72.36	Cost per hour spent on home visit (incl. qualification) PSSRU 2013 updated using HCHS index [128]
		MacMillan nurse	50 hours	£48.26	66.7% of community nurse cost [127]
		Drugs and equipment	--	£306.84	Marie Curie report figure of £240 increased for inflation using HCHS 2004-2016 [121, 128 129]
Hospital	55.8%	--	1 episode	£3,083.76	NHS Reference Costs 2015–2016, weight average of codes DZ17L-V (respiratory neoplasms with without/single/multiple intervention, CC 0-10+), non-elective long stay [117]
		--	0.92 excess bed days	£273.92	NHS Reference Costs 2015–2016, weight average of codes DZ17L-V (respiratory neoplasms with without/single/multiple intervention, CC 0-10+), excess bed days [117]
Hospice	16.9%	--	1 episode	£4,169.70	25% increase on hospital [127]

5.5.7 Adverse reaction unit costs and resource use

Adverse event unit costs (Table 99) were obtained from NHS reference costs using HRG codes applied in previous NICE appraisal. The unit costs were calculated as a weighted average of all HRG codes included for each event using the 'activity' provided in NHS reference costs.

Table 99: Adverse event unit costs

Adverse Event (AE)	AE Details		Unit Cost / Source	Rational
	HRG Codes	Description		

Anaemia	SA03G-H	Haemolytic Anaemia with CC Score 0-3+	£814.03 NHS reference costs 2015-16 [117]	TA411
	SA04G-L	Iron Deficiency Anaemia with CC Score 0-14+		
	SA05G-J	Megaloblastic Anaemia with CC Score 0-8+		
	SA08G-J	Other Haematological or Splenic Disorders, with CC Score 0-6+		TA399, TA391
Diarrhoea	FZ49D-E	Nutritional Disorders with Interventions, with CC Score 0-2+	£1,497.86 NHS reference costs 2015-16 [117]	TA391 TA440
	FZ49F-H	Nutritional Disorders without Interventions, with CC Score 0-6+		
	FZ91A-D	Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions, with CC Score 0-8+		
	FZ91E-H	Non-Malignant Gastrointestinal Tract Disorders with Single Intervention, with CC Score 0-9+		
	FZ91J-M	Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-11+		
Dyspnea	DZ19H	Other Respiratory Disorders with Multiple Interventions	£718.76 NHS reference costs 2015-16 [117]	TA420
	DZ19J-K	Other Respiratory Disorders with Single Intervention, with CC Score 0-5+		
	DZ19L-N	Other Respiratory Disorders without Interventions, with CC Score 0-11+		
Fatigue	WA17X	Other Admissions Related to Neoplasms with Intermediate CC	£1,499.09 [130] [127] Brown (2013) and NHS reference costs 2011-12 inflated with HCHS index	TA391
Leukopenia	SA08G-J	Other Haematological or Splenic Disorders, with CC Score 0-6+	£1,142.90 NHS reference costs 2015-16 [117]	TA391
Nausea	FZ13C	Minor Therapeutic or Diagnostic, General Abdominal Procedures, 19 years and over	£872.42 NHS reference costs 2015-16 [117]	TA411

Neutropenia	SA08G-J	Other Haematological or Splenic Disorders, with CC Score 0-6+ (assumed equal to leukopenia)	£1,142.90 NHS reference costs 2015-16 [117]	TA411 TA399
Pyrexia	WA05Z	Pyrexia of Unknown Origin with length of stay 5 days or more	£3,923.50 NHS reference costs 2013-14 inflated with HCHS index [131]	TA366 TA311
Thrombocytopenia	SA12G-K	Thrombocytopenia with CC Score 0-8+	£636.19 NHS reference costs 2015-16 [117]	TA399 TA440
Vomiting	FZ49D-91M	Assumed equal to diarrhoea	£1,497.86 NHS reference costs 2015-16 [117]	TA360 TA440

5.5.8 Miscellaneous unit costs and resource use

Allogeneic stem cell transplantation

The cost of alloSCT has been taken from Radford [28] as this was the preferred source by the committee in TA462. Radford [28] was a retrospective analysis that studied the cost and resource use in 40 cHL patients who had failed after autoSCT. A total of 15 patients subsequently received alloSCT and were followed up to date of death or to most recent follow-up (mean 3.44 years), with a mean total cost of £110,374. Of this total cost 31.5%–39.9% were due to the alloSCT procedure itself ranging between £34,783 and £44,059 per patient. Therefore, a substantial proportion of the cost was associated with additional follow-up cost so no additional disease management costs or AE costs were included in the model for post alloSCT patients.

5.6 Summary of base-case de novo analysis inputs and assumptions

5.6.1 Summary of base-case de novo analysis inputs

The full list variables used in the cost effectiveness analysis is presented in Appendix 16

5.6.2 For the base-case de novo analysis the company should ensure that the cost-effectiveness analysis reflects the NICE reference case as closely as possible

The base-case cost-effectiveness analysis reflects the NICE reference case as closely as possible.

5.6.3 Assumptions

A summary of the main assumptions applied within the economic model is provided in Table 100.

Table 100: Model assumptions

Assumption	Rationale
Baseline cohort characteristics were derived from KEYNOTE-087 and were assumed to be representative of RRcHL patients in the UK (age, gender, weight, body surface area)	Given the limited availability of literature in the late stage of this disease, KEYNOTE-087 was the most reliable source. Sensitivity analysis has been conducted to assess the impact of the uncertainty of these parameters
Model outcomes were re-evaluated over weekly cycle periods for a time horizon of up to 40-years	A weekly re-evaluation period was selected as the common denominator between the periods between treatment cycles for pembrolizumab and SoC treatments. Considered a lifetime time horizon in TA462. This was to ensure all important differences in costs and outcomes were reflected (NICE reference case. Despite >2% of the cohort being alive after 40 years, extending the time horizon further was associated with additional uncertainty. Scenario analysis was conducted with 50 year time horizon to test the sensitivity of this assumption.
A year in the model was measured at 52.1786 (365.25/7) weeks, with each weekly cycle equating to 0.2300 months or 0.0198 years	To accurately account for the length of 1 year
Discounting rates were applied over discrete time periods, e.g. fixed discounting in years 1, 2, 3	Aligned with the NICE reference case [76]

Costs and health outcomes were calculated using lifetable mid-cycle estimates, with the exception of one-off costs for subsequent treatment and adverse events, which apply at the start of the model time horizon, and the costs for drug acquisition, administration	Use of the state population at the start or end of a given cycle results in an under or over estimation of the state population for that given cycle. Therefore, a mid-cycle correction is included in the model to mitigate this inherent bias caused by the use of discrete time in state transition models. The lifetable mid-cycle correction method is used where the state population was calculated as the average of the start and end population. This method was chosen over the standard mid-cycle correction where half a cycle is added at the beginning of the model as this requires adjustment when patients are still alive at the end of the Markov trace and can cause issue with discounting Drug acquisition and administration costs were calculated using the number of patients occupying the PF state at the start of each relevant cycle, to reflect that therapy is given at fixed and discrete time points (e.g. every 3 weeks)
Response at week 12: All patients progression-free at week 12 that did not have a complete or partial response are assumed to have stable disease	By definition
Patients received alloSCT at week 12	Week 12 was chosen for the following reasons: The first tumour assessment in KEYNOTE-087 was 12 weeks after treatment initiation For the small proportion of patients in KEYNOTE-087 that received alloSCT the mean number of administration of pembrolizumab prior to alloSCT was [redacted] The finding for the clinician survey indicated from initiation of SoC to commencing alloSCT was 16 weeks (median) [64]
Patients treated with alloSCT did not experience disease progression	The omission of the PD state from the model simplifies the calculation of post-alloSCT survival, and hence the cure rates for alloSCT, as they can be derived directly from post-alloSCT survival data without the complications of modelling transitions between intermediary states such as PF and PD. In addition, the role of PFS in determining the QoL of patients who undergo alloSCT is unclear, given that longitudinal studies suggest that the time since alloSCT plays an important role in determining overall QoL, with an early deficit immediately after transplantation that is followed by a return to pre-transplantation levels by day 100, and stabilization or continuation of this improvement from day 100 up to 3 years of follow-up
Progression-free survival pre- and post- week 12: it was assumed that the proportional hazards assumption held	A large number of progression events occurred during the first 12 weeks of the SoC study (Cheah 2016). Therefore, it was not possible to estimate a hazard ratio (HR) between the two treatments after 12 weeks. This was due to the small number of patients at risk and low number of events after 12 weeks which was associated with substantial levels of uncertainty. Thereby, comparative efficacy was only assessed using a constant HR across both periods (pre- and post- 12 weeks). Thus, it was necessary to assume that proportional hazards held and acknowledge this as a limitation of the analysis.

The hazard ratio effects could be applied to both accelerated failure time and proportion-hazard models	Accelerated failure time models do not follow the proportional hazards assumption, and as such it is considered methodologically incorrect to apply a hazard ratio to these distributions. This inconsistency, while raised in the methods literature, has not precluded the combining of these distributions with hazards ratios in previous technology appraisals. Following past appraisals, we have applied hazard ratios to non-proportional hazard models, such as log normal, although it is acknowledged as a limitation of the analysis
The assessment of progression-free survival was based upon the blinded independent central review and not that of the study investigator	BICR was used as it was the primary end point in KEYNOTE-087
Overall survival for the first 12 weeks: was assumed equivalent across both treatment arms and was modelled using parametric distributions fitted to survival data from the KEYNOTE-087	This conservative assumption was made as any hazard ratio generated between pembrolizumab and SoC was subject to significant uncertainty given the low number of deaths observed in KEYNOTE-087.
Overall survival after alloSCT: patients were subject to the same survival irrespective of prior treatment or response	There is no evidence to suggest that the difference in the mechanism of action has any effect on the efficacy of alloSCT and overall survival data after alloSCT conditional on response prior to alloSCT was not available
Overall survival after alloSCT: the probability of death was not held constant with respect to time	The rate of death post-alloSCT is expected to vary with respect to time given that the procedure itself is associated with an excess mortality risk in the weeks immediately after transplantation, and that those who survive over the long-term are likely to be “cured” of their disease and hence experience a low mortality risk similar to the general population.
Post-progression survival: patients had the same prognosis (probability of death) at the point of progression irrespective of time to progression	Post-progression survival (PPS) was assumed constant due to both a lack of data to model a time dependent PPS and for simplicity as it removed the need of tracking patients within the Markov model. However, despite this limitation the predicted OS had a good level of face validity when compared to the observed OS from the Cheah study.
Post-progression survival: the same rate of post-progression mortality is applied to all patients independent of treatment assignment	The assumption was very conservative given the current OS rates of [redacted]% from KEYNOTE-087 cohort 1 and 2, respectively versus ~56% in Cheah at approximately 20 months. Despite some potential slight imbalances in the populations this extensive difference in the observed survival cannot be dismissed.
Probability of death (instantaneous hazard): The probability of death in any given cycle is greater than, or equal to, that of age/gender adjusted general population mortality	This control was implemented to avoid the mortality rate falling below that of the general population, as some of the parametric models plateau due to the long term overall survival observed in a portion of this population
AEs were assumed to accrue only once in the modelled time period	Grade 3+ adverse events can potentially lead to treatment discontinuation meaning patients remaining on treatment beyond the first year will be likely to be tolerating treatment well and not experiencing severe adverse events
Any AE relating to subsequent treatment is excluded from the analysis	Avoid over complicating the model for a minimal incremental impact

All-cause adverse events rates are included for pembrolizumab	It was unclear if the reported adverse events were treatment related for all the chemotherapy regimens, therefore it was conservative to assume all-cause incidence rates for pembrolizumab
Treatment-specific HSUVs for PF were conditional on the treatments response profile and are constant with respect to time	In line with the committee preference on a recent Hodgkin's Lymphoma appraisal (TA462) [57]
The HSUV for PD was constant and equivalent across treatments	In line with the committee preference on a recent Hodgkin's Lymphoma appraisal (TA462) [57]
A utility decrement was applied for the first 100 days after alloSCT	Pidala (2009) [81] identified that studies show patients QoL declines immediately following alloSCT and improve to baseline levels after 100 days
HSUV 100 days post alloSCT was conditional on the alloSCT response profile (irrespective of prior treatment or response)	In line with the assumptions for progression-free health states with patients on pembrolizumab and SoC
HSUVs decline over time in line with UK age/gender specific population norms	Given the long time horizon, patients would be expected to have a general decline in QoL.
A one-off disutility was applied in the first cycle of the model to account for the QALY loss from grade 3 or above AE	Grade 3+ adverse events can potentially lead to treatment discontinuation meaning patients remaining on treatment beyond the first year will be likely to be tolerating treatment well and not experiencing severe adverse events
The proportion of investigational agent and bendamustine usage was assumed from Cheah with all other patients being distributed equally between chemotherapy regimens	To allow the treatment costs to accurately reflect the efficacy data. However, the level of investigational agent usage within UK clinical practice is potentially lower (only 25% of clinicians surveyed had access to investigational agents) therefore the assumption should be considered conservative.
Patients discontinued SoC upon progression however, pembrolizumab was modelled based on ToT from KEYNOTE-087 as discontinuation prior to progression was observed	Progression-free survival generally exceeding that of ToT in KEYNOTE-087. Thus, on average, patient's discontinued pembrolizumab prior to progression and PFS is not a suitable proxy of ToT. Patients may discontinue treatment prior to progression due to a sufficient response or toxicity. ToT was not available for SoC.
The model assumed vial wastage (no vial sharing)	Assumption based on previous NICE submissions
There were no additional weekly follow-up costs included beyond the one-off cost of alloSCT	Radford (2017) [28] included all costs related to follow-up beyond the initial procedure.
A terminal cost was applied on death to all patients (except those whom receive alloSCT) to account for the additional intensive disease management prior to death	In line with previous NICE oncology submissions The cost of alloSCT is assumed to already include the cost of terminal care for a number of patients included in the Radford study.
Subsequent treatment costs were applied as a one-off cost on progression consisting of chemotherapy, palliative care and investigational agents	In line with a previous Hodgkin Lymphoma submission (TA462) [57]. There was no double counting with terminal care as palliative care had no cost.
AE associated with alloSCT are assumed to be incorporated in the cost of alloSCT	Radford (2017) [28] accounted for the cost of follow-up which will have captured any additional resource usage incurred from adverse events

AE costs are applied as a one-off cost in the first cycle of the model	Grade 3+ adverse events can potentially lead to treatment discontinuation meaning patients remaining on treatment beyond the first year will be likely to be tolerating treatment well and not experiencing severe adverse events

5.7 Base-case results

5.7.1 Base-case cost effectiveness analysis results

The results of the economic model are presented in Table 101 below. In the base case analysis, the estimated overall survival at 72 months was 28% and 22% for cohorts 1 and 2 respectively with pembrolizumab versus 15% with UK SoC. At the end of the 40-year time horizon there were 7.5% and 7.1% patients still alive in the pembrolizumab cohort and 5.5% and 5.4% in the UK SoC cohort in cohort 1 and 2 respectively. Patients treated with pembrolizumab accrued 4.497 and 4.072 QALYs compared to 3.223 and 3.200 among patients in the UK SoC cohort in cohorts 1 and 2 respectively.

5.7.2 Base-case incremental cost effectiveness analysis results

Table 101 below presents the base case incremental cost-effectiveness results, incorporating the CAA. The results show pembrolizumab to be cost-effective compared to UK SoC when considering a willingness to pay threshold of £50,000 per QALY. The corresponding incremental-cost-effectiveness ratio (ICER) when pembrolizumab is compared to UK SoC was £43,511 and 48,571 for cohorts 1 and 2 respectively. This ICER should be considered in the context of pembrolizumab being an end of life technology that presents an innovative nature (see Section 2.5 and Section 4.13).

Table 101 Base-case results

Technologies	Cohort	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
UK SoC	Cohort 1	52,017	4.864	3.223	-	-	-
	Cohort 2	51,424	4.832	3.200	-	-	-
Pembrolizumab	Cohort 1	107,459	6.252	4.497	55,442	1.274	43,511
	Cohort 2	93,732	5.775	4.072	42,308	0.871	48,571

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

5.7.3 Clinical outcomes from the model

In Table 102 the outcomes for pembrolizumab from KEYNOTE-087 and sources applied for the SoC arm have been compared to the outcomes from the model where possible. The model estimates similar percentages of patients in pre-progression and surviving at different points in time to those reported in KEYNOTE-087 and SoC sources (see Table 102), suggesting that, the model is able to replicate similar results. It should be noted that the comparisons made below are based on outcomes from the model if no alloSCT occurred as KEYNOTE-087 was not designed to allow for alloSCT. This means no alloSCT was assumed in order to derive the comparisons below as comparisons against KEYNOTE-087 post alloSCT would not be possible. The estimated model output comparison of OS after alloSCT at 5 years based on literature can be seen to be robust at 51.28% vs 53.47%.

Table 102: Comparison of model and trial outcomes

		Pembrolizumab		UK SoC	
Outcome		**Base case**	**KEYNOTE-087**	**Base case**	**Cheah et al**
% PFS at 1 Year *	Cohort 1	59.7%	[redacted]	4.1%	~7.5%
	Cohort 2	44.1%	[redacted]	4.9%	
OS at week 12	Cohort 1	98.96%	[redacted]	98.96%	~100%
	Cohort 2	98.78%	[redacted]	98.78%	
OS at 72 Months**	Cohort 1	28%	-	16%	15%
	Cohort 2	22%		16%	
OS after alloSCT 5 years		Base case	KEYNOTE-087	Base case	Lafferty et al
	Cohort 1	51.28%	-	51.28%	53.47%
	Cohort 2				
*using data post week 12 assuming no alloSCT as per KEYNOTE-087 design ** when no alloSCT is assumed as per assumption made about Cheah SoC arm					

Table 103 below show the response rates in the pembrolizumab arm in both cohorts 1 and 2 to be improved versus SoC and hence the levels of alloSCT observed to be increased (Table 104). The outcomes of the increased level of alloSCT can be seen in the improved mean time alive vs SoC in both cohorts 1 and 2 reported in Table 105. It can also be seen that patients who were unable to receive an alloSCT remain progression free longer on pembrolizumab vs SoC.

Table 103: Estimated response rates at 12 weeks

		CR	PR	SD
Pembrolizumab	Cohort 1	15.94%	42.03%	36.89%
	Cohort 2	8.64%	43.21%	38.94%
SoC	Cohort 1	15.19%	18.99%	37.90%
	Cohort 2	15.19%	18.99%	35.90%

Table 104: Estimated overall rate of alloSCT (%)

Pembrolizumab	Cohort 1	43.82%
	Cohort 2	40.05%
SoC	Cohort 1	30.67%
	Cohort 2	30.16%

Table 105: Mean survival times (months)

Mean time PF no alloSCT	Pembrolizumab	Cohort 1	10.0
		Cohort 2	6.38
	SoC	Cohort 1	1.3
		Cohort 2	1.36
Mean time PD no alloSCT	Pembrolizumab	Cohort 1	20.08
		Cohort 2	21.4
	SoC	Cohort 1	24.89
		Cohort 2	25.0
Mean time alive alloSCT	Pembrolizumab	Cohort 1	71.2
		Cohort 2	69.0
	SoC	Cohort 1	49.8
		Cohort 2	49.0

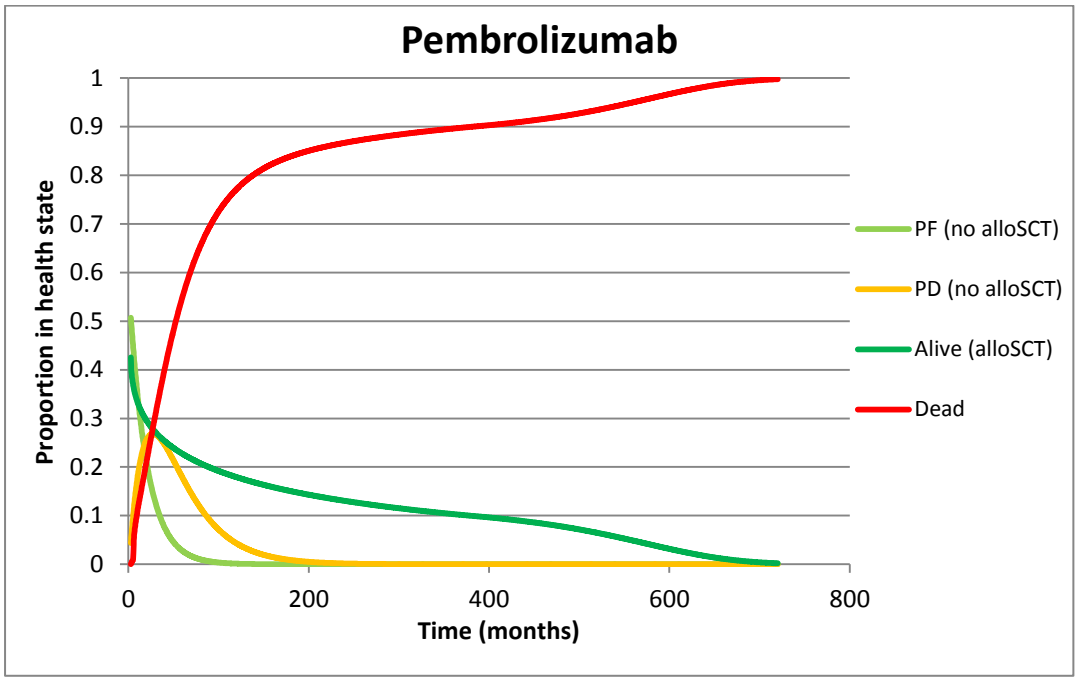
5.7.4 Markov traces

Figure 30 and Figure 31 below illustrates how patients move through the model over time from week 12 when treated with pembrolizumab or UK SoC, respectively. The diagrams show that patients spend longer in pre-progression and also that they survive longer on pembrolizumab compared the UK SoC.

5.7.5 Accrual of costs, QALYs and LYs over time

Figure 32 and Figure 33 shows how the costs, QALYs and life years accumulate over time, respectively. In the base case, QALYs are accrued over time according to the health state occupancy, as previously reported (see sections 5.2.2 and 5.4).

Figure 30: Cohort 1: Markov trace from week 12 for pembrolizumab and UK SOC



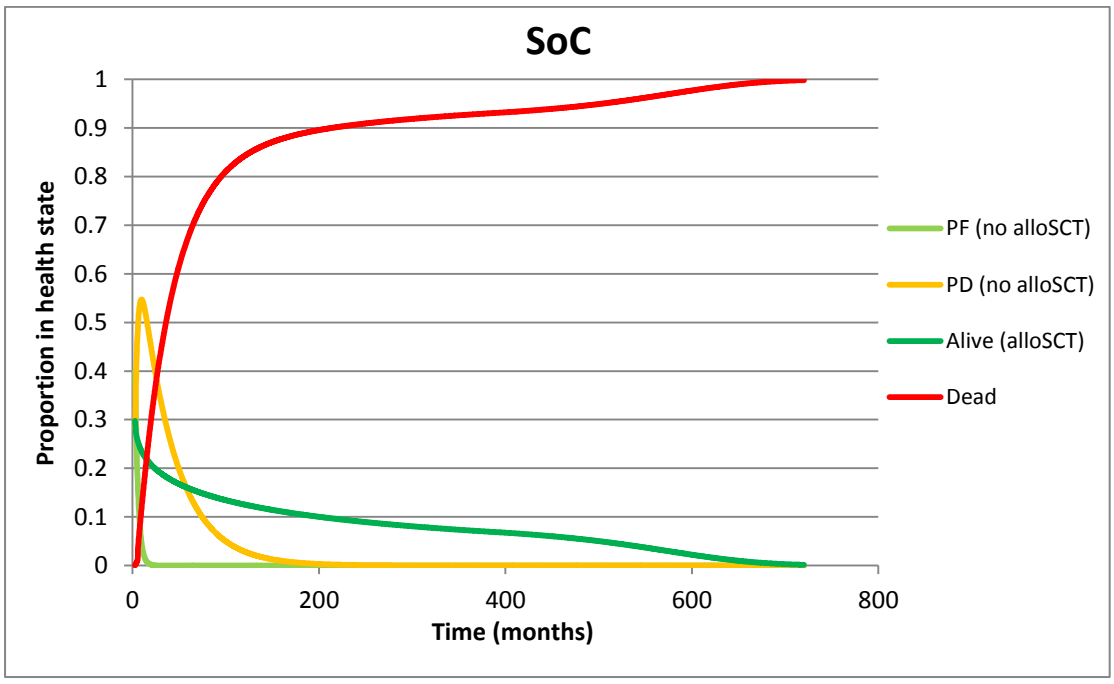
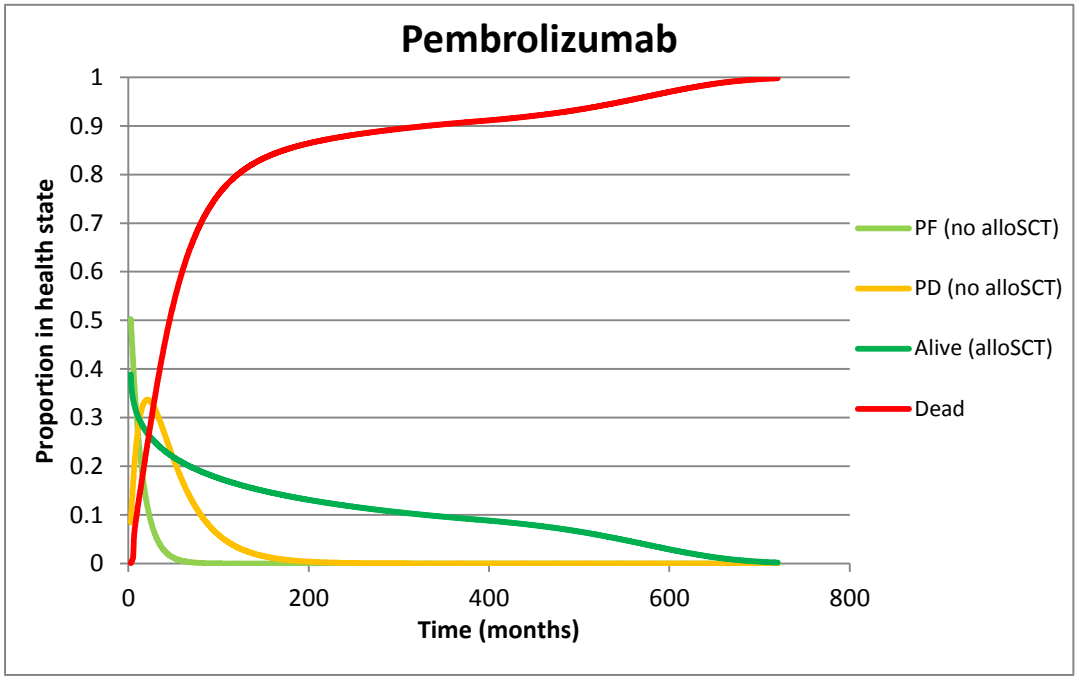


Figure 31: Cohort 2: Markov trace from week 12 for pembrolizumab and UK SoC



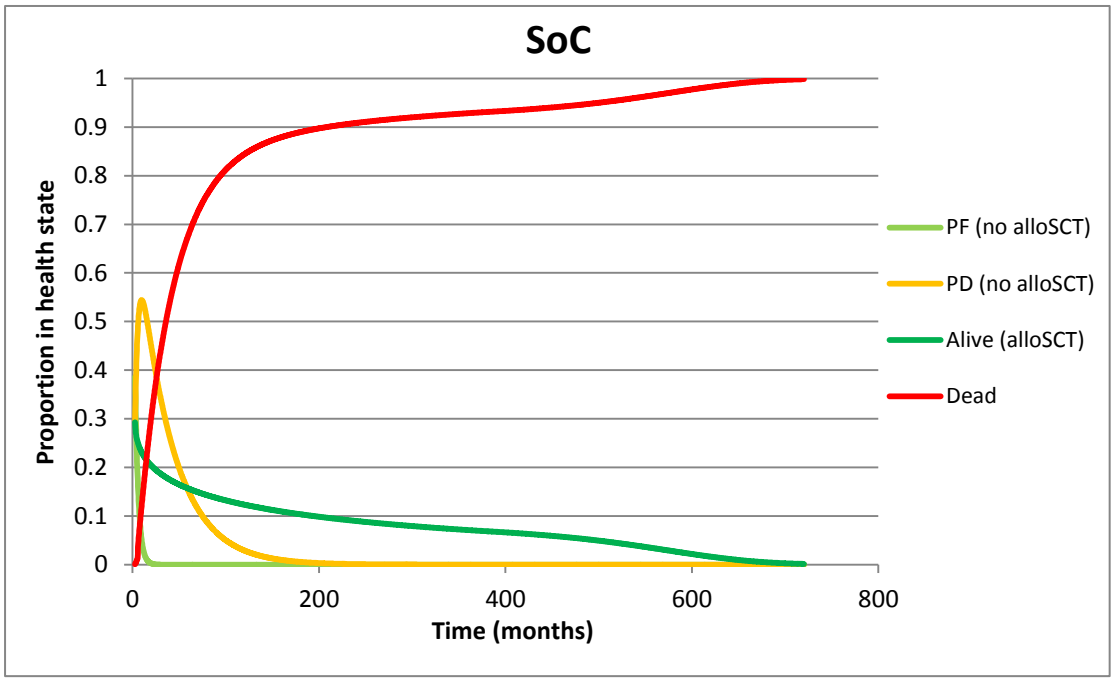
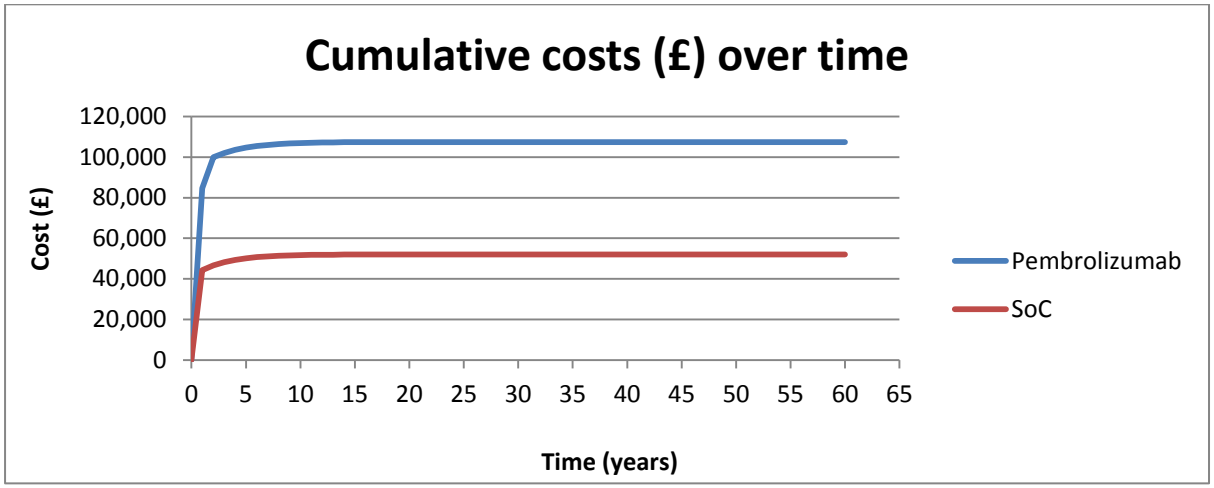


Figure 32: Cohort 1: Cumulative costs, QALYs and LYs over time



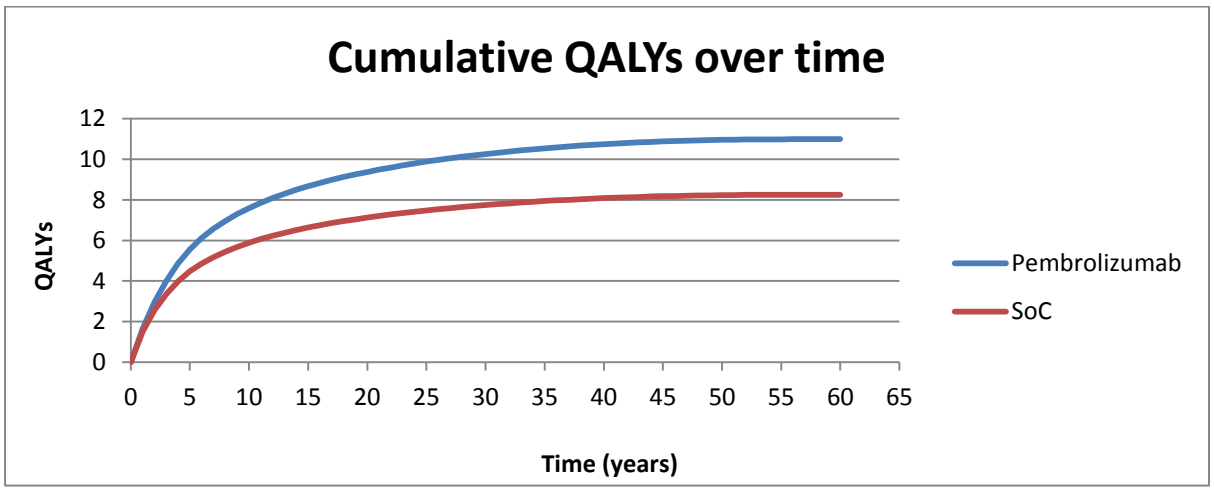
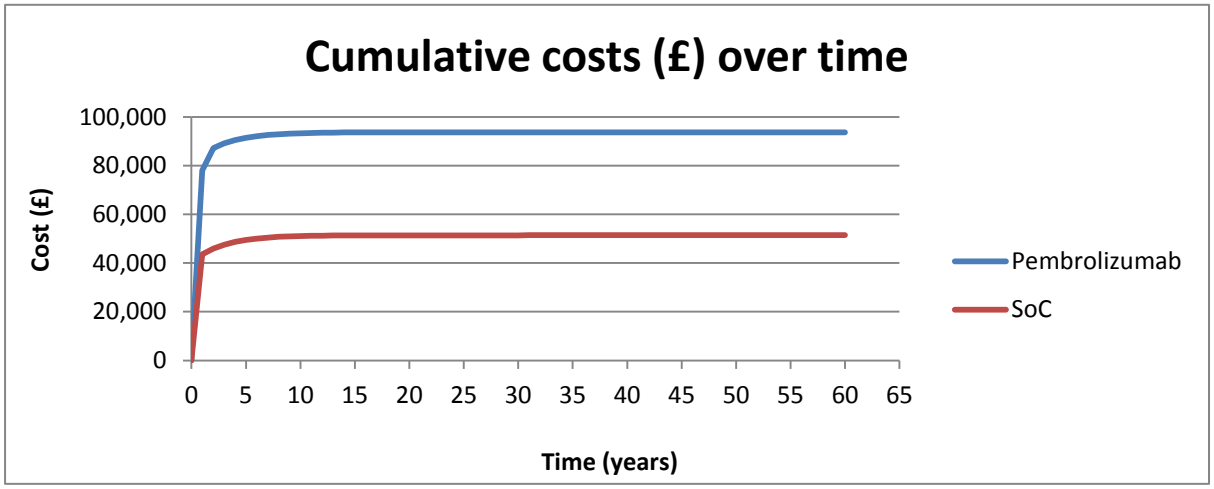
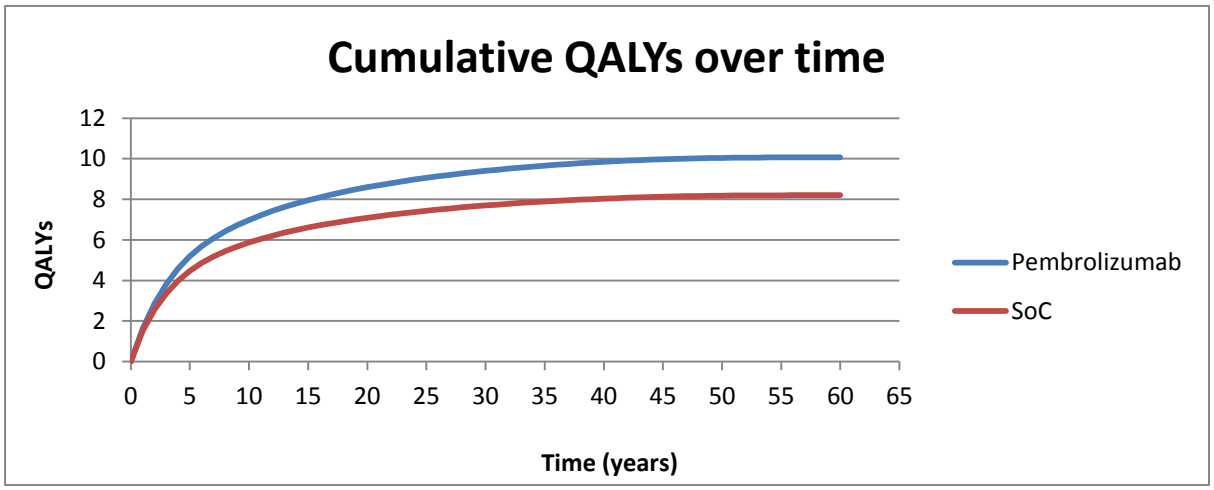


Figure 33: Cohort 2: Cumulative costs, QALYs and LYs over time





5.7.6 Disaggregated results of the base case incremental cost effectiveness analysis

Table 106 shows the disaggregated life years by health state. This shows that patients on pembrolizumab spend longer in both the pre-progression and alive health states compared to patients receiving UK SoC. Table 107 shows that the majority of costs in the pembrolizumab cohort are associated with treatment.

Table 106: Disaggregated life-years by health state (discounted)

	Cohort	Week 0 to week 12		Beyond week 12 (w/o alloSCT)		Beyond week 12 (w alloSCT) Alive
		PF	PD	PF	PD	Alive
Pembrolizumab	Cohort 1	0.226	0.003	0.799	1.455	3.770
	Cohort 2	0.227	0.002	0.519	1.582	3.446
SoC	Cohort 1	0.205	0.024	0.109	1.887	2.639
	Cohort 2	0.222	0.007	0.114	1.895	2.595

Table 107: Summary of predicted resource use by category of cost

	Cohort	Week 0-12				
		Terminal costs (£)	Acq. costs (£)	Admin. costs (£)	Sub. treat cost (£)	AE costs (£)
Pembrolizumab	Cohort 1	40.65	12,456	932.16	184.39	235.94
	Cohort 2	47.71	12,566	940.41	292.65	188.64
SoC	Cohort 1	40.65	946.16	1,130.43	1,225.94	1,945.74
	Cohort 2	47.71	990.60	1,193.28	1,101.72	1,945.74

	Cohort	Week 12 onwards			
		Terminal costs (£)	Acq. costs (£)	Admin. costs (£)	Sub. treat cost (£)
Pembrolizumab	Cohort 1	2,242	29,526	2,209	2,360
	Cohort 2	2,389	20,794	1,556	2,394
SoC	Cohort 1	2,777.25	22.60	85.75	2,003.29
	Cohort 2	2,791	22.15	83.81	1,930

	Cohort	Beyond week 12 (w alloSCT)		
		Disease management costs (£)	Terminal costs (£)	SCT costs (£)
Pembrolizumab	Cohort 1	0.00	0.00	48,363
	Cohort 2	0.00	0.00	44,204
SoC	Cohort 1	0.00	0.00	33,854
	Cohort 2	0.00	0.00	33,287

5.8 Sensitivity analyses

5.8.1 Probabilistic sensitivity analysis

To assess the uncertainty surrounding the variables included in the cost-effectiveness model, a probabilistic sensitivity analysis (PSA) was undertaken using 1,000 samples. The mean values, distributions around the means and sources used to estimate the parameters are detailed in Appendix 16.

Table 108: Incremental cost-effectiveness results based on probabilistic sensitivity analysis (discounted, with PAS)

Technologies	Cohort	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
UK SoC	Cohort 1	£53,491	3.219	-	-	-
	Cohort 2	£54,028	3.254	-	-	-
Pembrolizumab	Cohort 1	£106,702	4.438	£53,211	1.219	£43,653
	Cohort 2	£94,522	4.050	£40,493	0.796	£50,894

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

The incremental cost-effectiveness results obtained from the probabilistic sensitivity analysis are presented in Table 108, and the corresponding scatterplots and cost-effectiveness acceptability curves are presented in Figure 34 to Figure 37. The cost-effectiveness acceptability curve shows that there is an approximately 60% and 50% probability of pembrolizumab to be cost-effective when compared to UK SoC at the £50,000 per QALY threshold for cohort 1 and 2 respectively.

Figure 34: Cohort 1 Scatterplot of PSA results (1,000 simulations; results discounted, with PAS)

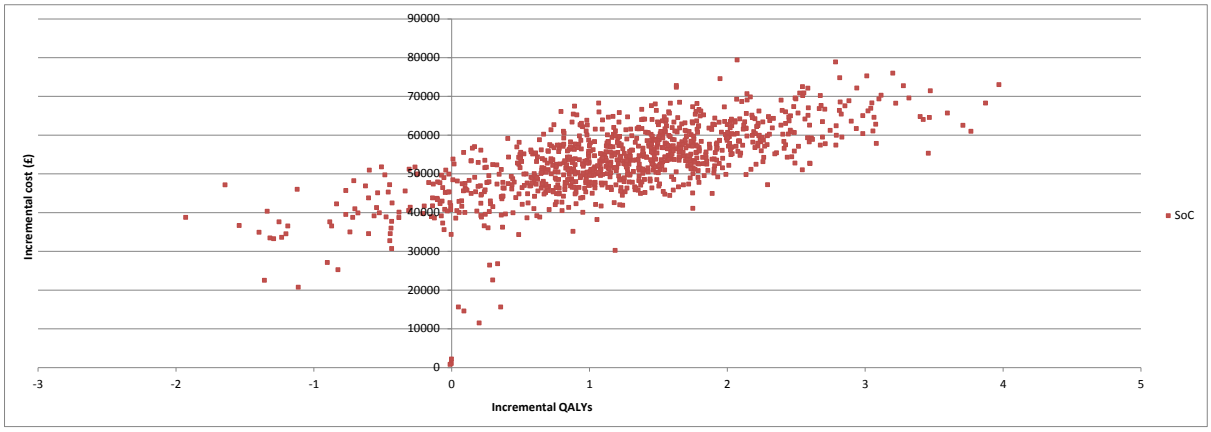


Figure 35: Cohort 2 Scatterplot of PSA results (1,000 simulations; results discounted, with PAS)

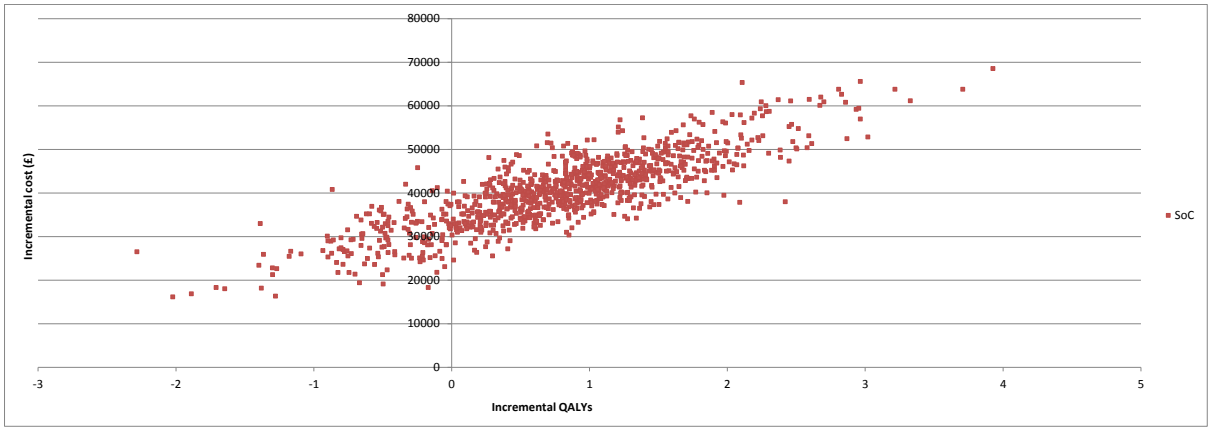


Figure 36: Cohort 1 Cost-effectiveness acceptability curve (results discounted, with PAS)

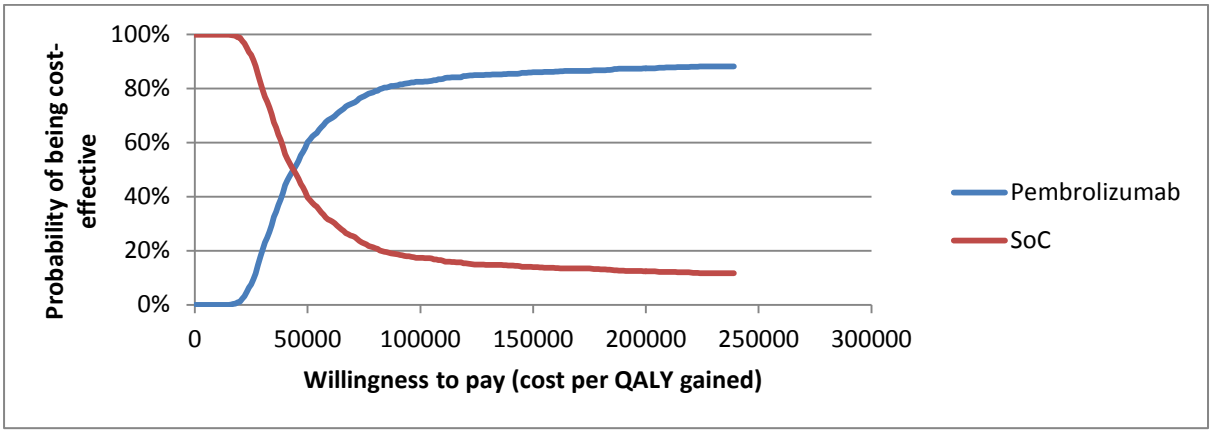
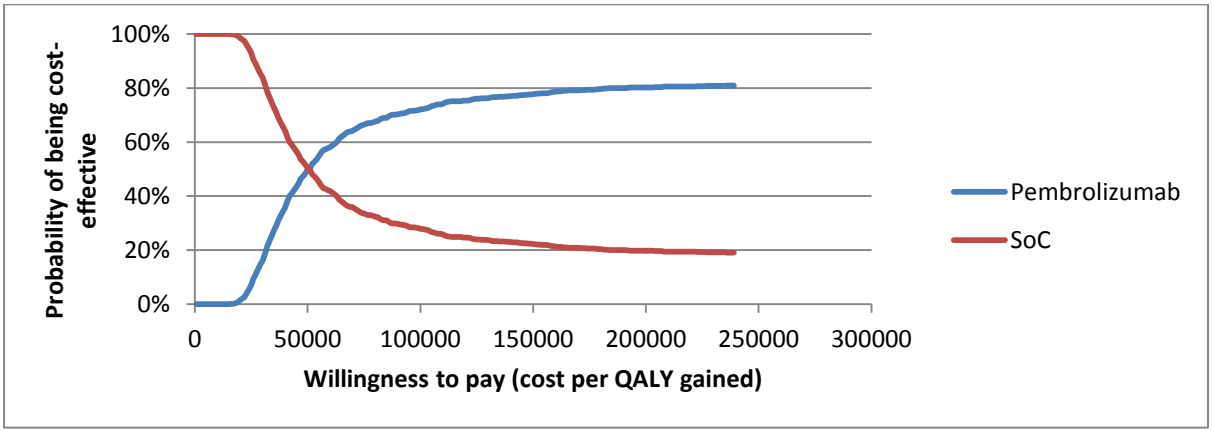


Figure 37: Cohort 2 Cost-effectiveness acceptability curve (results discounted, with PAS)



5.8.2 Deterministic sensitivity analysis

Deterministic sensitivity analyses were conducted for the key variables highlighted in Table 109 using the 5% and 95% confidence intervals for the variables where possible and +/- 10% otherwise.

Table 109: DSA input parameters

Parameter
Discount rate – Costs and Outcomes
Age (mean)
Female (%)
Body surface area (m²)
Health state utility values (PFS) - Pembrolizumab
Health state utility values (PFS) - SoC
Health state utility values (PFS) - Allogeneic SCT (<14 weeks*)
Health state utility values (PFS) - Allogeneic SCT (>14 weeks*)
Health state utility values PD - All treatments
Administration costs - Cost (£) per cycle - Pembrolizumab
Allogeneic stem cell transplant cost
Total progression free cost (£ per week)
Total cost following alloSCT (£ per week)
Total progressed disease cost (£ per week)
Total terminal cost (£)
Adverse event costs - Cost (£) per event
Percentage alloSCT given CR/PR/SD
Response at week 12 - Pembrolizumab – CR/PR
Response at week 12 - SoC – CR/PR odds ratio
Progression-free survival from Week 0 to 12 - SoC HR
Overall survival from Week 0 to 12 - SoC HR
PFS from week 12 given primary treatment beyond week 12 and no alloSCT - SoC HR

The results of the deterministic sensitivity analyses for pairwise comparisons of pembrolizumab vs. UK SoC are presented in Figure 38 below. These are presented with the CAA for pembrolizumab. In the majority scenarios in cohort 1 and cohort 2, the ICER for pembrolizumab vs SoC remained below the £50,000 WTP threshold. The inputs that most affect the ICERs are those related to the discount rate applied to outcomes and odds ratios applied to CR and PR at week 12 (Figure 38 and Figure 39). Plausible alternative scenarios have further been investigated in section 5.8.3, with relatively little impact on the cost effectiveness at a WTP threshold of £50,000.

Figure 38: Cohort 1 Tornado diagram presenting the results of the deterministic sensitivity analysis for the 20 most sensible variables (discounted results, with PAS)

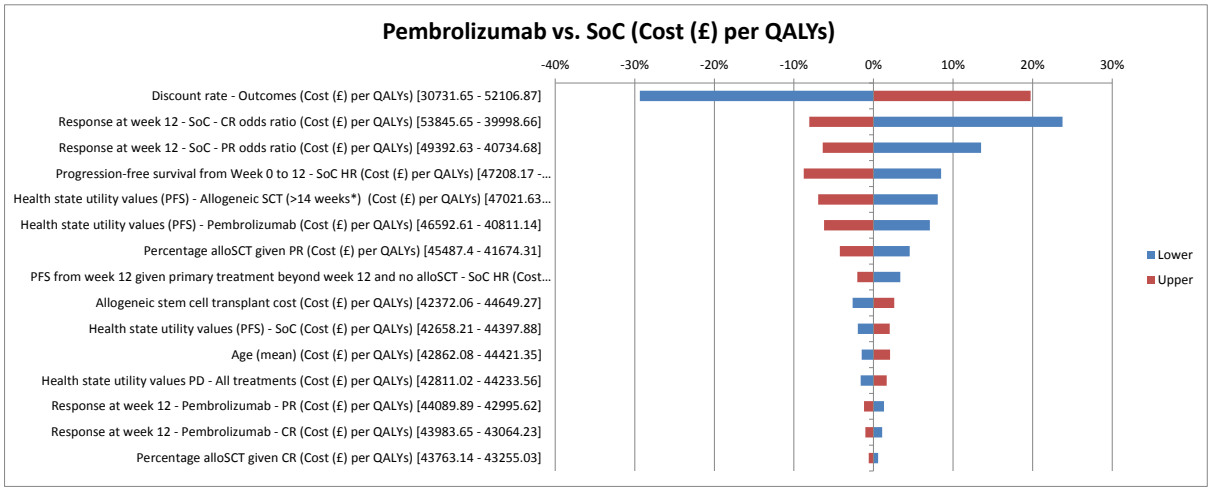
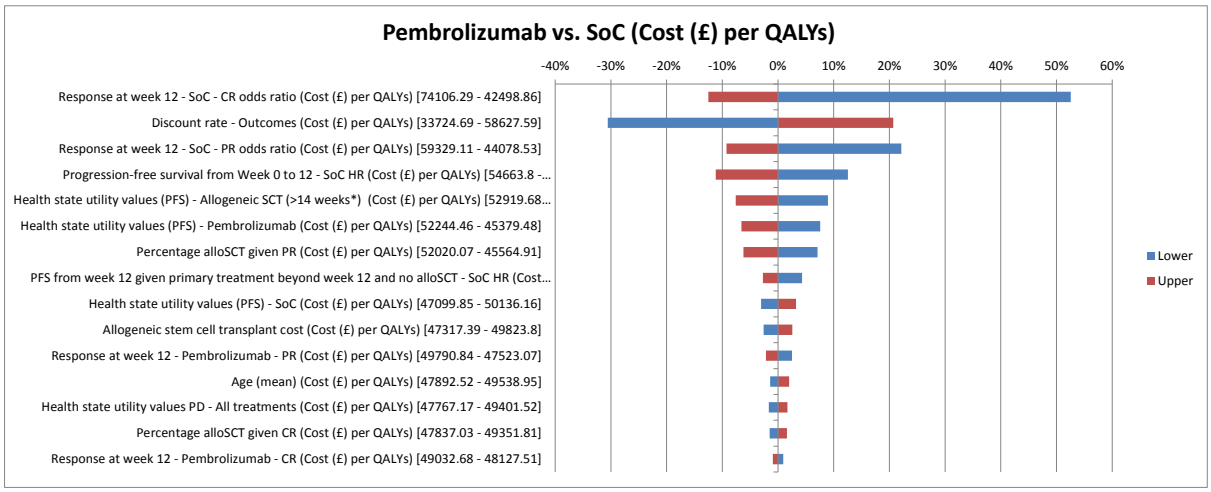


Figure 39: Cohort 2 Tornado diagram presenting the results of the deterministic sensitivity analysis for the 20 most sensible variables (discounted results, with PAS)



5.8.3 Scenario analysis

Alternative scenarios were tested as part of the sensitivity analysis to assess uncertainty regarding structural and methodological assumptions:

- Assessing BSC as a comparator as per the NICE scope [77] (scenario 1)
- Assessing different alloSCT rates (scenario 2)
    - 100% alloSCT in patients with CR, PR or SD (scenario 2a)
    - Alternative lower PR alloSCT rate from MSD clinician survey (scenario 2b)
- Using MAIC HR and OR rather than naïve ITC (scenario 3)
- Alternative extrapolation scenarios to estimate PFS and OS (scenario 4):
    - Considering a Weibull curve for week 0-12 PFS extrapolation in cohort 2 (4a).
    - Considering a Gompertz curve for week 12+ PFS extrapolation in cohort 2 (4b).
    - Considering a Lognormal curve following alloSCT (4c)
- Assessing varying the time horizon to 50 years (scenario 5)

Table 110: Results from the scenario analyses

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case	Cohort 1	£107,459	6.252	4.497	£52,017	4.864	3.223	£55,442	1.274	£43,511
	Cohort 2	£93,732	5.775	4.072	£51,424	4.832	3.200	£42,308	0.871	£48,571
Scenario 1	Cohort 1	£107,459	6.252	4.497	£51,188	4.864	3.223	£56,270	1.274	£44,161
	Cohort 2	£93,732	4.832	3.200	£50,713	4.832	3.200	£43,018	0.871	£49,387
Scenario 2a	Cohort 1	£119,943	8.503	6.768	89,436	7.175	5.474	£30,507	1.295	£23,564
	Cohort 2	£116,185	8.261	6.537	£87,472	7.053	5.364	£28,713	1.172	£24,492
Scenario 2b	Cohort 1	£106,221	6.029	4.272	£49,951	4.736	3.098	£56,270	1.173	£47,957
	Cohort 2	£91,431	5.520	3.819	£49,360	4.705	3.077	£42,070	0.742	£56,677
Scenario 3	Cohort 1	£107,459	6.252	4.497	£45,292	4.419	2.790	£62,166	1.707	£36,423
	Cohort 2	£93,732	5.775	4.072	£46,944	4.558	2.933	£46,787	1.139	£41,087
Scenario 4a	Cohort 2	£93,261	5.766	4.062	£51,234	5.814	3.175	£42,027	0.886	£47,410
Scenario 4b	Cohort 2	£93,439	5.688	4.000	£51,500	4.852	3.217	£41,938	0.783	£52,562
Scenario 4c	Cohort 1	£107,459	6.451	4.642	£52,016	5.003	3.324	£55,442	1.318	£42,075
	Cohort 2	£93,732	5.957	4.204	£51,423	4.969	3.300	£42,308	0.904	£46,812
Scenario 5	Cohort 1	£107,459	6.377	4.582	£52,016	4.951	3.283	£55,442	1.300	£42,651
	Cohort 2	£93,732	5.889	4.150	£51,423	4.918	3.259	£42,308	0.890	£47,516

Summary of sensitivity analyses results

The probability of pembrolizumab being the most cost-effective treatment at a threshold of £50,000 per gained QALY is 60% and 50% for cohort 1 and 2 respectively.

One-way sensitivity analyses showed that the inputs that most affect the ICER are those related to the discount rate applied to outcomes, and the odds ratios applied to CR and PR at week 12. This is to be expected given the potentially long term benefits associated with an increased number of patients being able to undergo alloSCT.

Scenario analysis showed that the cost-effectiveness of pembrolizumab is resilient to the sources of uncertainty assessed, including: selection of varied curves for extrapolation of PFS and OS, different rates of alloSCT applied, varying the time horizon and use of alternative outputs for PFS HR and response rates ORs derived from an MAIC. Applying BSC as the comparator in the model also maintained a stable ICER not dissimilar to the base case.

5.9 Subgroup analysis

5.9.1 Types of subgroups that are not considered relevant

No further subgroup analysis has been considered within this submission.

5.9.2 Analysis of subgroups

Not applicable.

5.9.3 Definition of the characteristics of patients in the subgroup

Not applicable.

5.9.4 Description of how the statistical analysis was carried out

Not applicable.

5.9.5 Results of subgroup analyses

Not applicable.

5.9.6 Identification of any obvious subgroups that were not considered

Not applicable.

5.10 Validation

5.10.1 Methods used to validate and quality assure the model

The outcomes of the pembrolizumab and the UK SoC arms of the KEYNOTE-087 trial have been compared to the outcomes from the model. For more details comparing the results generated from the model to the outcomes from the trial please refer to section 5.7.3.

Expert validation

The model structure, assumptions and rationale was critically reviewed by an independent health economics modelling expert. The methodology for which is provided as a reference[132].

5.11 Interpretation and conclusions of economic evidence

As previously discussed, within the context of relapsed or refractory cHL there are low patient numbers and short survival with the clinical pathway for cHL patients subject to considerable uncertainty and heterogeneity, particularly in the post autoSCT, post BV and autoSCT ineligible setting leading to a paucity of clinical evidence on which to base economic evaluation. In general, where no evidence has been identified, simple assumptions have been made base on independent sources such as British cHL guidelines, published literature and previous NICE appraisals which were then assessed for clinical plausibility. The rationale for each assumption is provided in Section 5.6. Sensitivity analysis around these assumptions were then undertaken, the conclusions from which represent the basis for the economic evaluation shown here.

The current analysis has been designed to be comparable with previous health economic analysis and reflect the most important treatment outcomes for patients in this therapy area in order to allow transparency and ease of review. These outcomes are survival (progression free and overall), response rates, side effects, symptom control and quality of life.

In the base case analysis, it was estimated that pembrolizumab use results in an additional 1.274 and 0.871 discounted QALYs and 1.388 and 0.943 discounted LYs versus SoC in Cohort 1 and 2, respectively.

Through clinician feedback on the UK clinical pathway, referred to throughout this submission, it was apparent that the goal for RRcHL patients in this late line of therapy is to achieve an adequate response to allow them to undergo alloSCT which is associated with long term improved outcomes and survival versus SoC alone. The economic analysis presented here has attempted to capture this clinical pathway and the benefits of pembrolizumab that can be realised versus SoC.

There was an increased level of response across all response rates in the pembrolizumab arm vs SoC allowing a higher level of alloSCT to occur (43.82% and 40.05% vs 30.67% and 30.16 in SoC for cohorts 1 and 2 respectively)

The model estimates that in those who receive an alloSCT, the mean time alive post alloSCT is increased in the pembrolizumab arm vs the SoC arm (71.2 and 69.0 months vs 49.8 and 49.0 months in cohorts 1 and 2 respectively).

The model also estimates that in those who cannot undergo an alloSCT, the benefits of pembrolizumab allow the time in progression free disease to be increased vs SoC (mean 10.0 and 6.38 months vs 1.3 and 1.36 months SoC in cohorts 1 and 2 respectively).

This indicates a considerable benefit to survival in cohorts 1 and 2 vs SoC, both in the ability to remain in progression free disease for longer if alloSCT is not an option and to be more likely to achieve an adequate response to allow an alloSCT associated with higher long term survival.

Incremental costs are expected to be £55,442 and £42,308 under base case assumptions and the resultant ICER was £43,511 and £48,571 for cohorts 1 and 2 respectively, which can be considered cost-effective at a WTP threshold of £50,000/QALY.

Several sensitivity analyses were performed to assess the impact of variation in all variables and assumptions applied within the model. The deterministic analysis and PSA showed pembrolizumab to be cost-effective in the majority of scenarios at a WTP threshold of £50,000/QALY. In addition, alternative inputs and assumptions were assessed as scenario analysis described in Section 5.8.3, with the majority of these ICERs remaining below £50,000/QALY.

Application of NICE end of life criteria to pembrolizumab in cHL

End of life criteria as applied by NICE are summarized as follows:

- 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 and extension to life.

The evidence presented in this submission highlight the paucity of data relevant to this small patient group. Furthermore, the results of the clinician survey (Section 4.11.1), and the recent stakeholder feedback, as per the committee meeting papers of ID972, suggest there is a substantial unmet need for patients with RRcHL who have failed/ considered ineligible for

autoSCT and subsequent BV treatment [75]. The case for end-of-life criteria is reported in Table 51 (section 4).

Recently, the NICE appraisal committee for nivolumab (TA462) concluded that within this indication (relevant to the current decision problem) it was plausible that the criteria for short life expectancy could apply [57]. Therefore, EoL criteria were factored into its decision making[57].

There is a general consensus that treatment options available at this later line of therapy, i.e. among those patients with RRcHL is limited and associated with poor outcomes. Therapies currently available for patients with RRcHL are associated with poor outcomes, although as explained there is a lack of available evidence in this area. Patients with RRcHL following ASCT had a median OS of up to 26.1 months, depending on therapy received and reduced to 17.1 months with novel agents excluded [9]. This further decreases for patients who do not achieve an initial response following ASCT [9]. Evidence describing the survival of patients who have RRcHL and who are also ineligible for ASCT is even more scarce and so we can assume that their OS would be expected to be lower given the likelihood that this patient population are most likely to be in the older disease peak period and have worse outcomes overall.

These outcomes are supported by a UK clinician survey (n=16). Clinicians reported that only a minority of patients with RRcHL who have either failed/ considered ineligible for autoSCT and have failed subsequent BV experience ORR, 35% and 31%, respectively. Within the same two patients groups clinicians reported that current SoC followed by alloSCT provided a median OS of 18.9 months and 14.2 months, respectively.

As mentioned above, recently NICE considered (ID 972) the relevance of Cheah et al 2016, which reported OS estimates of around two years. The appraisal committee agreed that although these data were not exactly generalisable to the UK setting, they were suitable for decision making [57]. However, this estimate of OS was skewed by the inclusion of investigational agents (47.4 months), when investigation agents are removed the median estimate of OS reduces to around 19 months [44]. Again, these estimates are broadly comparable to the findings reported above.

Median OS was [redacted] during KEYNOTE-087, but the small number of deaths occurring during this study indicates a substantially longer median survival than that offered by current therapies available to patients on the NHS. However, the small number of deaths reported during the current follow-up period (15.9 month) indicates a substantially longer median survival than that offered by current therapies. The OS rate at 15 months in cohort 1 and 2 was reported using Kaplan-Meier estimates at [redacted] respectively [45, 46].

Based on the available evidence, it can be considered that pembrolizumab meets both criteria for end of life, as specified by NICE.

5.11.1 Comparison with published economic literature

This is the first economic evaluation focused on assessing the cost-effectiveness of pembrolizumab for the treatment of patients with RRcHL. The economic evaluation reflects patients assessed in KEYNOTE-087 and is relevant to all groups of patients who could potentially benefit from use of the technology, as identified in the decision problem.

No study assessing the cost-effectiveness of pembrolizumab for the target population specified above was identified from the systematic literature review. It was therefore not possible to compare the results of the economic model developed in this submission with any available publication.

5.11.2 Relevance of the economic evaluation for all patient groups

The population included in the economic evaluation was consistent with the RRcHL eligible for pembrolizumab as per the licence. As mentioned previously, the KEYNOTE-087 trial, which assessed patients in line with the anticipated licenced indication, was used in the model. Therefore, the economic evaluation is relevant to all patients who could potentially use pembrolizumab within this licence.

5.11.3 Generalisability of the analysis to the clinical practice in England

The analysis is directly applicable to clinical practice in England since:

- The patient population in KEYNOTE-087 and the de novo economic evaluation are reflective of patients with RRcHL in the UK.
- The economic model structure is consistent with other oncology models submitted to NICE and has been validated with UK clinicians as representative of UK clinical practice.
- The resource utilisation and unit costs are reflective of UK clinical practice and were mainly derived from the NHS Reference Costs and previous NICE submissions, incorporating the feedback provided by the ERGs in recent NICE appraisals. These cost inputs are considered most appropriate to model the cost-effectiveness of pembrolizumab.

- Sensitivity analyses were conducted, considering alternative approaches to extrapolation and different data sources and scenarios related to the estimation of QALYs and costs.

- The OS projections of the model were validated with UK clinicians to ensure the clinical plausibility of the model and its applicability to UK clinical practice.

5.11.4 Strengths and weaknesses of the evaluation

The cost-effectiveness analysis makes use of the best available evidence to inform the model.

- Survival where available (PFS and OS) and response rates: Data from the KEYNOTE-087 trial was used in the economic evaluation for the pembrolizumab arm and adjusted for the SoC arm.

- SoC source: Cheah et al was used to derive comparative efficacy estimates using a naïve indirect treatment comparison vs KEYNOTE-087 pembrolizumab data. The study was chosen through a clinical SLR (see section 4) which was conducted and has also been validated for use in previous recent NICE TAs in this therapy area.

- With both naïve comparison and MAIC, there are a number limitations that may lead to uncertainty within the results. This is fully discussed in Section 4.10.14 - 4.10.16

- Survival post alloSCT: Data from best available literature sources was utilized validated through its use in recent NICE TAs.

- Estimation of utilities: Utility values were obtained from EQ-5D KEYNOTE-087 data applied to pembrolizumab and SoC arms. In order to fit with the model structure, utility at week 12 for PD did not show the utility decrement which would be expected hence the best available utility decrement available from literature was used (detailed in section 5.4.7).

- Treatment duration of pembrolizumab: The model assumed that patients will be treated for up to 24 months, i.e. 35 cycles, as defined as part of the KEYNOTE-087 protocol.

- Resource utilisation and unit costs used in the analysis are reflective of UK clinical practice.

Sensitivity analyses were conducted to inform the uncertainty around the above, which helped in understanding what key variables could potentially have a major impact on the cost-effectiveness results.

Since the approaches taken for modelling are, mostly conservative, the results presented here support the conclusion that, within the context of innovative end-of-life therapies, pembrolizumab is a cost-effective therapeutic option for the treatment of patients with RRcHL.

5.11.5 Further analyses

Ongoing studies are reported in Section 4.14

6.0 Assessment of factors relevant to the NHS and other parties

6.1 Analysis of any factors relevant to the NHS and other parties that may fall outside the remit of the assessments of clinical and cost effectiveness

There are no further factors relating to the decision problem which are relevant to the NHS but fall outside of the remit of the assessment.

6.2 Eligible population

In total, 106 patients with RRcHL are estimated to be eligible for treatment with pembrolizumab in each year (see Table 111 below). The steps followed to estimate these values are described below. The budget impact here assumes that patients treated with pembrolizumab receive the anticipated licensed dose of 200mg for an average of 13.3 cycles based on the economic model estimates. This estimated time on treatment duration is based on the maximum duration of treatment of pembrolizumab of 24 months and has been extrapolated using parametric modelling. Based on this and the expected short life expectancy of patients with RRcHL and short treatment time, the budget impact is based on the yearly incidence only.

Patients who have failed autoSCT and BV: Cohort 1

During 2015, 137 [133] autoSCT procedures were carried out in UK patients with HL. With 50% [32] of autoSCT patients estimated to fail to respond or relapsed and scaled to the population of England [134] this makes 58 patients. Of these 58 patients, 83% fail to respond to BV (less than CR) [135] and so these patients can be considered to be most likely to require treatment with pembrolizumab. This would equate to 48 patients eligible for pembrolizumab each year or 239 patients over 5 years summarized in Table 111.

Patients ineligible for autoSCT: Cohort 2

The incidence of HL in England is around 1790 new cases, equivalent to around 3.3 cases per 100,000 people [136]. Of these 15% [137] are thought to be relapsed/refractory to first line treatment and 30% [32] of those ineligible for an autoSCT. The rate of BV failure as in patients who are ineligible for an autoSCT (less than CR) is higher in this group at 83% [135], the number of patients ineligible for autoSCT and eligible for pembrolizumab equates to 67 each year or 334 over 5 years summarized in Table 111.

Table 111 Expected number of treatment eligible cases of cHL over five years

	Year 1	Year 2	Year 3	Year 4	Year 5
Patients who have failed autoSCT and BV: cohort 1					
UK (based on 2014)	137 [133]				
% Patients failing autoSCT	50% [32]				
Patients failing autoSCT	69				
% England	84% [134]				
Number of patients failing autoSCT (England)	58				
Patients failing BV*	83% [135]				
Estimated eligible cohort 1 treated	48	48	48	48	48
Patients ineligible for ASCT: cohort 2					
Incidence of HL England	1790[1]				
Fail 1st line treatment	15% [137]				
Ineligible for autoSCT	30% [32]				
Patients failing BV*	83% [135]				
Estimated eligible cohort 2 treated each year	67	67	67	67	67
Total estimated eligible cohort treated	**115**	**115**	**115**	**115**	**115**

autoSCT: autologous stem cell transplant, BV: brentuximab vedotin

*Assuming same rate of BV failure as in patients who have had an autoSCT

6.3 Current treatment options and uptake assumptions

In the context of RRcHL, with low patient numbers and short survival, the clinical pathway for HL patients is subject to some uncertainty and heterogeneity, particularly in the post ASCT and post BV setting. In light of this uncertainly and the lack of data surrounding comparator composition, the base case analysis assumes that established clinical management is comprised of standard chemotherapy (as per BCSH guidelines), bendamustine and investigational agents as per the economic analysis presented and validated with UK clinicians.

6.4 Assumed market share

Market share assumptions are detailed in Table 112 and are assumed to be in line with that in the economic analysis; for the purpose of the analysis, current chemotherapies are assumed to have equal market share. Pembrolizumab is assumed to replace SoC once introduced. Hence, a market share of 100% has been applied to provide a conservative estimation of the potential budget impact. Market share is assumed to remain constant over the 5 years.

Table 112 Standard of care: market share assumptions

	Percentage
Chemotherapy	38.40%
Bendamustine	18.50%
Investigational agents	43.10%

6.5 Other significant costs associated with treatment that may be of interest to commissioners

Technology costs, administration costs, AE costs and the cost of alloSCT associated with treatment with pembrolizumab and SoC are identical to those assumed in the cost-effectiveness model and are described in section 5.5.

6.6 Unit costs assumed and how they were calculated

All unit costs considered here estimate the annual budget to the NHS in England and are based upon the ones included in the economic analysis in section 5.5. The budget impact below for pembrolizumab and SoC is based on the economic model output for per patient costs (technology, administration, AEs and alloSCT).

6.7 Resource savings and other significant costs

In clinical practice, there may be cost savings associated with pembrolizumab therapy due to the simplified administration schedule. However, in order to provide a robust, conservative analysis, it is assumed that there are no significant savings associated with the use of pembrolizumab in RRcHL.

6.8 Estimated budget impact

Based on assumptions surrounding the number of patients eligible for the treatment, market share and uptake, the estimated budget impact to the NHS over the next 5 years associated with the use of pembrolizumab in this setting in the licensed patient population of cohort 1, cohort 2 and combined is shown in Table 113, Table 114 and Table 115 respectively.

Table 113 Cohort 1 Budget impact estimation (world with/world without)

	Year 1	Year 2	Year 3	Year 4	Year 5
Cohort 1 patients	48	48	48	48	48
World without pembrolizumab					
Total treatment costs	£46,525	£46,525	£46,525	£46,525	£46,525
Total administration costs	£59,007	£59,007	£59,007	£59,007	£59,007
Total adverse event costs	£39,548	£39,548	£39,548	£39,548	£39,548
Total alloSCT	£1,168,193	£1,168,193	£1,168,193	£1,168,193	£1,168,193
Total world without	£1,313,273	£1,313,273	£1,313,273	£1,313,273	£1,313,273
World with pembrolizumab					
Total treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total administration costs	£155,077	£155,077	£155,077	£155,077	£155,077
Total adverse event costs	£11,295	£11,295	£11,295	£11,295	£11,295
Total alloSCT	£2,196,151	£2,196,151	£2,196,151	£2,196,151	£2,196,151
Total world with	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference between the world with and the world without pembrolizumab					
Total treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total administration costs	£96,070	£96,070	£96,070	£96,070	£96,070
Total adverse event costs	(£28,253)	(£28,253)	(£28,253)	(£28,253)	(£28,253)
Total alloSCT	£1,027,958	£1,027,958	£1,027,958	£1,027,958	£1,027,958
Total budget impact	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 114 Cohort 2 Budget impact estimation (world with/world without)

	Year 1	Year 2	Year 3	Year 4	Year 5
Cohort 2 patients	67	67	67	67	67
World without pembrolizumab					
Total treatment costs	£67,933	£67,933	£67,933	£67,933	£67,933
Total administration costs	£86,555	£86,555	£86,555	£86,555	£86,555
Total adverse event costs	£55,232	£55,232	£55,232	£55,232	£55,232
Total alloSCT	£1,559,612	£1,559,612	£1,559,612	£1,559,612	£1,559,612
Total world without	£1,769,332	£1,769,332	£1,769,332	£1,769,332	£1,769,332
World with pembrolizumab					
Total treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Total administration costs	£174,481	£174,481	£174,481	£174,481	£174,481
Total adverse event costs	£12,612	£12,612	£12,612	£12,612	£12,612
Total alloSCT	£2,683,157	£2,683,157	£2,683,157	£2,683,157	£2,683,157
Total world with	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference between the world with and the world without pembrolizumab					
Total treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total administration costs	£87,927	£87,927	£87,927	£87,927	£87,927
Total adverse event costs	(£42,620)	(£42,620)	(£42,620)	(£42,620)	(£42,620)
Total alloSCT	£1,123,544	£1,123,544	£1,123,544	£1,123,544	£1,123,544
Total budget impact	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 115 Estimated budget impact of licensed patient populations cohort 1 and cohort 2 combined

Combined cohort 1 and 2	Year 1	Year 2	Year 3	Year 4	Year 5
Total RRcHL	115	115	115	115	115
Total world without pembrolizumab	£3,082,605	£3,082,605	£3,082,605	£3,082,605	£3,082,605
Total world with pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total budget impact	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

6.9 Estimates of resource savings

See section 6.1.

6.10 Highlight the main limitations within the budget impact analysis.

A number of assumptions were made in terms of proportion of patients treated in RRcHL, which may introduce uncertainty into the estimates of budget impact presented here. The costs assumed here are based on the time on treatment curve estimated from the economic model based on maximum treatment duration of pembrolizumab of 24 months. Additionally, the model is based on a closed cohort of patients based on the eligible population presented in Table 111. In total, 106 patients with RRcHL are estimated to be eligible for treatment with pembrolizumab in each year (In Table 111). The steps followed to estimate these values are described below. The budget impact here assumes that patients treated with pembrolizumab receive the anticipated licensed dose of 200 mg for an average of 13.3 cycles based on the economic model estimates. This estimated time on treatment duration is based on the maximum duration of treatment of pembrolizumab of 24 months and has been extrapolated using parametric modelling. Based on this and the expected short life expectancy of patients with RRcHL and short treatment time, the budget impact is based on the yearly incidence only.

References

1. UK Cancer Research. Hodgkin lymphoma incidence statistics, 2014. Published online 2016. Accessed August 2017
2. Moskowitz AJ PM-A, Kewalramani T, 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;2(146):1365-2141.
3. Graham P. Collins ANP, Christopher Pocock, Irfan Kayani, Anna Sureda, Tim Illidge, Kirit Ardeshna,David C. Linch, Karl S. Peggs. Guideline on the management of primary resistant and relapsed classical Hodgkin lymphoma. BJH, British Journal of Haematology. 2014;164(1):39-52.
4. Food Drugs Administration. U.S. Food and Drug Administration granted accelerated approval to pembrolizumab (KEYTRUDA), Merck and Co., Inc.) for the treatment of adult and pediatric patients with refractory classical Hodgkin lymphoma (cHL), or those who have relapsed after three or more prior lines of therapy2017. Available from: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm546893.htm
5. Robert C RA, Wolchok JD, Hodi FS, Hamid O, Kefford R, et al. . Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a phase 1 trial. Lancet. 2014:1109-17.
6. DM. P. The blockade of immune checkpoints in cancer immunotherapy. NatRevCancer. 2012;12(4):252-64.
7. Merck & Co. Pembrolizumab Summary of Product Characteristics (31 August 2017). http://www.medicines.org.uk/emc/medicine/30602: 2017.
8. Martinez C CC, Sarina B, et al. Identification of prognostic factors predicting outcome in Hodgkin's lymphoma patients relapsing after autologous stem cell transplantation. Ann Oncol. 2013;24(9):2430-4.
9. Bair SM SL, Nasta S, Nagle SJ, Landsburg DJ,. Outcomes of patients with relapsed/refractory Hodgkin lymphoma progressing after autologous stem cell transplant in the current era of novel therapeutics: A retrospective analysis Am J Hematol. 2017:1-6.
10. Food Drugs Administration. Pembrolizumab FDA accelerated approval. 2014.
11. Food Drugs Administration. Pembrolizumab (KEYTRUDA) Checkpoint Inhibitor2016. Available from: http://wayback.archive-it.org/7993/20170111231548/http://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm526430.htm.
12. Food Drugs Administration. Hematology/Oncology (Cancer) Approvals & Safety Notifications. 2016.
13. Food Drugs Administration. Pembrolizumab (KEYTRUDA) August 5th 2016 treatment of patients with recurrent or metastatic head and neck squamous cell carcinoma (HNSCC) with disease progression on or after platinum-containing chemotherapy2016. Available from: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm515627.htm.
14. Merck & Co. FDA Approves Merck's KEYTRUDA® (pembrolizumab) as First-Line Combination Therapy with Pemetrexed and Carboplatin for Patients with Metastatic Nonsquamous Non-Small Cell Lung Cancer (NSCLC), Irrespective of PD-L1 Expression2017. Available from: http://www.mrknewsroom.com/news-release/prescription-medicine-news/fda-approves-mercks-keytruda-pembrolizumab-first-line-combin.

15. Merck & Co. FDA Accepts Two sBLAs for Merck's KEYTRUDA® (pembrolizumab) for Locally Advanced or Metastatic Urothelial Cancer in Cisplatin-Ineligible First-Line and Second-Line Post-Platinum Failure Treatment Settings2017. Available from: http://www.mrknewsroom.com/news-release/oncology-newsroom/fda-accepts-two-sblas-mercks-keytruda-pembrolizumab-locally-advanced-.
16. Food Drugs Administration. U.S. Food and Drug Administration granted accelerated approval to pembrolizumab (KEYTRUDA®, Merck and Co., Inc.) in combination with pemetrexed and carboplatin for the treatment of patients with previously untreated metastatic non-squamous non-small cell lung cancer (NSCLC)2017. Available from: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm558048.htm.
17. Medicines, Healthcare Products Regulatory A. Pembrolizumab Melanoma Early Access to Medicines Scientific Opinion - Public Assessment Report. 2015.
18. Medicines, Healthcare Products Regulatory A. Pembrolizumab NSCLC Early Access to Medicines Scientific Opinion - Public Assessment Report. 2016.
19. Carol S. Portlock. Hodgkin Lymphoma (Hodgkin's Disease)2012. Available from: http://www.merckmanuals.com/professional/hematology-and-oncology/lymphomas/hodgkin-lymphoma.
20. The American Cancer Society . Treating Hodgkin Lymphoma General treatment information, (Chemotherapy for Hodgkin Lymphoma) 2017. Available from: https://www.cancer.org/cancer/hodgkin-lymphoma/treating/chemotherapy.html.
21. The American Cancer Society. What Is Hodgkin Lymphoma? 2017. Available from: https://www.cancer.org/cancer/hodgkin-lymphoma/about/what-is-hodgkin-disease.html.
22. Batlevi CLaay, et al. Novel therapy for Hodgkin lymphoma. Hematology Am Soc Hematol Educ Program. 2013:394-9.
23. The American Cancer Society. Signs and Symptoms of Hodgkin Lymphoma 2017. Available from: https://www.cancer.org/cancer/hodgkin-lymphoma/detection-diagnosis-staging/signs-and-symptoms.html.
24. Lymphoma Association. Symptoms of lymphoma in detail 2017. Available from: https://www.lymphomas.org.uk/printpdf/2619.
25. The American Cancer Society. Hodgkin Lymphoma Stages 2017. Available from: https://www.cancer.org/cancer/hodgkin-lymphoma/detection-diagnosis-staging/staging.html.
26. Paul P. Carbone HSK, Karl Musshoff, David W. Smithers and Maurice Tubiana. Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res. 1971;31(11):1860-1.
27. UK CR. Hodgkin lymphoma survival statistics (2008), Accessed 2017. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma/survival#heading-Three.
28. 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 Transplant. 2017;52(3):452-4.
29. Yabroff KR BC, Mariotto AB, Brown ML, Feuer EJ. Estimates and Projections of Value of Life Lost From Cancer Deaths in the United States. . JNCI Journal of the National Cancer Institute. 2008;100(24):1755-62.
30. Bradley CJ YK, Dahman B, Feuer EJ, Mariotto A, Brown ML. Productivity Costs of Cancer Mortality in the United States: 2000–2020. JNCI Journal of the National Cancer Institute. 2008;100(24):1763-70.
31. George A. Follows KMA, Sally F. Barrington, Dominic J. Culligan, Peter J. Hoskin, David Linch, Shalal Sadullah, Michael V. Williams, Jennifer Z. Wimperis. Guidelines for the first line management of classical Hodgkin lymphoma. BJH, British Journal of Haematology. 2014;166(1):34-49.

32. NICE. Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma TA446. https://www.nice.org.uk/guidance/ta446/resources/brentuximab-vedotin-for-treating-cd30positive-hodgkin-lymphoma-pdf-82604787983557: 2017.
33. Karyn A. Goodman ER, Victoria Serrano, Subhash Gulati, Craig H. Moskowitz, and Joachim Yahalom. Long-Term Effects of High-Dose Chemotherapy and Radiation for Relapsed and Refractory Hodgkin's Lymphoma. Journal of Clinical Oncology. 2008;26(32):5240-7.
34. Saskatchewan Cancer Agency. Provincial Hodgkin's Lymphoma Treatment Guidelines2012. Available from: http://www.saskcancer.ca/Hodgkin's%20Lymphoma%20v3.
35. D. A. Eichenauer AE, M. André, M. Federico, T. Illidge, M. Hutchings, & M. Ladetto on behalf of the ESMO Guidelines Working Group. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2014;25(3).
36. Scottish Medicines Consortium. brentuximab vedotin (Adcetris®) 50mg powder for concentrate for solution for infusion SMC No. (845/12)2014. Available from: https://www.scottishmedicines.org.uk/SMC_Advice/Advice/845_12_brentuximab_vedotin_Adcetris/brentuximab_vedotin_Adcetris_Full.
37. All Wales Medices Strategy Group. Brentuximab vedotin (Adcetris®) for the treatment of adult patients with relapsed or refractory CD30+ Hodgkin lymphoma following autologous stem cell transplant (ASCT) or following at least two prior therapies when ASCT or multi-agent chemotherapy is not a treatment option (Superseded by NICE TA446). 2015
38. Chen R ZP, Fanale M, et al. Phase II Study of the Efficacy and Safety of Pembrolizumab for Relapsed/Refractory Classic Hodgkin Lymphoma. J Clin Epidemiol 2017:2072.1316.
39. Armand P, Shipp MA, Ribrag V, Michot JM, Zinzani PL, Kuruvilla J, et al. Programmed Death-1 Blockade With Pembrolizumab in Patients With Classical Hodgkin Lymphoma After Brentuximab Vedotin Failure. J Clin Oncol. 2016.
40. MSD UK Ltd. Classical Hodgkin Lymphoma Advisory Board Meeting The King's Fund, London Monday 13 March 2017. MSD data on file: 2017.
41. London Cancer Alliance (Wesy and South). LCA Haemato-Oncology Clinical Guidelines, Lymphoid Malignancies. 2015.
42. Chen R ZP, Fanale M, et al. . Pembrolizumab For Relapsed/Refractory Classical Hodgkin Lymphoma (R/R CHL): Phase 2 KEYNOTE-087 Study. J Clin Oncol 2016;34:(suppl; abstr 7555).
43. Moskowitz CH ZP, Fanale MA, et al. Pembrolizumab in relapsed/refractory classical hodgkin lymphoma: Primary end point analysis of the phase 2 keynote-087 study. Blood. 2016;22:128.
44. Cheah CY CD, Horowitz S, et al. Patients with classical Hodgkin lymphoma experiencing disease progression after treatment with brentuximab vedotin have poor outcomes. Ann Oncol. 2016;27(7):1317-23.
45. Merck & Co.. Keytruda (MK-3475) relapsed or refractory classical hodgkin lymphoma HTA submission UK Report (July 2017). MSD data on file: 2017.
46. Merck & Co. KN-087 Efficacy Update Report (June 2017), A Phase II Clinical Trial of MK-3475 (Pembrolizumab in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (Database lock 21 March 2017). MSD Data on file: 2017.
47. Merck & Co. KEYNOTE-087/ KEYNOTE-013 Summary of Clinical Effiacy, Update Report for FDA (Datalock September 2016). MSD data on file: 2016.
48. Merck & Co. CSR: A Phase II Clinical Trial of MK-3475 (Pembrolizumab) in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL) (Database lock June 2016). MSD data on file: 2016.

49. Merck & Co. Protocol: A Phase II Clinical Trial of MK-3475 (Pembrolizumab) in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL). Data on file: 2015.
50. Cheson BD PB, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised Response Criteria for Malignant Lymphoma. J Clin Oncol. 2007;25:579-86.
51. Cheson BD CF, Schwartz L, Zucca E, Lister A et al. Initial Evaluation, Staging and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: Lugano Classification. J Clin Oncol. 2014;32:3064-5.
52. Johnston PB ID, Colgan JP, Laplant BR, et al. Phase II trial of oral mTOR inhibitor everolimus in relapsed Hodgkin lymphoma. Am J Hematol. 2010;85:320-5.
53. Fehniger TA LS, Trinkaus K, Seigel MS, et al. . A phase 2 multicenter study of lenalidomide in relapsed or refractory classical Hodgkin lymphoma. . Blood. 2011;19:5119-26.
54. Younes A SA, Ben-Yehuda D, Per Luigi Z, Tee Chaun O, et al. Panobinostat in Patients with Relapsed/Refractory Hodgkin's Lymphoma after Autologous Stem Cell Transplantation: Results of a Phase II Study. J Clin Oncol. 2012;30:2197-203.
55. Moskowitz A HJP, Miguel-Angel P, et al. Phase II Study of Bendamustine in Relapsed and Refractory Hodgkin lymphoma. J Clin Oncol. 2012;31:456-60.
56. GA Wells BS, D O'Connell, J Peterson, V Welch, M Losos, P Tugwell, . The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
57. NICE. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma TA462. https://www.nice.org.uk/guidance/ta462/resources/nivolumab-for-treating-relapsed-or-refractory-classical-hodgkin-lymphoma-pdf-82604902197445: 2017.
58. D. A. Assessing exchangeability in indirect and mixed treatment comparisons2011.
59. Ishak KJ PI, Benedict A. Simulation and matching-based approaches for indirect comparison of treatments. Pharmacoeconomics. 2015;33(6):537-49.
60. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9.
61. Precision Health Economics. MSD Data on file: Pembrolizumab for relapsed/refractory classical Hodgkin's lymphoma: Update of systematic literature review and matching-adjusted indirect comparison. 2017.
62. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton KJ. NICE DSU Technical support document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Regent Court, 30 Regent Street Sheffield, S1 4DA: 2016 December 2016. Report No.
63. Haematological Malignancy Research Network (HMRN), Haematological Malignancy Research Network: Proposal to Audit the Clinical Management and Outcome in Hodgkin Lymphoma (Proposal). MSD data on file: 2017.
64. MSD UK Ltd. Clinicians Survey Results, MSD UK and MedeConnect. MSD data on file: 2017.
65. MSD UK Ltd. Clinician Survey Questionnaire (MSD UK, MedeConnect). MSD data on file: 2017.
66. Merck & Co. Keytruda (MK-3475) relapsed or refractory classical Hodgkin Lymphoma HTA Submission Safety Endpoints (September 2016). 2016.
67. Merck & Co. Merck Data on file. Safety data GVHD of Keynote-087 September 2016 Data. 2017.
68. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. NEnglJ Med. 2015;372(21):2018-28.
69. Herbst RS, Baas P, Kim DW, Felip E, Perez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet. 2015.

70. Ribas A, Puzanov I, Dummer R, Schadendorf D, Hamid O, Robert C, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. Lancet Oncol. 2015;16(8):908-18.
71. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med. 2015;372(26):2521-32.
72. Muro K, Chung HC, Shankaran V, Geva R, Catenacci D, Gupta S, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. Lancet Oncol. 2016;17(6):717-26.
73. Nanda R, Chow LQ, Dees EC, Berger R, Gupta S, Geva R, et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. J Clin Oncol. 2016;34(21):2460-7.
74. Seiwert TY, Burtness B, Mehra R, Weiss J, Berger R, Eder JP, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. Lancet Oncol. 2016;17(7):956-65.
75. NICE. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma [ID972] Committee Papers. https://www.nice.org.uk/guidance/ta462/documents/committee-papers: 2017.
76. NICE. Guide to the methods of technology appraisal. 2013.
77. National Institute for Health and Care Excellence. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma Final scope. 2017.
78. Beth Woods ES, Stephen Palmer, Nick Latimer, Marta Soares. NICE Decision Support Unit: Partition survival analysis for decision modelling in health care: A critical review. 2017.
79. Merck & Co. Post Pembrolizumab AlloSCT treated patients' information table. MSD Data on File: 2017.
80. Lafferty N, Anandram S, Lawes N, Usman M, Liebersbach S, Benn K, et al. Allogeneic stem cell transplantaion in patients with Hodgkin Lymphoma - a retrospective single centre case series. Br J Haematol. 2017;176(Suppl 1):5-145.
81. Pidala J, Anasetti C, Jim H. Quality of life after allogeneic hematopoietic cell transplantation. Blood. 2009;114(1):7-19.
82. NICE. Bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation [TA311]. 2015.
83. NICE. Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's Bcell lymphoma [TA306]. National Institute for Health and Care Excellence. 2014.
84. Latimer N. NICE DSU Technical Support Document 14 : Survival analysis for economic evaluations alongside clinical trials - Extrapolation with patient-level data. http://www nicedsu org uk/ [Internet]. 2011.
85. NICE. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma [ID972] Committee Papers 2 2017 [Available from: https://www.nice.org.uk/guidance/ta462/documents/committee-papers-2.
86. Perrot A, Monjanel H, Bouabdallah R, Quittet P, Sarkozy C, Bernard M, et al. Impact of post-brentuximab vedotin consolidation on relapsed/refractory CD30+ Hodgkin lymphomas: a large retrospective study on 240 patients enrolled in the French Named-Patient Program. Haematologica. 2016;101(4):466-73.
87. ONS. National Life Tables: England 2013-15 2016 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables.
88. Moskowitz CH, Bertino JR, Glassman JR, Hedrick EE, Hunte S, Coady-Lyons N, et al. Ifosfamide, carboplatin, and etoposide: a highly effective cytoreduction and peripheral-blood progenitor-cell mobilization regimen for transplant-eligible patients with non-Hodgkin's lymphoma. Journal of Clinical Oncology. 1999;17(12):3776-85.

89. Zinzani P, Barbieri E, Visani G, Gherlinzoni F, Perini F, Neri S, et al. Ifosfamide, epirubicin and etoposide (IEV) therapy in relapsed and refractory high-grade non-Hodgkin's lymphoma and Hodgkin's disease. Haematologica. 2002;79(6):508-12.
90. Ferme C, Bastion Y, Lepage E, Berger F, Brice P, Morel P, et al. The MINE regimen as intensive salvage chemotherapy for relapsed and refractory Hodgkin's disease. Ann Oncol. 1995;6(6):543-9.
91. Sibon D, Ertault M, Al Nawakil C, de Bazelaire C, Franchi P, Brière J, et al. Combined ifosfamide, etoposide and oxalipatin chemotherapy, a low- toxicity regimen for first-relapsed or refractory Hodgkin lymphoma after ABVD/EBVP: a prospective monocentre study on 34 patients. British journal of haematology. 2011;153(2):191-8.
92. 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.
93. Chau I, Harries M, Cunningham D, Hill M, Ross PJ, Archer CD, et al. Gemcitabine, cisplatin and methylprednisolone chemotherapy (GEM- P) is an effective regimen in patients with poor prognostic primary progressive or multiply relapsed Hodgkin's and non Hodgkin's lymphoma. British journal of haematology. 2003;120(6):970-7.
94. Baetz T, Belch A, Couban S, Imrie K, Yau J, Myers R, et al. Gemcitabine, dexamethasone and cisplatin is an active and non-toxic chemotherapy regimen in relapsed or refractory Hodgkin's disease: a phase II study by the National Cancer Institute of Canada Clinical Trials Group. Annals of Oncology. 2003;14(12):1762-7.
95. Bartlett N, Niedzwiecki D, Johnson J, Friedberg J, Johnson K, Van Besien K, et al. Gemcitabine, vinorelbine, and pegylated liposomal doxorubicin (GVD), a salvage regimen in relapsed Hodgkin's lymphoma: CALGB 59804. Annals of Oncology. 2007;18(6):1071-9.
96. Aparicio J, Segura A, Garcera S, Oltra A, Santaballa A, Yuste A, et al. ESHAP is an active regimen for relapsing Hodgkin's disease. Annals of Oncology. 1999;10(5):593-5.
97. Rodriguez J, Rodriguez M, Fayad L, McLaughlin P, Swan F, Sarris A, et al. ASHAP: a regimen for cytoreduction of refractory or recurrent Hodgkin's disease. Blood. 1999;93(11):3632-6.
98. Josting A, Rudolph C, Reiser M, Mapara M, Sieber M, Kirchner H, et al. Time-intensified dexamethasone/cisplatin/cytarabine: an effective salvage therapy with low toxicity in patients with relapsed and refractory Hodgkin's disease. Annals of Oncology. 2002;13(10):1628-35.
99. Rigacci L, Fabbri A, Puccini B, Chitarrelli I, Chiappella A, Vitolo U, et al. Oxaliplatin-based chemotherapy (dexamethasone, high- dose cytarabine, and oxaliplatin)±rituximab is an effective salvage regimen in patients with relapsed or refractory lymphoma. Cancer. 2010;116(19):4573-9.
100. Moskowitz AJ, Hamlin Jr PA, Perales M-A, Gerecitano J, Horwitz SM, Matasar MJ, et al. Phase II study of bendamustine in relapsed and refractory Hodgkin lymphoma. Journal of Clinical Oncology. 2012;31(4):456-60.
101. Paul Dolan CG, Paul Kind, Alan Williams. A social tariff for EuroQoL: Results from a UK general population survey. The University of York Centre for Health Economics. 1995.
102. Merck & Co. Keytruda (MK-3475) relapsed or refractory classical Hodgkin Lymphoma HTA Submission UK - Utility Data. MSD Data on File: 2017.
103. 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. Leuk Lymphoma. 2015;56(6):1839-45.
104. Ramsey SD, Nademanee, A., Masszi, T., Holowiecki, J., Abidi, M., Chen, A., . . . Moskowitz, C. H. Quality of life results from a phase 3 study of brentuximab vedotin consolidation following autologous haematopoietic stem cell transplant for persons with Hodgkin lymphoma. . Br J Haematol. 2016;175(5):860-7.

105. Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. Health Qual Life Outcomes. 2010;8:50.
106. Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. Lung Cancer. 2008;62(3):374-80.
107. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. Br J Cancer. 2006;95(6):683-90.
108. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.
109. Shingler SL, Swinburn P, Lloyd A, Diaz J, Isbell R, Manson S, et al. Elicitation of health state utilities in soft tissue sarcoma. Qual Life Res. 2013;22(7):1697-706.
110. Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. Eur J Health Econ. 2013;14(5):749-59.
111. NICE. Ticagrelor for preventing atherothrombotic events after myocardial infarction [TA420]. 2016.
112. NICE. Paclitaxel as albumin-bound nanoparticles in combination with gemcitabine for previously untreated metastatic pancreatic cancer [TA360]. 2015.
113. Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. Bone Marrow Transplant. 2015;50(9):1241-9.
114. Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. Center for Helath Economics Discussion Paper Series. 1999;Discussion Paper 172.
115. Radford J, Johnson, R., Mckay, P., Malladi, R., Peggs, K., Sureda, A., . . . Kerrigan, M. Treatment pathways and resource use associated with the management of recurrent Hodgkin Lymphoma after autologous stem cell transplantation. Haematologica. 2013;98(38).
116. NICE. Pembrolizumab for advanced melanoma not previously treated with ipilimumab [TA366]. 2015.
117. Department of Health DoH. NHS reference costs 2015 to 2016 2016 [Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.
118. NICE. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy [TA428]. 2017.
119. Department of Health DoH. Drugs and pharmaceutical electronic market information (eMit). 2017.
120. British National Formulary BNF. British National Formulary: June 2017 2017 [Available from: https://bnf.nice.org.uk/.
121. Curtis L. Unit Costs of Health and Social Care: Personal Social Services Research Unit, University of Kent, Canterbury; 2016.
122. NICE. Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab [TA357]. National Institute for Health and Care Excellence. 2015.
123. Burton C, Linch D, Hoskin P, Milligan D, Dyer M, Hancock B, et al. A phase III trial comparing CHOP to PMitCEBO with or without G-CSF in patients aged 60 plus with aggressive non-Hodgkin's lymphoma. British journal of cancer. 2006;94(6):806.
124. Karasek P, Skacel T, Kocakova I, Bednarik O, Petruzelka L, Melichar B, et al. Gemcitabine monotherapy in patients with locally advanced or metastatic pancreatic cancer: a prospective observational study. Expert opinion on pharmacotherapy. 2003;4(4):581-6.
125. Pfreundschuh M, Trümper L, Kloess M, Schmits R, Feller AC, Rübe C, et al. Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of elderly patients with aggressive lymphomas: results of the NHL-B2 trial of the DSHNHL. Blood. 2004;104(3):634-41.

126. Mead GM, Barrans SL, Qian W, Walewski J, Radford JA, Wolf M, et al. A prospective clinicopathologic study of dose-modified CODOX-M/IVAC in patients with sporadic Burkitt lymphoma defined using cytogenetic and immunophenotypic criteria (MRC/NCRI LY10 trial). Blood. 2008;112(6):2248-60.
127. 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. 2013.
128. Curtis L. Unit Costs of Health and Social Care: Unit Costs of Health and Social Care; 2013.
129. Marie Curie Cancer Care. Understanding the cost of end of life care in different settings. 2004.
130. Department of Health DoH. NHS reference costs 2011 to 2012 2012 [Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.
131. Department of Health DoH. NHS reference costs 2013 to 2014 2014 [Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.
132. MSD UK Ltd. Data on file, Health Economist Expert Validation of cHL Cost Effectiveness Model Methodology. 2017.
133. Transplantation BSoBaM. 2015 Activity 2017 [Available from: http://bsbmt.org/2015-activity/.
134. ONS. Population Estimates for UK, England and Wales, Scotland and Northern Ireland: mid-2016 2016 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2016.
135. Adam Gibb CJ, Adrian Bloor,Samar Kulkarni,Tim Illidge,Kim Linton,and John Radford. Brentuximab vedotin in refractory CD30+ lymphomas: a bridge to allogeneic transplantation in approximately one quarter of patients treated on a Named Patient Programme at a single UK center. haematologica. 2013;98(4):611-14.
136. UK Cancer Research. Hodgkin Lymphoma incidence statistics 2014 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma/incidence#collapseZero.
137. Diehl V FJ, Pfreundschuh M, Lathan B, Paulus U. Standard and increased-dose BEACOPP chemotherapy compared with COPPABVD for advanced Hodgkin's disease. New Eng J Med. 2003;348(24):2386- 95.

Appendices

Please see relevant word document

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]

Alternative data source feasibility for cohort 2

In response to NICE's query relating to a potential comparative evidence source for cohort 2 of KEYNOTE-087, MSD explored data presented by Eyre et al. 2017, titled: *'Results of a multicentre UK-wide retrospective study evaluating the efficacy of brentuximab vedotin in relapsed, refractory classical Hodgkin lymphoma in the transplant naïve setting'.* This was not included in the company SLR as it was not available at the time of the search.

KEYNOTE-087 reported patients (Cohort 2) who were unable to achieve a complete or partial response to salvage chemotherapy and did not receive autoSCT; patients must have relapsed or failed to respond to brentuximab vedotin (BV). Patients within Eyre et al:

- Were transplant naïve
- Had received at least two prior lines of therapy
- Received BV with the intention of a subsequent stem cell transplant consolidation.

Note that patients in KEYNOTE-087 had failed to respond or had relapsed following treatment with BV, and as such were eligible to receive pembrolizumab. It is unlikely that patients within Cohort 2 would be considered fit to receive SCT unlike the majority of patients reported in Eyre et al. who received SCT. Prior lines of therapy differed markedly between populations; 96.3% of patients within KEYNOTE-087 had received ≥3 lines if prior therapy (mean and median of 4 lines of therapy); compared with patients of Eyre et al. with a median of 2 prior lines of therapy, with 29% (n=29) receiving 3 or 4 lines of prior therapy. This demonstrates that the population of KEYNOTE-087 was more heavily treated and further advanced in their disease course..

It is feasible that approximately 30 patients from Eyre et al. are relevant to the decision problem. This is the *"30% (n = 30) who received BV followed by additional therapy. Typically, patients did not proceed to SCT because of inadequate treatment response [stable or progressive disease (n = 24)]"*. Even if appropriate, to include this data MSD would require further information relating to the subsequent intervention received, and would also require new aggregate data for this patient population alone; i.e. new baseline characteristics table, and outcomes reported according to this group only.

In summary, due to the limited data available, and a lack of IPLD it is not feasible to consider this evidence. Furthermore, MSD would need a KM curve from the point of initiation of subsequent therapy in **1)** the patients with deferred SCT (n=27) and **2)** patients who received this in the "none" group (n=30); this information does not appear to have been captured.

Cost-effectiveness analysis

In response to the request from NICE (dated 19th December 2017), MSD has amended the economic model used in the original submission to 2 models (new 12 week model and new 24 week model) with the following:

Requested by NICE:

- Analysis of the possibility of having an allogeneic stem cell transplant at 12 and 24 weeks
- Analysis including a progressive disease state after allogeneic stem cell transplants

Amended in line with the ERG preferred assumptions:

- 'Fixing errors 1 and 2'
- Removal of patient characteristics from PSA

All other ERG assumptions have been reflected in scenario analysis below.

24 week model update

The new 24 week model shows the cost-effectiveness analysis of pembrolizumab vs SoC with alloSCT occurring at week 24 and the inclusion of a progressed disease heath state post alloSCT.

All parameters or assumptions which have been updated for the week 24 model are described in the following sections and relate to the following bullet points. Anything not mentioned is as it was in the original company submission model. For the new 24 week model, the following data have been updated from week 12 to 24.

- 0-24 week PFS parametric distribution updated
- 0-24 week OS Hazard Ratio (HR) updated
- PFS and ToT non alloSCT pathway post 24 week from KEYNOTE-087

- Response rates from KEYNOTE-087
- Odds ratios for response rates
- Pembrolizumab utility values updated using week 24 response rates

Progression-free survival (week 0 to 24) – Pembrolizumab

For the analysis of week 0 to 24, parametric models were fitted using all observed data from study initiation given that only a small number of events occurred in the first 24 as per the original company submission. PFS for SoC was estimated by applying the HR described in the original company submission (section 4) ([redacted])

For the new 24 week model, the following assumptions were necessary to update from the original 12 week company submission model (detail below):

- PFS distribution week 0 to 24 has been updated from the week 12 distribution to exponential in order to match the observed Kaplan Meier (KM) data at week 24 as closely as possible. This is in line with the selection criteria used in the original company submission.
    - Cohort 1 KM at week 24 was [redacted]% progression free with the exponential curve predicting closest to this at 79.02%
    - Cohort 2 KM at week 24 was [redacted]% progression free with the exponential curve predicting closest to this at 66.83%

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of PFS at week 24 are shown for each distribution in Table 1.

Table 1: Summary of the goodness of fit qualities of the survival models (cohort 1; PFS prior 24 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087
AIC	329.4049	326.7148	328.5895	326.2135	327.1309	328.312	-
Rank	6	2	5	1	3	4	-

BIC	331.639	331.183	333.0577	330.6817	331.5992	335.0143	-
Rank	4	2	5	1	3	6	-
% at week 24	79.02%	84.86%	83.68%	84.44%	82.98%	84.31%	[redacted]%

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 1.

Figure 1. PFS (BIRC) extrapolations (cohort 1; PFS 24 weeks)

[redacted]

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of PFS at 24 weeks are shown for each distribution in

Table **2**.

Table 2. Summary of the goodness of fit qualities of the survival models (cohort 2; PFS prior 24 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087
AIC	482.0749	474.5633	477.3447	474.8582	470.9785	465.038	-
Rank	6	3	5	4	2	1	-
BIC	484.4694	479.3522	482.1336	479.6471	475.7674	472.2214	-
Rank	6	3	5	4	2	1	-
% at week 24	66.83%	74.60%	74.07%	72.33%	71.53%	60.93%	[redacted]%

The fit of the parametric models to the Kaplan-Meier data is shown graphically in

Figure **2**.

Figure 2: PFS (BIRC) extrapolations (cohort 2; prior 24 weeks)

Overall survival (week 0 to 24) – Pembrolizumab

The analysis of OS from week 0 to 24 are as per the company submission since no additional events had occurred between 12 and 24 weeks and the original data in the company submission model was fitted to all observed data so this is as it was in the original company submission.

Contrary to the original company submission model, it was necessary to update the hazard ratio for overall survival (OS). The HR for OS for the week 0 to 12 model was originally assumed to be 1.00. However, given that the alloSCT decision point is not until week 24 in the new week 24 model, the number of patients alive on SoC would be substantially overestimated (98% vs. ~88% estimated from digitized Cheah OS curve). As we do not have a HR for OS from March 2017 data cut, we have used an OS HR for week 0-24 pooled from cohorts 1&2 data of 13.13 (95% CI (3.07-56.04)) as this is the only available evidence based estimate and is statistically significant. There are limitations associated with this since it is based on June 2016 data and is a pooled estimate, however given the low number of overall death events from week 0-24, we do not expect this estimate to have changed. A scenario analysis has been provided using a HR of 1.00 for OS week 0-24 also.

Response rates

Response rates were applied at week 24 in the model to apportion patients that were progression-free into CR, PR or SD. The proportions of patients with either CR/PR were estimated directly from observed data (presented in this section) with the remaining progression-free patients that had not achieved responses were assumed to occupy the SD node.

The cohort specific response rates at week 24 from KEYNOTE-087 are presented in Table 3.

Table 3: KEYNOTE-087 number of complete and partial responders at week 24

Response	Cohort 1	Cohort 2
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]

Odds ratios

As in the original company submission, the comparative response of SoC was estimated via ORs estimated from a naïve indirect comparison. The associated ORs estimated for cohort 1 and 2 are presented below in Table 4.

Table 4: Odds ratios for response at week 24

Response	Cohort 1		Cohort 2	
	Mean	SE	Mean	SE
CR	[redacted]	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]	[redacted]

Progression-free survival (post 24 weeks) – Pembrolizumab

The analysis of PFS from week 24, involved fitting parametric models to the observed data in KEYNOTE-087 from week 24 as per the original company submission.

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of PFS over time post 24 weeks are shown for each distribution in Table 5.

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 3.

Table 5: Summary of the goodness of fit qualities of the survival models (cohort 1; PFS post 24 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort 1)
AIC	188.7993	190.1	190.4944	191.489	194.5751	190.884	-
Rank	1	2	3	5	6	4	-
BIC	190.6912	193.8837	194.278	195.2726	198.3588	196.5595	-
Rank	1	2	3	4	6	5	-
Median (months)	14.03	15.64	12.88	18.63	24.84	13.80	-
Mean (months)	20.58	27.17	14.56	70.09	107.74	15.63	-
% at 1 year	56.24%	58.04%	54.14%	60.31%	62.03%	55.86%	[redacted]
% at 2 years	31.46%	37.83%	18.95%	44.80%	50.79%	25.89%	-
% at 5 years	5.44%	12.32%	0.00%	26.15%	35.97%	0.00%	-
% at 10 years	0.29%	2.40%	0.00%	15.95%	25.97%	0.00%	-

Figure 3: PFS (BIRC) cohort 1 from week 24 extrapolations

The exponential was applied in the base case for the same reason as in the original company submission.

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of PFS over time after week 24 are shown for each distribution in Table 6.

Table 6: Summary of the goodness of fit qualities of the survival models (cohort 2; PFS post 24 weeks)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	KEYNOTE-087 (cohort 2)
AIC	232.1969	228.5293	226.7207	231.2036	234.1444	224.632	-
Rank	5	3	2	4	6	1	-
BIC	234.0681	232.2717	230.4631	234.946	237.8868	230.2456	-
Rank	4	3	2	5	6	1	-
Median (months)	8.51	8.28	8.74	8.28	8.28	8.05	-
Mean (months)	12.61	9.62	8.83	14.69	15.19	8.23	-
% at 1 year	39.09%	30.98%	27.65%	35.95%	38.27%	20.81%	[redacted]%
% at 2 years	15.14%	2.92%	0.00%	13.63%	17.02%	0.00%	-
% at 5 years	0.86%	0.00%	0.00%	2.86%	3.44%	0.00%	-
% at 10 years	0.01%	0.00%	0.00%	0.83%	0.67%	0.00%	-

The fit of the parametric models to the Kaplan-Meier data is shown graphically in

Figure 4.

Figure 4: PFS (BIRC) cohort 2 from week 24 extrapolations

In the base case analysis, the exponential was applied with the gompertz considered in scenario analysis given the uncertainty in the tail of the Kaplan-Meier curve in line with the original company submission.

Mortality pre-progression

As in the original company submission.

Post-progression survival

At the time of analysis, the number of patients that progressed in KEYNOTE-087 was judged to be too small to support robust analysis of post-progression survival. This was modelled as per the original company submission.

Progression free survival post alloSCT

Progressed disease post-alloSCT has been modelled using a partitioned survival model informed by progression-free survival and overall survival data from Lafferty (2017). The base case parametric model for progression-free survival has been modelled using the Weibull as this was the conservative option of the two parametric models previously recommended by the ERG in TA462 (Weibull or log-normal).

Overall survival post-AlloSCT

In line with original company submission.

Time on treatment post-24 weeks

Figure 5: Kaplan-Meier Analysis of ToT and PFS from KEYNOTE-087 cohort 1

Figure 6: Kaplan-Meier Analysis of ToT and PFS from KEYNOTE-087 cohort 2

As in the original company submission, the duration of pembrolizumab treatment is modelled via the simulation of ToT data from week 24 onwards in KEYNOTE-087 and extrapolated to a maximum time period of 24 months. The "best fitting" model for cohort 1 and cohort 2 is described below. It was assumed that PFS was a reasonable proxy for ToT for SoC as no treatment discontinuation data was available for SoC.

Cohort 1

A summary of the goodness of fit statistics and modelled probabilities of ToT over time are shown for each distribution in Table 7.

Table 7: Summary of the goodness of fit qualities of the survival models

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma
AIC	268.4502	270.2743	270.1346	271.8474	274.57	270.7554
Rank	1	3	2	5	6	4
BIC	270.4575	274.289	274.1493	275.8621	278.5846	276.7774
Rank	1	3	2	4	6	5
Median (months)	11.96	12.19	11.50	13.34	14.72	11.50
Mean (months)	13.16	13.31	12.63	13.94	14.17	12.20
% at 1 year	51.06%	51.67%	49.91%	53.56%	54.71%	49.69%

The fit of the parametric models to the Kaplan-Meier data is shown graphically in

Figure **7**.

Figure 7: ToT cohort 1 from week 24 extrapolations

The exponential was applied for the same reasons and in line with the original company submission.

Cohort 2

A summary of the goodness of fit statistics and modelled probabilities of ToT over time are shown for each distribution in Table 8.

Table 8: Summary of the goodness of fit qualities of the survival models

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma
AIC	344.2997	340.5774	345.0553	343.1128	344.8381	341.362
Rank	4	1	6	3	5	2
BIC	346.325	344.6281	349.106	347.1635	348.8888	347.4381
Rank	2	1	6	3	5	4
Median (months)	6.67	6.21	6.21	6.21	5.98	6.67
Mean (months)	9.09	9.64	9.76	10.39	10.38	8.84
% at 1 year	30.29%	34.27%	32.99%	37.38%	37.92%	32.26%

The fit of the parametric models to the Kaplan-Meier data is shown graphically in Figure 8.

Figure 8: ToT cohort 2 from week 24 extrapolations

The exponential was chosen to maintain consistency with the base case PFS distribution and in line with the original company submission.

Utility data

Utility values have been updated for week 24 response rates. All other utility values are in line with the original base case submission.

Table 9: Week 24 utility values for cost-effectiveness analysis

	Response utility values from KEYNOTE-087	Pembrolizumab (cohort 1)		Pembrolizumab (cohort 2)		SoC	
		Response proportion [KEYNOTE -087]*	Input	Response proportion [KEYNOTE -087]*	Input	Response proportion [Cheah 2016]	Input
Complete response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial response	[redacted]	[redacted]		[redacted]		[redacted]	
Stable disease	[redacted]	[redacted]		[redacted]		[redacted]	
*patients that had progressed at week 24 were excluded; those who were not assessed were assumed to have stable disease							

Please note that all new analyses presented in this document are based on the original model assumptions and parameters, submitted by MSD, with the aforementioned alterations in line with the ERG and NICE requests.

Base-case results

Table 10 shows the base case results from the original company submission alongside the new week 24 and week 12 model base case results.

Table 10 Base-case results

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs (£)	Total LYs	Total QALYs	Total costs (£)	Total LYs	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
Base case (company submission)	Cohort 1	107,459	6.252	4.497	52,017	4.864	3.223	55,442	1.274	43,511
	Cohort 2	93,732	5.775	4.072	51,424	4.832	3.200	42,308	0.871	48,571
Base case new 24 week model	Cohort 1	106,051	6.025	3.612	34,320	3.407	1.813	71,730	1.799	39,880
	Cohort 2	89,726	5.427	3.154	33,217	3.236	1.731	56,509	1.423	39,714
Base case new 12 week model	Cohort 1	107,459	6.252	4.328	52,018	4.864	3.097	55,441	1.231	45,034
	Cohort 2	93,733	5.775	3.917	51,425	4.832	3.077	42,307	0.840	50,353

Sensitivity analysis

Probabilistic sensitivity analysis (PSA), deterministic sensitivity analysis (DSA) and scenario analysis have been run on the base case analysis in line with the original company submission.

Probabilistic sensitivity analysis

Week 24

Table 11 Probabilistic sensitivity analysis results week 24

Scenario	Cohort	Pembrolizumab		SoC				
		Total costs (£)	Total QALYs	Total costs (£)	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
Base case (company submission)	Cohort 1	106,672	4.361	46,723	2.857	59,949	1.505	39,841
	Cohort 2	92,941	3.875	45,391	2.771	47,550	1.105	43,049
Base case new 24 week model	Cohort 1	105,655	3.593	33,494	1.678	72,161	1.915	37,682
	Cohort 2	90,581	3.176	35,129	1.784	55,452	1.392	39,828

Figure 9 Cost effectiveness Plane week 24 cohort 1

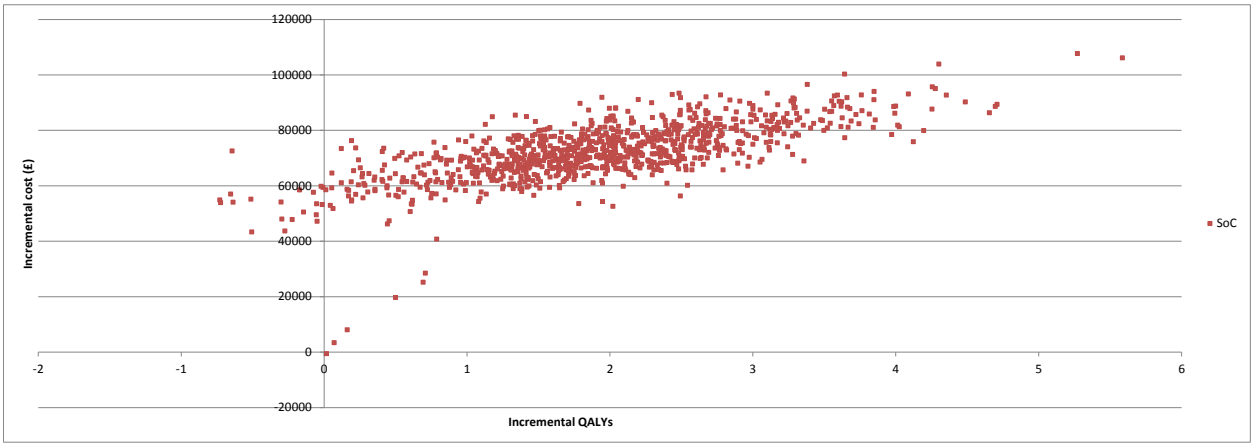


Figure 10 Cost effectiveness Plane week 24 cohort 2

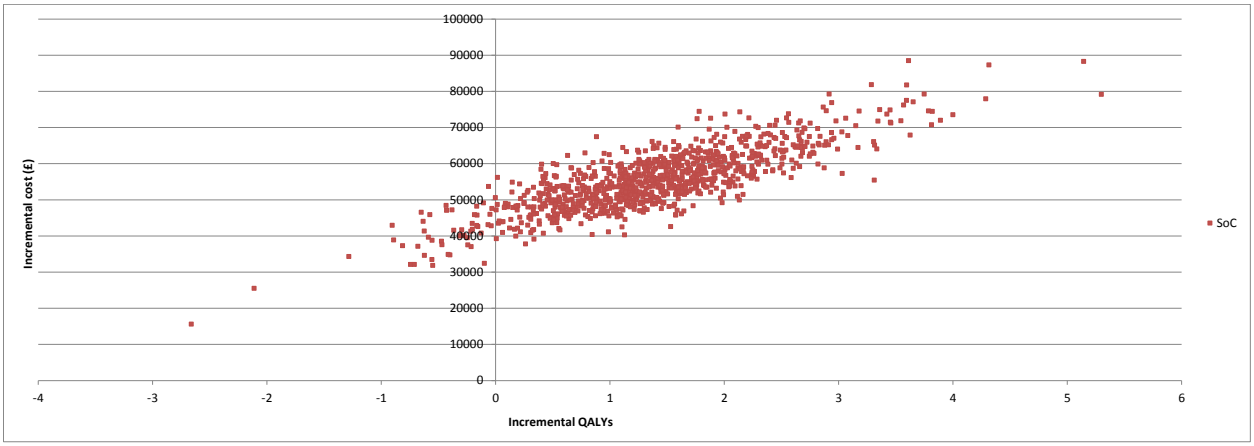


Figure 11 Cost effectiveness acceptability curve week 24 cohort 1

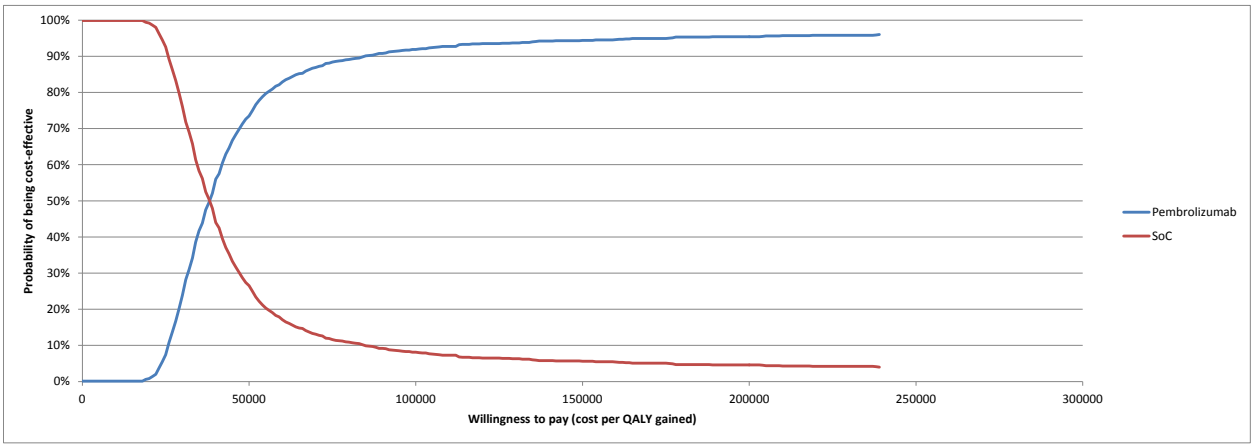
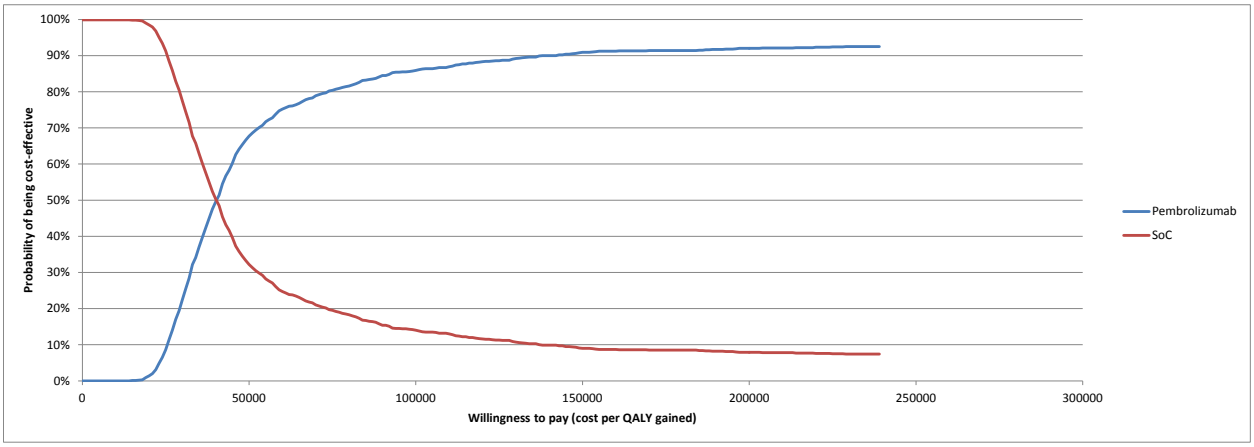


Figure 12 Cost effectiveness acceptability curve week 24 cohort 2



Week 12

Table 12 Probabilistic sensitivity analysis results week 12

Scenario	Cohort	Pembrolizumab		SoC				
		Total costs (£)	Total QALYs	Total costs (£)	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
Base case (company submission)	Cohort 1	106,672	4.361	46,723	2.857	59,949	1.505	39,841
	Cohort 2	92,941	3.875	45,391	2.771	47,550	1.105	43,049
Base case new 12 week model	Cohort 1	106,624	4.032	53,830	2.968	52,794	1.065	49,588
	Cohort 2	94,424	3.720	53,810	2.977	40,614	0.742	54,704

Figure 13 Cost effectiveness plane week 12 cohort 1

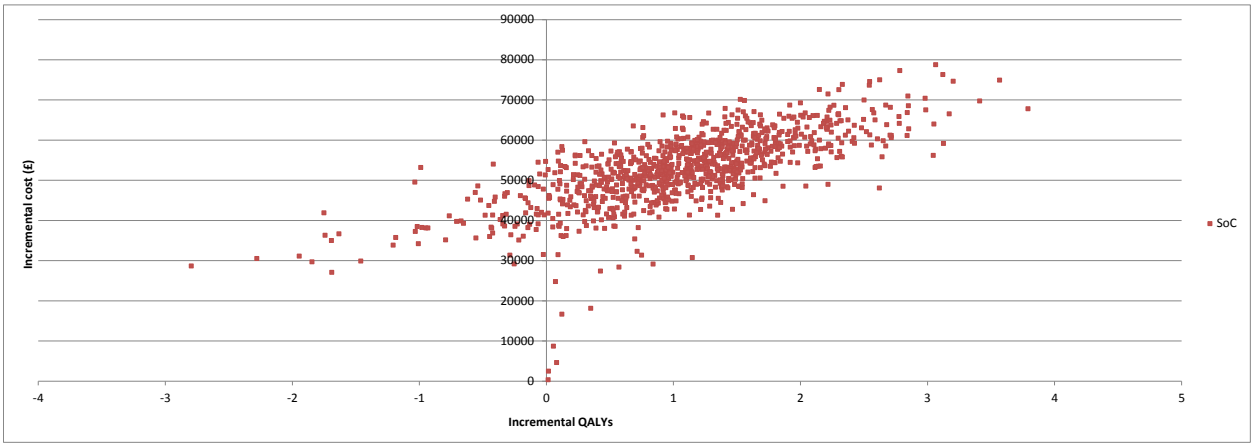


Figure 14 Cost effectiveness plane week 12 cohort 2

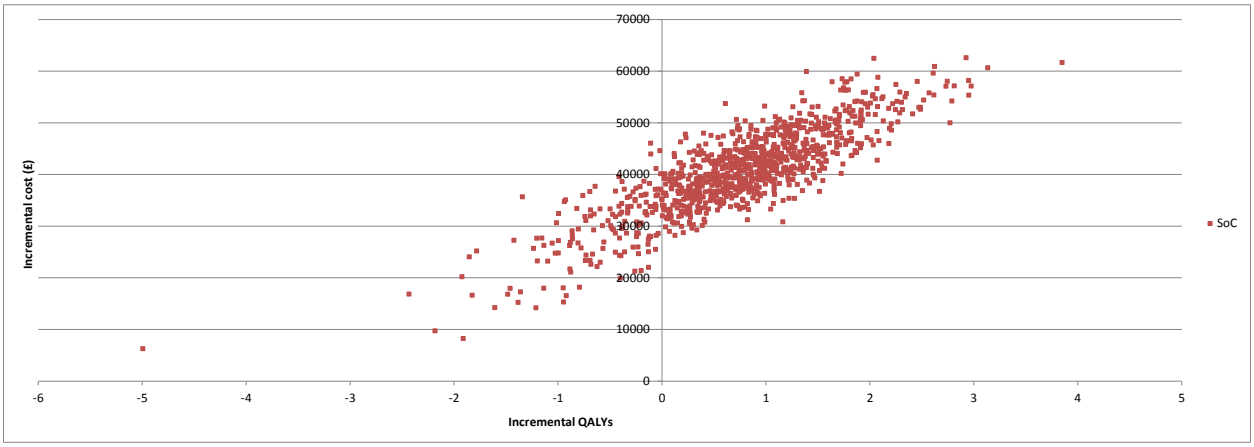


Figure 15 Cost effectiveness acceptability curve week 12 cohort 1

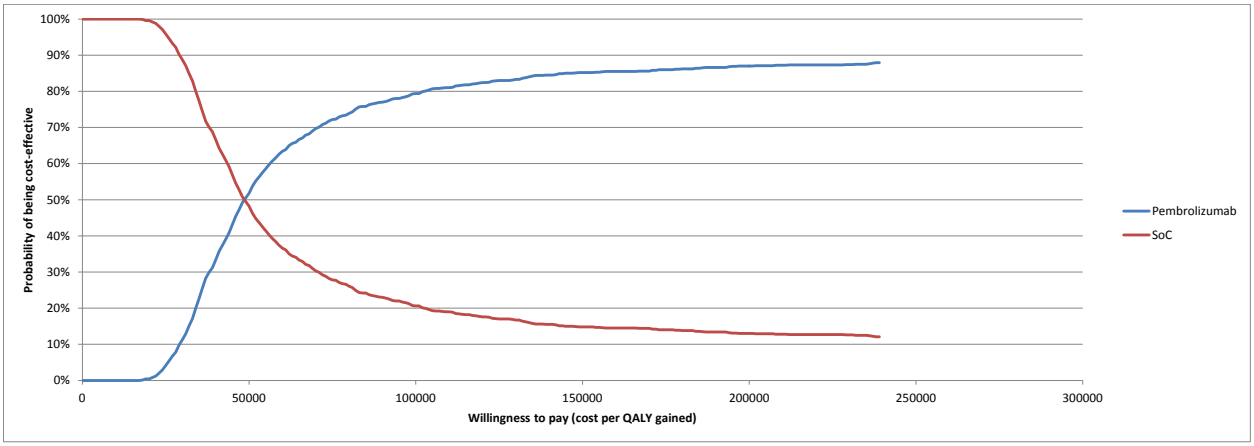
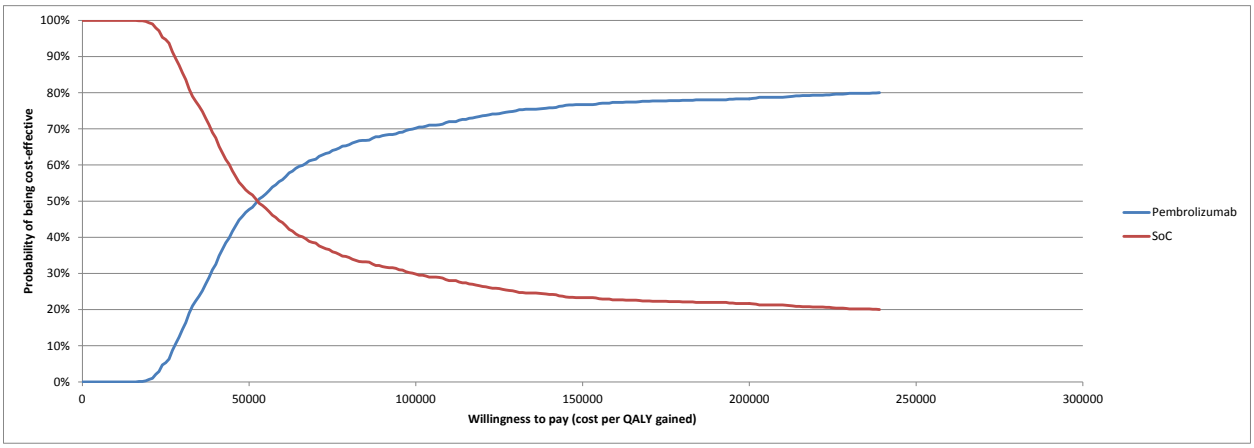


Figure 16 Cost effectiveness acceptability curve week 12 cohort 2



Deterministic sensitivity analysis

Week 24

Figure 17 Week 24 Tornado diagram cohort 1

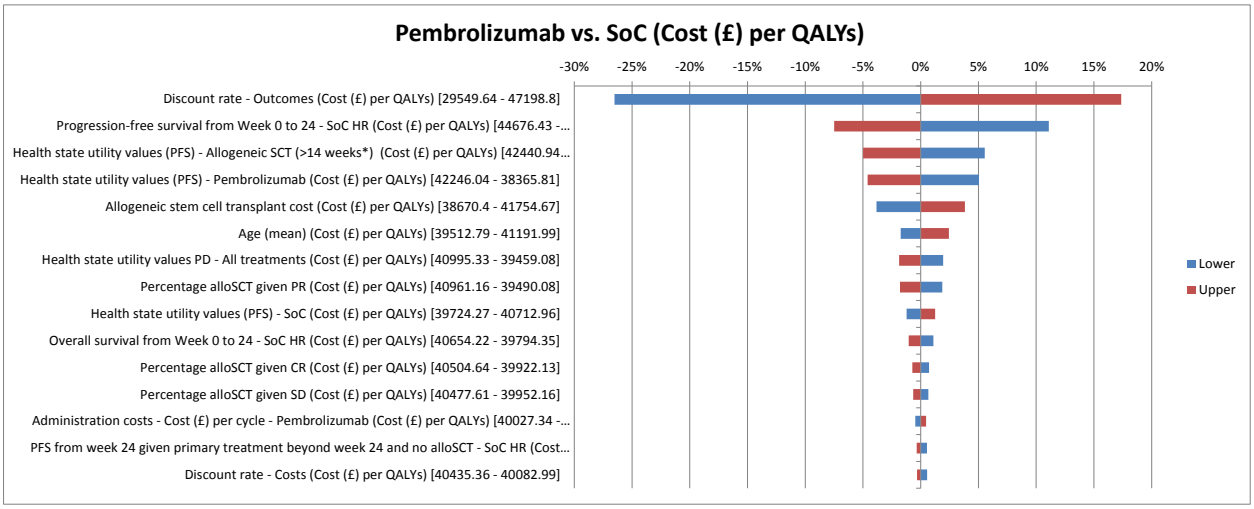
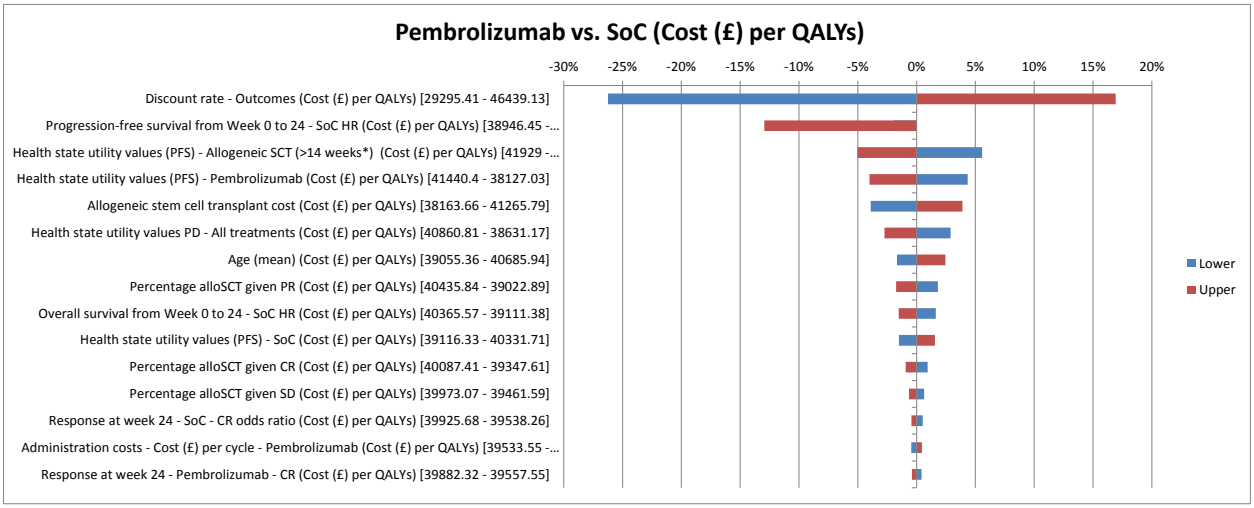


Figure 18 Week 24 Tornado diagram cohort 2



Week 12

Figure 19 Week 12 Tornado diagram cohort 1

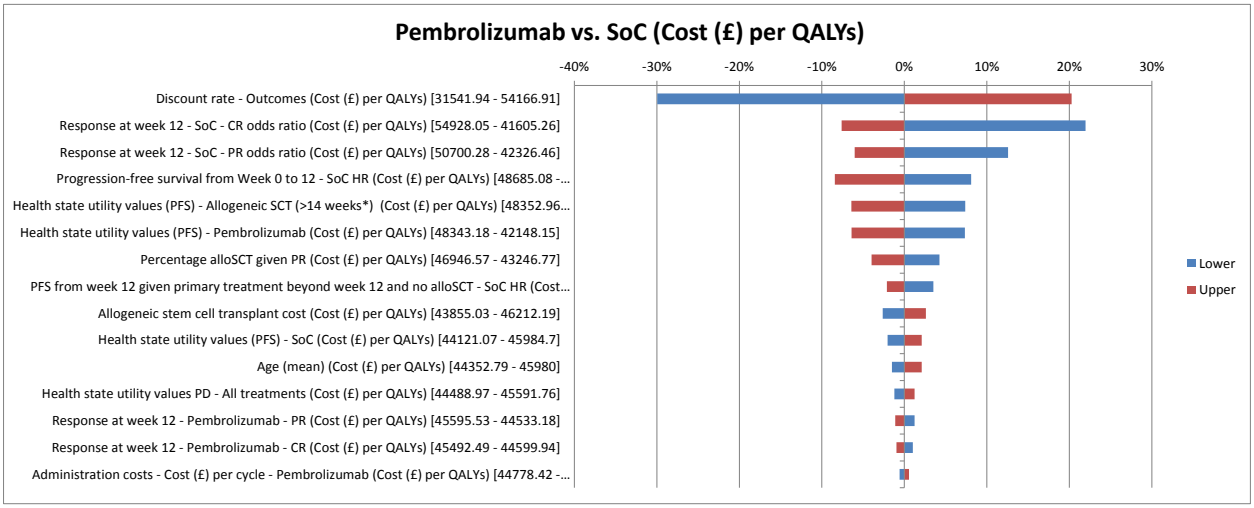
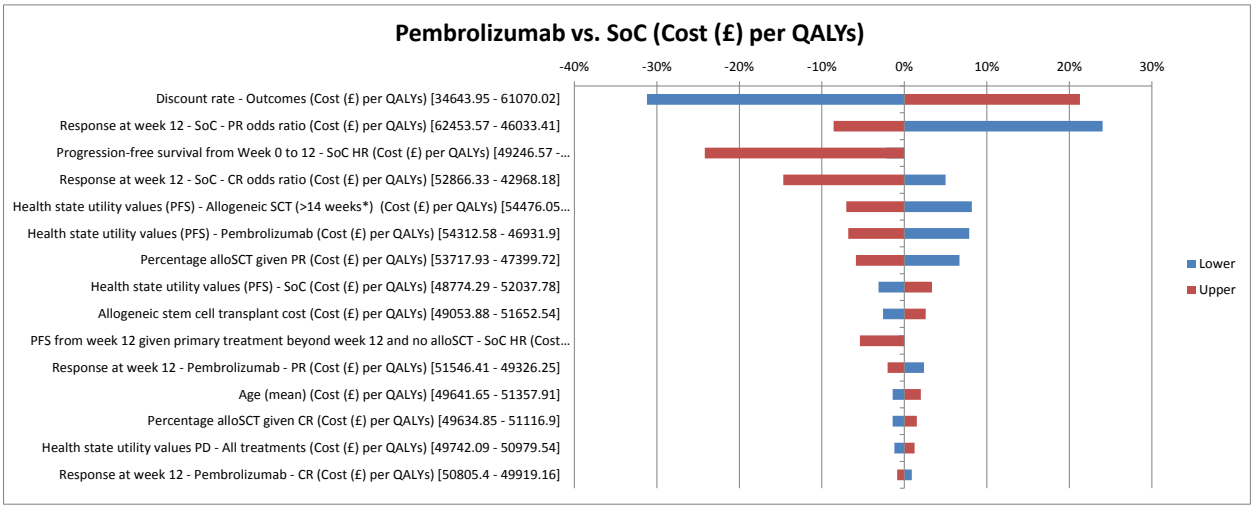


Figure 20 Week 12 Tornado diagram cohort 2



Scenario analysis

The following scenario analysis has been conducted as per ERG comments:

1. Inclusion of results of mixed modelling of utilities by response status in KEYNOTE-087

2. Inclusion of long term monitoring costs post alloSCT using the same assumptions applied in TA462; a monthly cost of £91.69.

3. Use of MSD survey means for alloSCT only (CR: 56.79%, PR: 43.93%, SD: 18.36%)

4. Time horizon of 50 years.
5. Distributions for pre-12 weeks OS to reflect ERG (and applying these distributions for 0-24 week OS)
    - Cohort 1: exponential (5a)
    - Cohort 2: lognormal (5b)

The following scenario analysis has been conducted in line with the original company submission scenario analysis if not already included above.

6. Assessing BSC as a comparator as per the NICE scope
7. Assessing different alloSCT rates
    - 100% alloSCT in patients with CR, PR or SD (7a)
    - Alternative lower PR alloSCT rate from MSD clinician survey (7b)
8. Use of MAIC HE and OR rather than naïve ITC
9. Alternative extrapolation scenarios to estimate PFS and OS at week 12 (12 week model only)
    - Considering a Weibull curve for week 0-12 PFS extrapolation in cohort 2 (9a)
    - Considering a Gompertz curve for week 12+ PFS extrapolation in cohort 2 (9b)
    - Considering a Lognormal curve following alloSCT (9c)
10. Use of an OS HR 0-24 weeks = 1.00 (24 week model only)
11. ERG combined preferences (1-5)

Week 24

Table 13 New 24 week model scenario analysis

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs (£)	Total LYs	Total QALYs	Total costs (£)	Total LYs	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
1	Cohort 1	106,051	6.025	4.503	34,320	3.407	2.538	71,730	1.965	36,505
	Cohort 2	89,726	5.427	4.050	23,617	3.160	2.373	66,109	1.677	39,411
2	Cohort 1	109,777	6.025	3.612	26,670	3.331	1.581	83,107	2.031	40,927
	Cohort 2	92,957	5.427	3.154	34,565	3.236	1.731	58,391	1.423	41,037
3	Cohort 1	103,624	5.667	3.388	31,797	3.254	1.710	71,827	1.678	42,794
	Cohort 2	86,464	5.082	2.931	30,752	3.086	1.631	55,711	1.301	42,836
4	Cohort 1	106,051	6.140	3.655	34,320	3.450	1.829	71,730	1.826	39,290
	Cohort 2	89,726	5.526	3.191	33,217	3.277	1.747	56,509	1.445	39,118
5a	Cohort 1	106,104	6.035	3.617	34,893	3.518	1.863	71,211	1.754	40,610
5b	Cohort 2	90,021	5.484	3.180	36,469	3.864	2.013	53,552	1.166	45,910
6	Cohort 1	106,050	6.025	3.612	33,328	3.407	1.821	72,721	1.791	40,612
	Cohort 2	89,725	5.427	3.154	32,217	3.236	1.739	57,507	1.415	40,645
7a	Cohort 1	117,665	7.739	4.681	42,724	3.918	2.157	74,940	2.524	29,687
	Cohort 2	104,908	7.030	4.192	41,426	3.735	2.066	63,482	2.125	29,870
7b	Cohort 1	104,676	5.822	3.485	32,868	3.319	1.754	71,808	1.731	41,471
	Cohort 2	87,951	5.239	3.033	31,799	3.149	1.673	56,152	1.359	41,304
8	Cohort 1	106,051	6.025	3.612	33,369	3.349	1.774	72,682	1.838	39,553
	Cohort 2	89,726	32,655	5.427	32,655	3.202	1.708	57,070	1.446	39,473
10	Cohort 1	106,051	6.025	3.612	37,520	4.025	2.091	68,530	1.521	45,048
	Cohort 2	89,726	5.427	3.154	37,128	3.991	2.070	52,598	1.084	48,523
11	Cohort 1	106,721	5.768	4.317	33,526	3.399	2.532	73,195	1.784	41,021
	Cohort 2	89,408	5.218	3.898	35,134	3.748	2.795	54,274	1.103	49,220

Week 12

Table 14 New 12 week model scenario analysis

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs (£)	Total LYs	Total QALYs	Total costs (£)	Total LYs	Total QALYs	Inc. costs (£)	Inc. QALYs	ICER (£)
1	Cohort 1	107,459	6.252	4.740	52,018	4.864	3.660	55,441	1.080	51,319
	Cohort 2	93,733	5.775	4.375	51,425	4.832	3.637	42,307	0.738	57,308
2	Cohort 1	111,622	6.252	4.328	54,931	4.864	3.097	56,690	1.231	46,047
	Cohort 2	93,733	5.775	3.917	46,652	4.538	2.811	47,080	1.106	42,551
3	Cohort 1	105,128	5.832	3.941	45,920	4.487	2.756	59,208	1.184	49,987
	Cohort 2	89,745	5.333	3.514	45,464	4.465	2.744	44,281	0.769	57,548
4	Cohort 1	107,459	6.377	4.432	52,018	4.951	3.170	55,441	1.262	43,917
	Cohort 2	93,733	5.889	4.012	51,425	4.918	3.149	42,307	0.864	48,980
5a	Cohort 1	107,497	6.259	4.331	52,055	4.871	3.100	55,441	1.231	45,033
5b	Cohort 2	93,968	5.809	3.933	51,608	4.866	3.092	42,359	0.840	50,413
6	Cohort 1	107,459	6.252	4.328	51,188	4.864	3.105	56,270	1.223	46,006

	Cohort 2	93,732	5.775	3.917	50,713	4.832	3.085	43,018	0.832	51,692
7a	Cohort 1	119,943	8.503	6.403	89,436	7.175	5.189	30,507	1.215	25,118
	Cohort 2	116,185.	8.261	6.187	87,472	7.053	5.087	28,713	1.100	26,105
7b	Cohort 1	106,222	6.029	4.122	49,952	4.736	2.982	56,269.	1.141	49,322
	Cohort 2	91,431	5.520	3.684	49,361.	4.705	2.962	42,070	0.723	58,221
8	Cohort 1	107,459	6.252	4.328	51,052	4.804	3.043	56,407	1.285	43,892
	Cohort 2	93,733	5.775	3.917	46,652	4.538	2.811	47,080	1.106	42,551
9a	Cohort 2	93,262	5.766	3.907	51,235	4.814	3.053	42,026	0.855	49,159
9b	Cohort 2	93,262	5.766	3.907	51,235	4.814	3.053	42,026	0.855	49,159
9c	Cohort 1	107,459	6.451	4.488	52,018	5.003	3.209	55,441	1.279	43,347
	Cohort 2	93,733	5.957	4.063	51,425	4.969	3.187	42,307.	0.876	48,282
11	Cohort 1	108,530	5.936	4.501	48,305	4.562	3.428	60,225	1.072	56,160
	Cohort 2	93,025	5.455	4.132	47,958	4.566	3.432	45,066	0.700	64,353

Conclusions

As previously discussed in the original company submission, within the context of relapsed or refractory cHL there are low patient numbers and short survival within the context of the current clinical pathway. There is considerable uncertainty and heterogeneity, particularly in the post autoSCT, post BV and autoSCT ineligible settings leading to a paucity of clinical evidence on which to base the economic evaluation. For the updated cost-effectiveness analysis shown here, assumptions have been kept in line with the original company submission and sensitivity analysis around these assumptions and those raised by the ERG were then undertaken.

In the original company submission base case analysis, it was estimated that pembrolizumab use resulted in an additional 1.274 and 0.871 discounted QALYs and 1.388 and 0.943 discounted LYs versus SoC in Cohort 1 and 2, respectively. Base case ICERs were £43,511 and £48,571 for cohorts 1 and 2 respectively.

On adding a progressed disease heath state post alloSCT and looking at a different later time point for alloSCT of 24 weeks, the resultant ICERs are £39,880 and £39,714 for cohort 1 and 2 respectively which are lower than that observed in the original base case 12 week model and that of the updated 12 week model (including a progressed disease heath state post alloSCT).

For the 24 week model, several sensitivity analyses were performed to assess the impact of variation in all variables and assumptions applied within the model. The deterministic analysis and PSA showed pembrolizumab to be cost-effective in the majority scenarios at a WTP threshold of £50,000/QALY (~75% and ~67% probability of being cost effective for cohort 1 and 2 respectively). In addition, all scenario analysis from the original company submission and ERG assumptions, except for the inclusion of a progressed disease health state leading to alloSCT as per confirmation from clinicians that this would not be appropriate, were tested and found to produce ICERs of less than £50,000/QALY. Scenario analysis 11 considered all of the ERG assumptions together (including altering of parametric models suggested for the 12 week model which may not be deemed appropriate for the new 24 week model) and produced ICERs of £41,021 and £49,220. Without the aforementioned parametric model ERG updates, but including all other ERG assumptions for the week 24 model – the ICERs were £39,633 and £37,351 for cohort 1 and 2 respectively.

On adding a progressed disease heath state post alloSCT at the original alloSCT time point of 12 weeks, the resultant ICERs are £45,033 and £50,353 for cohort 1 and 2, respectively. It is to be expected that the new week 12 model ICERs would increase from the original week 12 base case ICER since there has been the introduction of a progressed disease health state post alloSCT incurring lower utility versus the original 12 week model alive and dead post alloSCT. The fact that the inclusion of a progressed disease health state affects the ICER by only ~£2,000/QALY should provide increased confidence in the model estimates.

Several sensitivity analyses were performed to assess the impact of variation in all variables and assumptions applied within the model. The deterministic analysis and PSA showed pembrolizumab to be cost-effective in the many of the scenarios at a WTP threshold of £50,000/QALY (~53% and ~47% probability of being cost effective for cohort 1 and 2 respectively). In addition, many scenario analyses were tested and found to produce ICERs of less than £50,000/QALY.

The new 24 week time point model also estimates that in those who cannot undergo an alloSCT, the benefits of pembrolizumab allow the time in progression free disease to be increased vs SoC (mean 8.18 and 4.16 months vs 0.30 and 0.31 months SoC in cohorts 1 and 2 respectively). The overall response rates (CR, PR and SD) for pembrolizumab are 79.4% and 67.4% vs 23.9% and 23.4% for SoC in cohorts 1 and 2, respectively. This indicates a considerable benefit to survival in cohorts 1 and 2 vs SoC, both in the ability to remain in progression free disease for longer if alloSCT is not an option and to be more likely to achieve an adequate response to allow an alloSCT associated with higher long term survival. These values are similar to the original company base case and for the new 12 week model.

It should also be noted that at the time of this analysis and the original company submission, the number of patients that progressed in KEYNOTE-087 was judged to be too small to support robust analysis of post-progression survival. In the absence of data, external literature sources that reported mortality in patients with RRcHL who had progressed on treatment were identified from the clinical systematic review. As explained in the company submission, this was found to be Cheah et al, which reported a median OS of 25.2 months. The committee appraising TA462 concluded that in the UK, this estimate was likely to be lower.

In the economic analysis presented here and in the company submission, post-progression survival (PPS) was assumed constant due to both a lack of data to model a time dependent PPS and for simplicity, as it removed the need of tracking patients within the Markov model. However,

despite this limitation the predicted OS had a good level of face validity when compared to the observed OS from the Cheah study. For both weeks 24 and 12, after the maximum follow-up of Cheah (72 months) approximately 15% of patients were alive, which was correlated to the predicted SoC OS in the model at 72 months of approximately 15% in both cohorts 1 and 2.

The assumption that there was no post-progression survival benefit was conservative given the current OS rates of [redacted]% from KEYNOTE-087 cohort 1 and 2, respectively versus ~56% in Cheah at approximately 20 months. Despite some potential slight imbalances in the populations this extensive difference in the observed survival cannot be dismissed.

In regard to the differences observed between TA462 and the economic analyses presented here and in the original company submission, the model used in TA462 attributed a large proportion of its survival benefit of nivolumab compared to SoC on extrapolated OS directly from the pivotal trial in which the data were very immature, the committee were concerned about this as noted in the FAD. The extrapolation of OS data past trial follow up would have predicted a significantly higher number of total life years than the current approach, which conservatively assumed an equivalent post progression survival between both arms estimated from the SoC publication. The extrapolation of KEYNOTE-087 OS data was not deemed appropriate for this submission since the number of deaths were too low for the data to be considered mature enough to extrapolate in this way hence a different and more conservative model method was applied. In addition, the SoC treatment costs in TA462 were also assumed to be higher than those used in this submission (£10,477 vs ~£2,000). The use of the same SoC treatment costs in this economic analysis would also have produced a lower ICER.

Overall, the reduction in ICER at week 24 and marginal increase in week 12 versus the original company base-case should provide increased confidence in the model estimates. Although complex, the reduction in ICER at week 24 is, in a large part be due to the fact that patients on SoC progress more quickly than those on pembrolizumab and hence at week 24 there are likely to be less eligible SoC patients to transplant versus week 12 which was originally used as a conservative estimate of the two data collection points in KEYNOTE-087. In all likelihood, it is expected that the true transplant time-point will vary patient to patient and the average will lie somewhere between the two.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]

Dear [redacted],

The Evidence Review Group, Kleijnen Systematic Reviews, and the technical team at NICE have looked at the submission received on 05 September 2017 from Merck Sharp & Dohme UK Ltd. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **19 October 2017**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals [https://appraisals.nice.org.uk/request/34753 on 'NICE Docs/Appraisals'].

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Thomas Walker, Technical Lead (Thomas.walker@nice.org.uk). Any procedural questions should be addressed to Donna Barnes, Project Manager (Donna.Barnes@nice.org.uk).

Yours sincerely

Janet Robertson
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Literature searching

A1. *Company submission page 47 figure 3:* The search flow diagram depicted in section 4.1.4, figure 3 does not appear to include the 112 hits reported for the WHO ICTRP search. Please confirm if these were screened at this stage.

Included and excluded studies

A2. The systematic review in the company submission excluded studies based on the intervention not reflecting UK practice (as described on page 42). Which treatments were taken to reflect UK practice and how was this determined?

A3. *Company submission page 43 table 6:* For the purposes of the systematic review, how was best supportive care defined?

The KEYNOTE-087 study

A4. **Priority question**: Please provide the full clinical study reports for June 2016 and all updates. The referenced clinical study report is a summary version only.

A5. **Priority question**: Efficacy data in the company submission is from a cut-off date of 21 March 2017 and safety data is from a cut-off date of 25 Sept 2016. Please confirm that these data are the latest available from the trial.

A6. When are further overall and progression free survival data from this trial expected to be available?

A7. *Company submission page 160:* The company submission states that 'KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2'. Please explain this statement. Is the population in the trial representative of a UK population who would be suitable for allogeneic stem cell transplant?

A8. Please provide a study flow chart of KEYNOTE-087 encompassing the information on pages 63 and 64 of the company submission, including reasons for discontinuation of treatment, trial discontinuation and any subsequent allogeneic stem cell transplant.

A9. *Company submission page 91:* The company submission states that patients in Cheah et al. 2016[1] had received a median of 3 (range 0-9) prior lines of therapy

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

before commencing treatment with brentuximab vedotin, compared to a median of 4 (range 1-12), including treatment with brentuximab vedotin, in Cohorts 1 and 2 from KEYNOTE-087. Please provide a histogram of the number of prior therapies in KEYNOTE-087 and Cheah et al., 2016. Please also comment on how this compares to UK practice.

A10. **Priority question**: Please provide a break-down for KEYNOTE-087 of the reasons why people did not receive autologous stem cell transplant prior to trial entry, and please explain how many people in KEYNOTE-087 were autologous stem cell transplant ineligible in each cohort.

A11. **Priority question** *Company submission page 74*: Progression free survival is defined as time from first dose (of pembrolizumab) to disease progression, or death, whichever occurs first. Please clarify the time period between earlier treatment failure with brentuximab vedotin to first dose with pembrolizumab and how variable this was between patients in the KEYNOTE-087 trial.

A12. Did any patients who achieved a complete response in the trial subsequently progress and require retreatment with pembrolizumab?

The Cheah 2016 study

A13. **Priority question**: Cheah et al. includes patients who received investigational agents, which could include pembrolizumab. Therefore:

a. Please clarify how many of the patients receiving investigational agents received pembrolizumab. If possible, please provide a break-down of the investigational agents used in this study.

b. Please do an analysis (for all outcomes using the naïve and MAIC analyses) in which patients receiving pembrolizumab in Cheah et al. are excluded. Or, if this number is not known, please exclude all patients receiving investigational agents.

A14. *Company submission appendix 8:* There appears to be a discrepancy between data presented in appendix 8 of the company submission and the Cheah et al. publication. Please confirm the sample size from Cheah et al. that was used for matching was 83, as stated in the tables in Appendix 8.

A15. From the Cheah et al. study sample used for matching (appendix 8), how many patients had received prior autologous stem cell transplant? How many of those who did not receive a transplant were not eligible for the procedure?

A16. **Priority question**:

a. Did you consider creating two cohorts from the Cheah et al. study to reflect the two cohorts within the KEYNOTE-087 study?

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. If possible, please provide separate cohort analyses, using separate data from the two cohorts in the Cheah et al. study.

Indirect and mixed treatment comparisons

A17. **Priority question** *company submission page 93*: The company submission states that 'Weights were estimated for participants from KEYNOTE-087 so that their weighted mean baseline characteristics matched those observed in each of the comparator studies featured in the pairwise comparisons'. Please specify which comparator studies were used for which pairwise comparison.

A18. **Priority question**: Please provide full individual patient data (IPD) data necessary to perform the MAIC in a format that can be entered in R (with all necessary R packages), including all variables for which data were available in both KEYNOTE-087 and Cheah et al. 2016.

A19. **Priority question**: Please explain why 12 weeks was chosen as a time point for progression free survival. Is there a clinical rationale for this?

A20. *Company submission page 96:* In table 28, for cohort 1 please check the upper 95% CI of the hazard ratio for the naïve comparison as this seem too high; [redacted].

Further information

A21. *Company submission page 102:* In section 4.11.1 a UK Clinician Survey is described. Please provide full data from the survey, including individual results for each question and number of participants that contributed to each question/outcome.

Section B: Clarification on cost-effectiveness data

Please note that some of the following questions may request changes/amendments to the base case model, which should be addressed before any of the further scenario analyses requested in this section are undertaken. These are: B5, B15, B16, B21, B25, B27.

Literature searching

B1. *Company submission appendix 12:* Regarding the Medline/Embase strategy reported in Appendix 12, please clarify if this was a single search conducted simultaneously over both the Embase and Medline individual databases or was it a single search of Embase conducted on the understanding that it now contains all records from Medline.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B2. *Company submission page 135 and appendix 12:* The table of searched databases in Appendix 12 includes a search of EconLit. The table and the search flow diagram presented in in section 5.1.2 for identified cost-effectiveness studies reports no hits for EconLit; however, the search flows in both section 5.4.3 (figure 28 page 187) and 5.5.2 (figure 29 page 198) report 34 studies identified from EconLit. Please provide the full strategy, or strategies, used to search EconLit.

B3. *Company submission page 131:* Further additional searches are reported in section 5.1.1 for the ASCO, ESMO and ISPOR conference proceedings, as well as an additional search of the NICE website. However these searches are not reported in any of the search flow diagrams in sections 5.1.2, 5.4.3 or 5.5.2 and no strategies are reported in Appendix 12. Please provide full details.

Model structure

B4. **Priority question**: In the model structure, it is assumed that patients can only receive allogeneic stem cell transplant at 12 weeks, and no other time points. It is further assumed that the probability of allogeneic stem cell transplant is conditional on complete response (CR)/partial response (PR)/stable disease (SD)/progressive disease (PD).

a. Please justify this assumption given that the time to response in KEYNOTE-087 ranged between [redacted] and [redacted] months and had a mean of [redacted] weeks (that is, it can differ substantially from the 12-week time point, and is, in fact not accounting for half of the patients receiving allogeneic stem cell transplant) and discuss the potential bias caused by not allowing allogeneic stem cell transplant at earlier and/or later time points.

b. Given the assumption that allogeneic stem cell transplant would occur at 12 weeks and that the median follow-up in KEYNOTE-087 is approximately 16 months, it should have been possible to estimate allogeneic stem cell transplant rates directly based on KEYNOTE-087. Please explain why this was not done.

c. The introduction of the decision tree node at 12 weeks necessitates further assumptions, including separate estimation of pre- and post-12 week relative effectiveness. Please comment on uncertainty and potential bias introduced by using this model structure.

d. Please provide a scenario analysis incorporating time to allogeneic stem cell transplant for pembrolizumab directly estimated from KEYNOTE-087 into the state transition model, removing the decision nodes at week 12. If you believe that allogeneic stem cell transplant probabilities derived from KEYNOTE-087

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

are an under-estimate, these could be calibrated to reflect UK practice, for example by using clinical expert opinion.

Treatment effectiveness

B5. **Priority question:** There are inconsistencies in the distribution of patients across the different response states in the company submission.

a. Please provide a corrected version of Table 62 from the company submission.

b. The estimation of the distribution of patients in the different response states appears incorrect. In particular, the proportions in the stable disease state in both cohorts in the model are significantly larger than those observed in the KEYNOTE-087 study, as reported in Table 19 of the company submission. This results in significant over-estimation of the numbers of people having allogeneic stem cell transplant in both cohorts. Please submit a version of the model in which response rates are corrected and explain the correct proportion of responders, and how they have been calculated, referring to the clinical part of the company submission.

B6. **Priority question** *company submission page 162***:** The company submission states that for progression free survival, parametric models were fitted to the observed data in KEYNOTE-087 from week 12. However, in the model file it appears that parametric models were fitted to all the data, beginning from week 0.

a. Please clarify what was done.

b. Please provide progression free survival and overall survival curves fitted to the entire study data for both cohorts and include them in the model.

c. Please estimate comparative effectiveness using the entire study data (from week 0 to end of follow-up).

B7. *Company submission page 161:* In the clinician survey on the probabilities of allogeneic stem cell transplant conditional on response status, some clinicians had indicated that even with progressed disease patients could be eligible for allogeneic stem cell transplant.

a. Please explain why this was not incorporated in the model.

b. Please provide a scenario analysis in which patients in the progressed disease health state could also receive allogeneic stem cell transplant.

c. Please provide the data used from the previously conducted clinician survey by BMS (for nivolumab) and explain the methods of how the two were

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

combined – and how it was ensured that the same experts were not reflected twice.

d. Please perform a scenario analysis in which only the data from the new survey is used.

B8. **Priority question** *company submission page 141***:** The hazard ratio for progression free survival post-12 weeks is the single most important parameter as per the company's deterministic sensitivity analyses. However, it is based on a strong assumption: it was assumed to be the same as in the pre-12 weeks period because *"a PFS HR from week 12 to end of follow-up could not be estimated given the low number of events post week 12 observed in Cheah"* (page 141). This suggests that the HR in the pre-12 week period is potentially an over-estimate of comparative effectiveness of pembrolizumab versus standard of care.

a. Please comment on the potential bias this introduces, given that most events in Cheah et al. (2016)[1] had already happened in the first 12 weeks and that therefore the HR after 12 weeks would be expected to be much less in favour of pembrolizumab – and also in light of choosing different distributions for modelling pre- and post-12 week PFS.

b. Please state whether the use of any alternatives to using the same HR as in the pre-12 week period were considered. If any alternative were considered, please provide details and an explanation of why they were not used.

c. Please provide one scenario analysis (both cohorts) using HR = 1 for the post-12 week period and another using the estimated HR from the indirect treatment comparison (ITC), reflecting the uncertainty associated with it.

B9. *Company submission figures 17 and 21:* In cohort 2, patients progress significantly faster than in cohort 1. Please explain what causes patients in cohort 2 to start progressing at approximately [redacted] weeks (company submission figure 17 page 156) and for all of them to have progressed at approximately [redacted] months (company submission figure 21 page 167), as opposed to much longer estimates for time to progression in cohort 1?

B10. *Company submission page 169:* The calculation of post progression survival was based on data from Cheah et al. (instead of KEYNOTE-087). Please provide a scenario analysis using data from the KEYNOTE-087 study to calculate post progression survival.

B11. **Priority question:** Lafferty et al. 2017[2] is an abstract reporting on a single centre experience of 13 patients with Hodgkin Lymphoma who had allogeneic stem cell transplant. According to page 29 of the company submission, there were 23 patients in the KEYNOTE-087 study who had allogeneic stem cell transplant.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

a. Please explain why Lafferty et al. (2017) was preferred over the KEYNOTE-087 study to estimate post-allogeneic stem cell transplant overall survival (OS), proportions for complete response (CR)/partial response(PR)/stable disease (SD)/progressive disease (PD) as well as for acute graft vs host disease after allogeneic stem cell transplant.

b. Please estimate post-allogeneic stem cell transplant OS, proportions for CR/PR/SD/PD as well as for acute graft vs host disease after allogeneic stem cell transplant based on the KEYNOTE-087 study and use this in a scenario analysis.

c. Please comment on whether expert opinion was consulted to validate the OS estimates resulting from this study.

B12. In TA 462 investigational agents were excluded from the Cheah data analysis. Referring to Question A.13, please provide a scenario analysis in which patients receiving pembrolizumab, or if this is not possible investigational agents, are excluded.

B13. **Priority question:** Please provide a scenario analysis in which, in line with the marketing authorisation for pembrolizumab, patients can continue treatment after 24 months.

Adverse events

B14. *Company submission page 182 onwards*: Adverse event incidence rates for standard of care were derived from different sources reporting adverse events associated with chemotherapy regimens and bendamustine. Please provide more information on how the weighted average for standard of care was derived.

Health-related quality of life

B15. **Priority question** *company submission page 185 onwards***:** Long term utility values have been estimated from KEYNOTE-087 using observations from week 12 only. These utility values might not reflect long-term utility values.

a. Please use a mixed model based on utility data from KEYNOTE-087 to estimate utility values separately for complete response, partial response, stable disease and progressive disease, using all available observations from the KEYNOTE study for participants before they have allogeneic stem cell transplant.

b. Please use a mixed model based on utility data from KEYNOTE-087 to estimate utility values for post-allogeneic stem cell transplant, using all available observations from the KEYNOTE study for participants after they

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

have allogeneic stem cell transplant. If possible also provide an additional analysis distinguishing between <14 weeks and >14 weeks post-allogeneic stem cell transplant, or alternatively distinguishing between patients with and without acute graft vs host disease.

B16. Please explain why the proportions of responders (complete response and partial response) used to calculate utilities for the progression free health state (that is, used in calculations in cells G18 and H18 of the model; "NonClinicalData" worksheet) are different from the proportion of responders used to calculate the proportions of patients receiving allogeneic stem cell transplant at week 12 (that is, proportions in rows 38-41; "ClinicalData" worksheet). Please provide a corrected model, if necessary.

B17. The lack of a post-progression health state in the allogeneic stem cell transplant group potentially causes a bias in the estimation of quality of life.

a. The one-year progression free survival after allogeneic stem cell transplant reported by Lafferty et al. (2017)[2] was 54%. Please indicate how the occurrence of progression is reflected in quality of life in the post allogeneic stem cell transplant health state (from which patients cannot transit to a progressive disease health state). Although, in the model, patients have a disutility during the first 14 weeks post-allogeneic stem cell transplant, after this period patients are assumed to have a utility of 0.865 while a utility of [redacted] is assumed for the progressive disease health state.

b. Please explain why the allogeneic stem cell transplant utilities were assumed to be driven by complete response/partial response/stable disease (given that this utility is calculated based on the complete response/partial response/stable disease proportions and the 12-week utility values).

B18. Please explain why only the disutility for acute graft vs host disease was used from Kurosawa et al. (2015)[3] instead of using the utilities from this study for the post allogeneic stem cell transplant health state (that is, utility values of 0.65 and 0.80 for <14 weeks and >14 weeks post allogeneic stem cell transplant respectively).

Resource use and costs

B19. For the calculation of standard of care costs, it is assumed that all chemotherapy agents contribute an equal proportion of treatment to standard of care. Please justify this assumption.

B20. According to the company submission and the model file, no doses of pembrolizumab or standard or care were missed.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

a. Please provide a justification for this assumption and explain whether alternative assumptions were considered.

b. Please provide a scenario analysis considering missed doses.

B21. In the company submission it is stated that baseline cohort characteristics of the KEYNOTE-087 cohort are used in the model (in table 100). The model however shows a body surface area (BSA) of 1.85 was used which does not match with the mean BSA of the KEYNOTE-087 cohort (as reported in the company submission table 56 page 150). Please explain this discrepancy, and correct the model if necessary.

B22. *Company submission page 214*: In the model, one-off costs are applied upon treatment with allogeneic stem cell transplant. These costs are taken from Radford et al. (2017)[4], a study reporting on costs in 14 relapsed or refractory classical Hodgkin Lymphoma patients treated with allogeneic stem cell transplant, where mean follow-up was 3.44 years (the proportion of patients that died was not stated). Given the uncertainty on what proportion of patients died in Radford et al. (2017), it seems impossible to assess the assumption that terminal care costs are already included in the one-off cost. Furthermore, given the follow-up period of maximum 5 years and the fact that in the model 40% of patients treated with allogeneic stem cell transplant are alive at 10 years, it is questionable if the one-off cost accurately captures monitoring and subsequent treatment costs allogeneic stem cell transplant.

a. Please explain why a one-off cost was deemed appropriate over the use of individual costs for allogeneic stem cell transplant intervention and subsequent treatment, monitoring, adverse events and terminal care costs. Please also explain why these costs are assumed to be reflected by the one-off cost.
b. Please provide a scenario where the one-off cost for allogeneic stem cell transplant is replaced by costs accounting for the allogeneic stem cell transplant intervention and post- allogeneic stem cell transplant period separately.
c. It is stated in the company submission on page 209 that palliative care costs within best supportive care are set to £0 in line with assumptions made in TA462,[5] please provide an explanation why this assumption is appropriate for this appraisal.
d. According to the company submission on page 212, costs of terminal care come from a previous HTA assessment on non-small cell lung cancer that have been updated. Please provide a justification for why data from a different patient population was used and explain how the costs were updated. Furthermore, please clarify whether the inclusion of terminal care

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

costs results in double-counting of costs, given that patients receive treatment until they die.

Cost effectiveness results

B23. *Company submission page 220:* Please provide the UK standard of care clinical outcomes which appear to be missing from Table 102.

B24. Please provide the total LYs for each comparator obtained from the probabilistic sensitivity analysis.

B25. Please explain how the cost and effects of best supportive care were estimated in scenario analyses 1. Please provide a model file in which best supportive care can be selected as a comparator, and both best supportive care and standard of care can be included simultaneously in the probabilistic sensitivity analysis.

Model validation

B26. Please provide a detailed cross validation with TA462,[5] commenting on all differences in modelling assumptions, data sources used and model predictions, at least for standard of care, using publicly available data on the inputs and results from TA462.

B27. Please provide model files without hidden worksheets, rows and columns.

B28. Please confirm whether the model files for cohort 1 and cohort 2 are exactly the same.

References

[1] Cheah CY, Chihara D, Horowitz S, Sevin A, Oki Y, Zhou S, et al. Patients with classical Hodgkin lymphoma experiencing disease progression after treatment with brentuximab vedotin have poor outcomes. *Ann Oncol* 2016;27(7):1317-23.

[2] Lafferty N, Anandram S, Lawes N, Usman M, Liebersbach S, Benn K, et al. Allogeneic stem cell transplantaion in patients with Hodgkin Lymphoma: a retrospective single centre case series. *Br J Haematol.* 2017;176(Suppl 1):5-145.

[3] Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplant* 2015;50(9):1241-9.

[4] 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 Transplant* 2017;52(3):452-4.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

[5] National Institute for Health and Care Excellence. *Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma: NICE technology appraisal guidance TA462 [Internet]*. London: NICE, 2017 [accessed 18.9.17] Available from: https://www.nice.org.uk/guidance/ta462/resources/nivolumab-for-treating-relapsed-or-refractory-classical-hodgkin-lymphoma-pdf-82604902197445

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

46291Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]

Dear [redacted]

The Evidence Review Group, Kleijnen Systematic Reviews, and the technical team at NICE have looked at the submission received on 05 September 2017 from Merck Sharp & Dohme UK Ltd. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **19 October 2017**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals [https://appraisals.nice.org.uk/request/34753 on 'NICE Docs/Appraisals'].

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Thomas Walker, Technical Lead (Thomas.walker@nice.org.uk). Any procedural questions should be addressed to Donna Barnes, Project Manager (Donna.Barnes@nice.org.uk).

Yours sincerely

Janet Robertson
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Literature searching

A1. *Company submission page 47 figure 3:* The search flow diagram depicted in section 4.1.4, figure 3 does not appear to include the 112 hits reported for the WHO ICTRP search. Please confirm if these were screened at this stage.

The searches of the WHO ICTRP occurred at the same time as the update to the main database searches.

As reported in section 4.14 of the manufacturer submission: After searching the International Clinical Trials Registry platform for ongoing clinical trials, there were 112 hits captured (Search strategy and results can be found in Appendix 2. A total of 85 records were removed for being a duplicates, 13 for not being ongoing, and 13 were excluded based on population, resulting in 1 ongoing trial (NCT03077828) of interest. Full details of this study are reported on page 130 of the manufacturer's submission.

These were not included in the PRISMA as they were seen as providing contextual information as opposed to data that could be used to assess comparative efficacy or safety.

Included and excluded studies

A2. The systematic review in the company submission excluded studies based on the intervention not reflecting UK practice (as described on page 42). Which treatments were taken to reflect UK practice and how was this determined?

MSD included comparators listed in the NICE final scope (March 2017) considered to represent UK clinical practice. This comprises: single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin and best supportive care.

A3. *Company submission page 43 table 6:* For the purposes of the systematic review, how was best supportive care defined?

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The number of patients with rrcHL at this line of therapy are low, and in the absence of established clinical guidelines it was not possible to define best supportive care. Therefore, in an attempt to maximise the identification and inclusion of evidence no definition was specified for best supportive care and all eligible studies that reported this would have been included.

The KEYNOTE-087 study

A4. **Priority question**: Please provide the full clinical study reports for June 2016 and all updates. The referenced clinical study report is a summary version only.

MSD has now provided the full clinical study report for June 2016 [redacted] Please note that the summary update reports utilising September 2016 and March 2017 data are not available as full documents as these were created at the requested of regulatory bodies.

A5. **Priority question**: Efficacy data in the company submission is from a cut-off date of 21 March 2017 and safety data is from a cut-off date of 25 Sept 2016. Please confirm that these data are the latest available from the trial.

This is correct. As per discussions with MSD and NICE in March 2017, an updated efficacy report (database lock 21st March 2017) was used within the submission. Please note that an updated safety report was not conducted using the March database lock, and for this reason MSD has presented safety data based on the September 2016 database lock.

A6. When are further overall and progression free survival data from this trial expected to be available?

[redacted]

A7. *Company submission page 160:* The company submission states that 'KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2'. Please explain this statement. Is the population in the trial representative of a UK population who would be suitable for allogeneic stem cell transplant?

The CSR (June 2016) for KEYNOTE-087 provides full details relating to study design, study objectives, and exploratory endpoints; these did not include the subsequent

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

investigation of patients treated with pembrolizumab who were then treated with a stem cell transplant.

The use of stem cell transplant would have been at the discretion of the treating physician on a per patient basis. Stem cell transplant data were captured as a result of the follow up described within the CSR, which states:

"Subjects were treated with pembrolizumab 200 mg Q3W until documented disease progression, unacceptable adverse event(s), intercurrent illness that prevents further administration of treatment, investigator's decision to withdraw the subject, subject withdraws consent, pregnancy of the subject, noncompliance with trial treatment or procedure requirements, or administrative reasons".

"After the end of treatment, each subject was followed for 30 days for adverse event monitoring (serious adverse events [SAE] and events of clinical interest [ECI] were collected for 90 days after the end of treatment). Subjects who discontinued treatment for reasons other than disease progression had post-treatment follow-up for disease status until disease progression, initiating a non-study cancer treatment, withdrawing consent, or becoming lost to follow-up. All subjects were followed by telephone contact for overall survival until death, withdrawal of consent or the end of the study, whichever came first".

As described in Section 3.5 of the submission patients with rrcHL at this stage of the care pathway patients are heavily pre-treated. The recent recommendation of Brentuximab Vedotin enables the patient population under consideration within this submission to be treated with pembrolizumab. Note that all enrolled patients of KEYNOTE-087 had received prior treatment with Brentuximab Vedotin, a requirement for inclusion; and therefore, could be considered to represent the current population as treated within UK clinical practice. Furthermore, this was validated with a UK clinician, who confirmed the relevance of this population in relation to UK clinical practice.

As reported in Section 4.3 of the submission, UK patient numbers were low (n=14). Therefore, to support feedback received from UK clinicians (MSD held advisory board), and take into consideration the recent recommendation of Nivolumab TA462 that reported *"the committee understood from the clinical experts and patient organisations that nivolumab had the potential to act as salvage therapy to enable allogeneic stem cell transplant after both autologous stem cell transplant and Brentuximab Vedotin"* MSD reported the number of patients within a UK setting who received allogenic stem cell transplant following treatment with pembrolizumab.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A8. Please provide a study flow chart of KEYNOTE-087 encompassing the information on pages 63 and 64 of the company submission, including reasons for discontinuation of treatment, trial discontinuation and any subsequent allogeneic stem cell transplant.

As per the subject disposition table in the March 2017 efficacy update report [redacted] recorded to have discontinued treatment due to bone marrow transplant. [redacted].

A9. *Company submission page 91:* The company submission states that patients in Cheah et al. 2016[1] had received a median of 3 (range 0-9) prior lines of therapy before commencing treatment with brentuximab vedotin, compared to a median of 4 (range 1-12), including treatment with brentuximab vedotin, in Cohorts 1 and 2 from KEYNOTE-087. Please provide a histogram of the number of prior therapies in KEYNOTE-087 and Cheah et al., 2016. Please also comment on how this compares to UK practice.

Please find a baseline characteristics table for prior lines of therapy below. This reports the frequency of patients reporting 1-12 prior lines of therapy. The distribution of prior treatment in the Cheah et al., 2016 study cannot be presented graphically as MSD do not have access to the patient level data.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

[redacted]

	COHORT 1		COHORT 2	
	n	(%)	n	(%)
Subjects in population	69		81	
Prior Lines of Therapy				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
(Database Cutoff Date: 25SEP2016).				

MSD sought clarity on the treatment experience of RRcHL within the UK, and was advised that typically patients within the UK would have received between 3 and 4 prior lines of therapy, including Brentuximab Vedotin, before commencing therapy with a PD-L1.

MSD was informed that patients would receive: 1) first line therapy chemotherapy, 2) second line therapy following relapse, which may or may not include stem cell transplant, 3) potentially a second salvage therapy to allow for transplant, or 4) Brentuximab Vedotin, 5) PD-L1 (this is aligned with positioning of pembrolizumab after BV and the KEYNOTE-087 population reporting ~3 or 4 lines of prior therapy).

A10. **Priority question**: Please provide a break-down for KEYNOTE-087 of the reasons why people did not receive autologous stem cell transplant prior to trial entry, and please explain how many people in KEYNOTE-087 were autologous stem cell transplant ineligible in each cohort.

This information is not reported within the company CSR. However, it was possible to use INFORM (a data capture system) to check for these data. It should be noted that data collected within this system is not used to inform regulatory agencies, generate

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

publications or answer external queries. Therefore the following information should be considered as an indicative response.

[redacted]

The above were validated with a UK clinician.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A11. **Priority question** *Company submission page 74*: Progression free survival is defined as time from first dose (of pembrolizumab) to disease progression, or death, whichever occurs first. Please clarify the time period between earlier treatment failure with brentuximab vedotin to first dose with pembrolizumab and how variable this was between patients in the KEYNOTE-087 trial.

Summarised in Table 1 are the baseline characteristics of KEYNOTE-087 Cohort 1 and 2 for the duration of time since last dose of Brentuximab Vedotin and first dose of pembrolizumab.

[redacted]

	COHORT 1	COHORT 2
Subjects in population	[redacted]	[redacted]
[redacted]		
Subjects with data	[redacted]	[redacted]
Mean	[redacted]	[redacted]
SD	[redacted]	[redacted]
Median	[redacted]	[redacted]
Range	[redacted]	[redacted]
(Database Cutoff Date: 25SEP2016).		

A12. Did any patients who achieved a complete response in the trial subsequently progress and require retreatment with pembrolizumab?

Using the September database lock 2016, based on subject disposition, [redacted] who started a 2nd course of pembrolizumab.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The Cheah 2016 study

In response to questions below, please note that the Cheah et al. 2016 publication was identified from a SLR and was not conducted by MSD. Further, MSD did not have access to patient level data from this study. This precludes a number of requested analyses throughout our response to clarification questions.

A13. **Priority question**: Cheah et al. includes patients who received investigational agents, which could include pembrolizumab. Therefore:

a. Please clarify how many of the patients receiving investigational agents received pembrolizumab. If possible, please provide a break-down of the investigational agents used in this study.

As reported in the Nivolumab clarification question A16; specific intervention subgroup data for Cheah et al. 2016 was not available. The manufacturer reported the impact of removing investigational agents (n=28) from the overall Cheah et al. population resulted in the following outcomes:

- The CR results for the whole study population was 15.2% (12/79) vs. the removal of investigational agents (n=28) left 8/51 at 15.7%.
- PR also increased with the removal of investigational agents to 23.5% (12/51) vs. 19% (15/79).
- Median PFS for the overall population was 3.5 months, investigational agents reported a PFS of 2.4 months; therefore, removal of this subgroup would increase the PFS overall.

Within the Southampton ERG report (18th Jan 2017) the authors of the Cheah et al. study confirmed *"that only a couple of patients in the study received a PD-1 inhibitor"; furthermore, the report states".* The ERG was not aware of a more appropriate data source for the comparator population".

MSD chose to present a conservative comparative effectiveness estimate that did not exclude patients who had been treated with investigational agents.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. Please do an analysis (for all outcomes using the naïve and MAIC analyses) in which patients receiving pembrolizumab in Cheah et al. are excluded. Or, if this number is not known, please exclude all patients receiving investigational agents.

As per question A13 investigational agents did not include pembrolizumab. As MSD does not have access to patient level data, it was not possible to conduct analyses excluding a PD-1 inhibitor, which may have related to a couple of patients within the investigation agents group.

In addition, the KM curve for PFS is based on data from "all patients" and therefore a subgroup analysis removing patients receiving investigational agents could not be conducted. However, it was possible to do an analysis of response removing patients who received investigational agents. The results of this are presented in Table 1 and Table 2.

Table 1: Summary of comparisons of objective response rates (best overall response) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 2: Summary of comparisons of objective response rates (12-weeks) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

Table 3: Summary of comparisons of complete response (best overall response) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 4: Summary of comparisons of complete response (12-weeks) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

Table 5: Summary of comparisons of partial response (best overall response) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 6: Summary of comparisons of partial response (12-weeks) for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size/effective sample size, n	ORR with pembrolizumab	ORR with SOC	Odds ratio (95% CI)
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
1 and 2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; ORR, objective response rate; SOC, standard of care

A14. *Company submission appendix 8:* There appears to be a discrepancy between data presented in appendix 8 of the company submission and the Cheah et al. publication. Please confirm the sample size from Cheah et al. that was used for matching was 83, as stated in the tables in Appendix 8.

Correct, the discrepancy is within the values reported by Cheah et al. and not the manufacturer submission. The n=83 referenced in Appendix 8 is the number of patients with data regarding treatment at disease progression following BV. As baseline characteristics for this cohort specifically are not presented, the distribution of characteristics from all patients with available data at the time of documented progression following treatment with BV were applied and used for matching (note that the number of patients with data varied from 89 for age to 46 for lymphocyte count). In addition, the proportion of females and median number of prior lines of therapy were only presented for patients before commencement of BV, therefore these values were again assumed to also apply to the post-BV treatment cohort.

A15. From the Cheah et al. study sample used for matching (appendix 8), how many patients had received prior autologous stem cell transplant? How many of those who did not receive a transplant were not eligible for the procedure?

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

These data are not available within the Cheah et al. publication. The available information for the 97 patients who met the study inclusion criteria reported a median PFS following initial therapy of 10 months (range 0-106 months).

The author's state that at second remission, autologous stem cell transplant occurred in 66/97 patients (68%), and four patients (4%) received an allogeneic stem cell transplant.

The publication states *"at the time of second remission, ASCT was carried out in 66 (68%) patients and alloSCT in 4 (4%) patients. Of the remaining 27 patients who did not undergo consolidative transplant, the primary reason was failure to respond to therapy (n = 21, 75%), age or co-morbidities (n = 1, 4%), failed mobilization (n = 1, 4%), patient decision (n = 1, 4%), financial reasons (n = 1, 4%) or reason unknown (n=2, 7%). The median PFS following initial stem cell transplant was 6.6 months (range 1–67); a further 10 patients underwent alloSCT after ASCT, but before treatment with BV. Thus, in total, 14 (14%) patients underwent alloSCT before receiving BV. Details regarding the outcome of the last therapy before BV were available in 84 patients, of whom 31 (36%) were refractory".*

A16. **Priority question**:

a. Did you consider creating two cohorts from the Cheah et al. study to reflect the two cohorts within the KEYNOTE-087 study?

MSD do not have access to individual patient level data for Cheah et al. and therefore it was not possible to determine cohorts using the same inclusion criteria as were applied to cohorts 1 and 2 in KEYNOTE-087.

b. If possible, please provide separate cohort analyses, using separate data from the two cohorts in the Cheah et al. study.

This was not possible for the reasons noted in part a.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Indirect and mixed treatment comparisons

A17. **Priority question** *company submission page 93*: The company submission states that 'Weights were estimated for participants from KEYNOTE-087 so that their weighted mean baseline characteristics matched those observed in each of the comparator studies featured in the pairwise comparisons'. Please specify which comparator studies were used for which pairwise comparison.

Apologies, the methods text was written before it was clear that data would only be available from this study, therefore MSD anticipated that it may be necessary to carry out multiple pairwise comparisons with studies featuring interventions contained in the final scope.

However, as per the PRISMA flow diagram and evidence presented only one study was used to conduct a pairwise comparison with KEYNOTE-087: Cheah et al., 2016

A18. **Priority question**: Please provide full individual patient data (IPD) data necessary to perform the MAIC in a format that can be entered in R (with all necessary R packages), including all variables for which data were available in both KEYNOTE-087 and Cheah et al. 2016.

MSD has uploaded an [redacted] data along with the associated R code. Please note that MSD does not have access to IPD for Cheah et al. 2016, and therefore cannot provide this.

A19. **Priority question**: Please explain why 12 weeks was chosen as a time point for progression free survival. Is there a clinical rationale for this?

Please see the response provided B4. *This reflects the current care pathway as reported in the UK, which has been validated by an MSD advisory board.*

A20. *Company submission page 96:* In table 28, for cohort 1 please check the upper 95% CI of the hazard ratio for the naïve comparison as this seem too high; [redacted]

Apologies, this was an error in reporting, the value for the HR (95% CI) is: [redacted]

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Further information

A21. *Company submission page 102:* In section 4.11.1 a UK Clinician Survey is described. Please provide full data from the survey, including individual results for each question and number of participants that contributed to each question/outcome.

The results slide deck has been uploaded to NICE.doc. Each slide reports the: question asked, results, and the effective sample size for each question. As reported within the STA submission document the numbers of patients treated by the clinicians included was small.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Section B: Clarification on cost-effectiveness data

Please note that some of the following questions may request changes/amendments to the base case model, which should be addressed before any of the further scenario analyses requested in this section are undertaken. These are: B5, B15, B16, B21, B25, B27.

Please note that all scenario analyses presented in the clarification document are based on the original base case model, submitted by MSD, with the following addition:

- B15a – inclusion of results of mixed modelling of utilities by response status in KEYNOTE-087
- B27 – all worksheets, rows and columns unhidden

Full responses to queries B5, B15b, B16, B21, B25, and B27 are provided in later sections of the document. No further changes to the base case model were made due to either data limitations or because no errors were identified.

Literature searching

B1. *Company submission appendix 12:* Regarding the Medline/Embase strategy reported in Appendix 12, please clarify if this was a single search conducted simultaneously over both the Embase and Medline individual databases or was it a single search of Embase conducted on the understanding that it now contains all records from Medline.

The first search strategy covers evidence from both Embase and Medline using the embase.com interface. This search strategy is presented on page 111 through page 114, (Title: "Search strategy for Embase® and MEDLINE® database for economic and utility review (searched via Embase.com on 12th July 2017)")

Additionally, the MEDLINE® In-Process database was searched to ensure that non-indexed citations were retrieved using PubMed platform. This search strategy is provided on page 114 through page 119 (Title: "Search strategy for MEDLINE® In-Process searched via PubMed® platform for economic and utility review (as searched on 12th July 2017)")

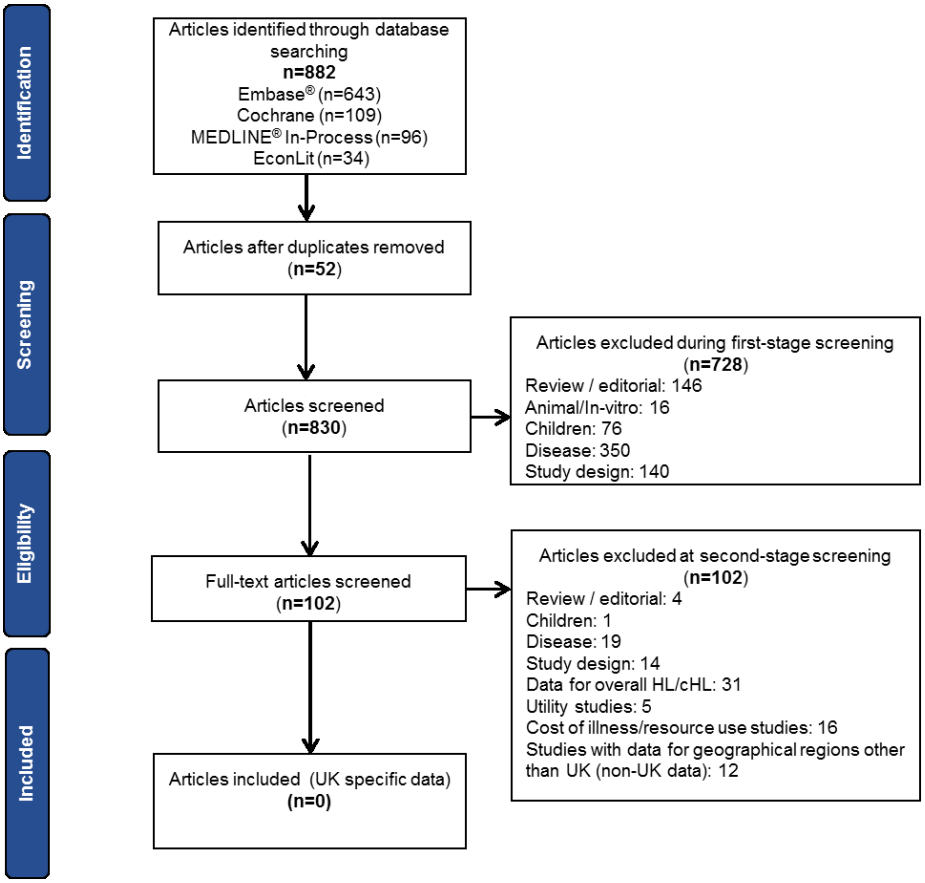
B2. *Company submission page 135 and appendix 12:* The table of searched databases in Appendix 12 includes a search of EconLit. The table and the search flow diagram presented in in section 5.1.2 for identified cost-effectiveness studies reports no hits

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

for EconLit; however, the search flows in both section 5.4.3 (figure 28 page 187) and 5.5.2 (figure 29 page 198) report 34 studies identified from EconLit. Please provide the full strategy, or strategies, used to search EconLit.



Search strategy for EconLit is attached. In brief, this search was based on indication of interest i.e. Hodgkin's lymphoma, irrespective of any specific review type and yielded 34 search hits. This search hit number is specified in PRISMA flow of all the reviews.

B3. *Company submission page 131:* Further additional searches are reported in section 5.1.1 for the ASCO, ESMO and ISPOR conference proceedings, as well as an additional search of the NICE website. However these searches are not reported in any of the search flow diagrams in sections 5.1.2, 5.4.3 or 5.5.2 and no strategies are reported in Appendix 12. Please provide full details.

The conference proceeding and NICE websites were hand-searched using specific keywords and the associated abstracts were screened for inclusion. The keywords used for each conference proceeding is listed in the excel file provided in response to B2.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Model structure

B4. **Priority question**: In the model structure, it is assumed that patients can only receive allogeneic stem cell transplant at 12 weeks, and no other time points. It is further assumed that the probability of allogeneic stem cell transplant is conditional on complete response (CR)/partial response (PR)/stable disease (SD)/progressive disease (PD).

a. Please justify this assumption given that the time to response in KEYNOTE-087 ranged between [redacted] and [redacted] months and had a mean of [redacted] weeks (that is, it can differ substantially from the 12-week time point, and is, in fact not accounting for half of the patients receiving allogeneic stem cell transplant) and discuss the potential bias caused by not allowing allogeneic stem cell transplant at earlier and/or later time points.

The model was designed to represent the current clinical care pathway for RRcHL in the UK as closely as possible within the limited data available and may be considered a simplistic representation of what would occur in UK clinical practice. MSD accepts that while for some patients, response to Pembrolizumab may be achieved earlier or later than week 12, the chosen time point in the model conservatively assumes that those with a response after this time point would not be considered for alloSCT use. The extent to which pembrolizumab would be used differently to standard of care in UK clinical practice is also not currently known.

The 12 week time point for assessing the uptake of alloSCT was based on the results of the UK clinician survey (median time to SCT is approximately 12 weeks). This was further validated at an advisory board meeting and discussions with clinicians where UK clinicians stated that patients would be transplanted as soon as they showed a CR or PR.

AlloSCT data from KENOYTE-087, summarised in response to question d), further indicate that most transplants occur within the first 6 months, with treatment lasting on average [redacted] weeks prior to transplant. This indicates that the decision to undergo alloSCT is taken around the 12 week assessment, but because of the time required to identify a donor and schedule the procedure, SCT may be performed after week 12 and before week 24. The use of a 12 week decision node provides an accurate representation of the timing of decisions to transplant, and from an economic viewpoint adequately reflects the time at which treatment was stopped in KEYNOTE-087. The exact timing of each individual transplantation procedure in the first year of the analysis is unlikely to significantly bias results over a lifetime analysis.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. Given the assumption that allogeneic stem cell transplant would occur at 12 weeks and that the median follow-up in KEYNOTE-087 is approximately 16 months, it should have been possible to estimate allogeneic stem cell transplant rates directly based on KEYNOTE-087. Please explain why this was not done.

This was not possible for the following reasons. KEYNOTE-87 was a predominantly non-UK study and not set up to investigate alloSCT. Expert testimony given during TA462 indicated that the UK rates of stem cell transplant were much higher than in other geographies, namely, the US. This is illustrated by higher rates of alloSCT among UK patients within KEYNOTE-087 cohort 1 [redacted] compared with the overall population [redacted] (March 2017 data cut). With limited UK patient numbers it is difficult to draw robust conclusions from the KEYNOTE-087 rates of alloSCT and hence the proportions of patients that received alloSCT conditional on response (CR/PR/SD) were estimated from clinician surveys.

c. The introduction of the decision tree node at 12 weeks necessitates further assumptions, including separate estimation of pre- and post-12 week relative effectiveness. Please comment on uncertainty and potential bias introduced by using this model structure.

The model provides the flexibility to vary treatment effect across the first 12 weeks and the subsequent time period. This structure does not necessitate the use of different comparative effects across time periods. For example, if the proportional hazards assumption holds, then the same effect may apply both prior to and after the 12 week decision node. This is the assumption applied in the base case, and is a common assumption used across many oncology submissions.

The model structure was developed to accurately capture the uptake of alloSCT as a function of response, based on UK specific alloSCT rates and following UK practice for alloSCT use in RRcHL. As detailed in B4a, experts state that alloSCT use would occur around week 12 in UK practice dependent on response. This assumption has been applied to both arms of the analysis, and is not expected to significantly bias the incremental cost-effectiveness results.

If data were available, or existing data could be calibrated (which in itself is uncertain and at risk of bias), the modelling of alloSCT use over time (removing the 12 week decision node) would significantly increase the complexity of the analysis by requiring the use of tunnel states or patient-level simulation modelling to capture time-varying survival rates post-alloSCT.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

RRcHL is a rare disease with limited data that prohibits the use of complex modelling techniques. In light of the uncertainties, a simpler decision-tree approach was adopted.

As described in B4a, MSD believes the use of a 12 week decision node provides an accurate representation of the timing of decisions to transplant and that the exact timing of each individual transplantation procedure in the first year of the analysis is unlikely to significantly bias results over a lifetime analysis.

d. Please provide a scenario analysis incorporating time to allogeneic stem cell transplant for pembrolizumab directly estimated from KEYNOTE-087 into the state transition model, removing the decision nodes at week 12. If you believe that allogeneic stem cell transplant probabilities derived from KEYNOTE-087 are an under-estimate, these could be calibrated to reflect UK practice, for example by using clinical expert opinion.

Kaplan-Meier plots of the time from first dose to alloSCT in KEYNOTE-087 are provided in Figure 1 (cohorts 1 and 2 combined),

Figure 2 (cohort 1) and

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Figure 3 (cohort 2).

A scenario analysis that incorporates these data in the model could not be performed for the following reasons:

- As outlined in response to question B4 b), the alloSCT rates in KEYNOTE-087 significantly underestimates the expected alloSCT rates in UK clinical practice.
- The limited UK patient numbers precludes any robust analysis of time to alloSCT for a UK population.
- The manufacturer does not have access to Kaplan-Meier data on the time to alloSCT from Cheah et al to compare transplantation probabilities between pembrolizumab and standard of care.

Figure 1: Kaplan-Meier plot for time from first dose to allogeneic stem cell transplantation in the combined cohorts 1 and 2 of KEYNOTE-087

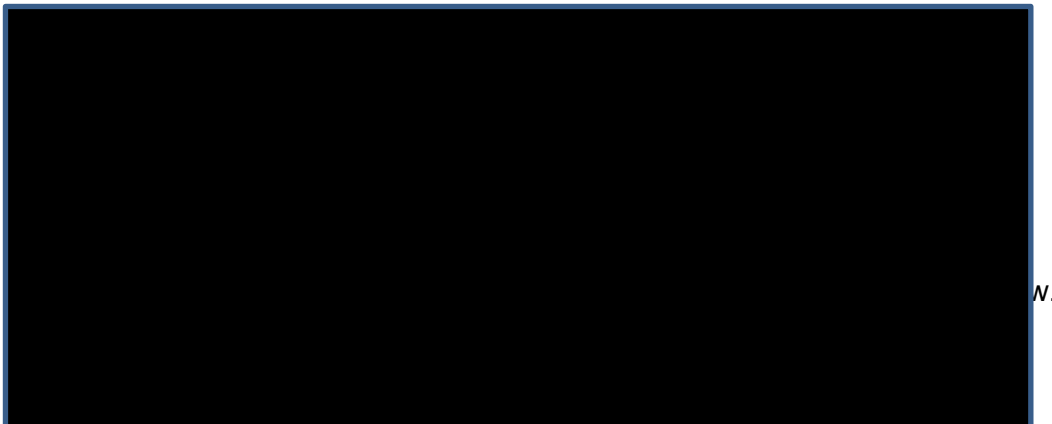
NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Figure 2: Kaplan-Meier plot for time from first dose to allogeneic stem cell transplantation in cohort 1 of KEYNOTE-087

Figure 3: Kaplan-Meier plot for time from first dose to allogeneic stem cell transplantation in cohort 2 of KEYNOTE-087



NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Treatment effectiveness

B5. **Priority question:** There are inconsistencies in the distribution of patients across the different response states in the company submission.

a. Please provide a corrected version of Table 62 from the company submission.

An updated version of table 62 is provided below.

Response	N	N
Cohort 1		
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]
Cohort 2		
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]

b. The estimation of the distribution of patients in the different response states appears incorrect. In particular, the proportions in the stable disease state in both cohorts in the model are significantly larger than those observed in the KEYNOTE-087 study, as reported in Table 19 of the company submission. This results in significant over-estimation of the numbers of people having allogeneic stem cell transplant in both cohorts. Please submit a version of the model in which response rates are corrected and explain the correct proportion of responders, and how they have been calculated, referring to the clinical part of the company submission.

The model assumes that any patient that fails to achieve complete or partial response, and does not progress or die by week 12 of the time horizon would occupy the stable disease state. By definition, this includes any patient with a non-evaluable response status. This assumption applies to both arms of the model.

Table 19 of the company submission outlines the proportion of patients in each response category based on the raw trial data. This includes the category of non-evaluable response. In the model, this group are included as part of the stable disease state. For example, the base case model prediction that 37% of patients treated with pembrolizumab have stable disease at week 12 is consistent with the sum of stable disease ([redacted] %) and non-evaluable response ([redacted] %) at week 12, from Table 19.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

As can be seen in Table 19, for cohort 1, the values applied in the model are predicted well. For cohort 2, the model under predicts occupancy in PD state. This is not an error but a consequence of the survival models used in the base case and the unusual shape of PFS for this population with a lot of events at or around week 12. We believe the discrepancy to be explained as there is a window for which events at week 12 (day 84) can be collected (day 74 to 126).

A comparison of Table 19 and model predictions is provided below:

Status	Table 19 of submission (cohort 1) N (%)	Model predictions (cohort 1) %	Table 19 of submission (cohort 2)	Model predictions (cohort 2)
Complete response	[redacted]	15.94%	[redacted]	8.6%
Partial response	[redacted]	42.0%	[redacted]	43.2%
Stable disease	[redacted]	36.9% *(~27.5%+8.7%)*	[redacted]	38.9% (~18.5% + 8.6%)
Non-evaluable	[redacted]	Not reported (combined in stable disease)	[redacted]	Not reported (combined in stable disease)
Progressed disease	[redacted]	4.10%	[redacted]	7.9%
Death	Not reported	1.04%	Not reported	1.22%

B6. **Priority question** *company submission page 162*: The company submission states that for progression free survival, parametric models were fitted to the observed data in KEYNOTE-087 from week 12. However, in the model file it appears that parametric models were fitted to all the data, beginning from week 0.

a. Please clarify what was done.

The parametric models used to predict PFS in weeks 0-12 of the time horizon were modelled on the observed data set beginning in week 0 given the small number of events occurred in the first 12 weeks as described on page 153 of the manufacturer submission. Whilst the statistical analysis was performed on the entire observed data, the resulting parametric models are used to predict PFS during the first 12 weeks only. The parametric model was chosen based on its statistical fit and ability to predict the most comparable rate of patient's progression free as the observed KEYNOTE-087 data at week 12.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The parametric models used to predict PFS from week 12 to lifetime were modelled on the observed data from weeks 12 to end of study follow-up as described on page 162 of the manufacturer submission. Any patients that experienced an event prior to week 12 were excluded from the analysis. Only PFS from KEYNOTE-087 was modelled beyond week 12 on the basis that OS was too immature to provide robust and plausible life-time extrapolations. External data from Cheah et al was used to model post-progression survival.

b. Please provide progression free survival and overall survival curves fitted to the entire study data for both cohorts and include them in the model.

The parametric models used to predict PFS and OS during the first 12 weeks of the evaluation are based on the parametric survival analysis of the entire study data. These data are available in the submitted model, and are summarised in pages 153-159 of the dossier.

c. Please estimate comparative effectiveness using the entire study data (from week 0 to end of follow-up).

The results of the comparative effectiveness analysis using the entire study follow-up are available in section 4.10 of the dossier (and repeated below).

Table 7: Summary of comparisons of progression-free survival for pembrolizumab versus SoC for the entire study scenario[61]

Cohort	Comparison	Sample size/effective sample size, n	Pembrolizumab		Hazard ratio (95% CI)
			Events, n	Censored, n	
1	Naïve	69	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
2	Naïve	81	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
1 and 2	Naïve	150	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; MAIC, matching-adjusted indirect comparison.

B7. *Company submission page 161:* In the clinician survey on the probabilities of allogeneic stem cell transplant conditional on response status, some clinicians had indicated that even with progressed disease patients could be eligible for allogeneic stem cell transplant.

a. Please explain why this was not incorporated in the model.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The MSD clinician survey returned responses that suggested alloSCT may occur in persons with PD. The clinical plausibility of alloSCT use in PD was further discussed with UK clinicians who advised that this would not be considered standard UK clinical practice. This is supported by the available UK data in KEYNOTE-087. Of the [redacted] UK patients in KEYNOTE-087 that underwent alloSCT (cohorts 1 and 2), [redacted] had complete response, [redacted] had partial response, [redacted] had stable disease and [redacted] was non-evaluable at the time of their last response measure prior to transplant. None of the UK patients had PD prior to alloSCT. On this basis, alloSCT in PD was not incorporated in the model.

b. Please provide a scenario analysis in which patients in the progressed disease health state could also receive allogeneic stem cell transplant.

As described within the main submission document (page 160-161), and when taking into consideration the feedback of clinical experts and the limited data available within the MSD clinician survey; the generalisability/ relevance of this scenario analysis to the UK population is limited. However, please find results for the scenario (Table 8) as per the new base case model including the transition of progressed disease patients to alloSCT as per the result generated in the MSD clinician survey.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 8 Scenario analysis using base case alloSCT rates and progressed disease alloSCT rate from MSD clinician survey

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
AlloSCT: PD 12.14%	Cohort 1	£107,388	6.182	4.712	£47,422	4.576	3.466	£59,966	1.246	£48,133
	Cohort 2	£93,034	5.650	4.305	£46,632	4.533	3.435	£46,402	0.870	£53,332

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

c. Please provide the data used from the previously conducted clinician survey by BMS (for nivolumab) and explain the methods of how the two were combined – and how it was ensured that the same experts were not reflected twice.

We are unable to provide any further detail or comment on the methodology of the BMS clinician survey as this was marked academic in confidence by the manufacturer. MSD are aware of these data having attended the public committee meeting.

As detailed on page 161 of the MSD manufacturer submission, an average of the two means generated from both the MSD clinician survey and that of the BMS clinician survey was used in the base case economic model.

MSD accepts, given the specialist nature of RRcHL, that it is possible for both surveys to have included the same clinical experts. Given the ACIC nature of the BMS survey it is not possible to comment in relating to the MSD survey. The MSD survey was conducted by a third party and responses were anonymous.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

d. Please perform a scenario analysis in which only the data from the new survey is used.

Please find the results as per the new base case model (Table 9) of the scenario analysis requested below using AlloSCT rates from the MSD clinician survey only.

Table 9 Scenario analysis using alloSCT rates from the MSD clinician survey only

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
AlloSCT: CR 56.79% PR 43.93 SD 18.36%	Cohort 1	£104,697	5.754	4.391	£39,883	4.114	3.116	£64,814	1.276	£50,806
	Cohort 2	£88,480	5.193	3.960	£39,099	4.073	3.085	£49,382	0.875	£56,419

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B8. **Priority question** *company submission page 141*: The hazard ratio for progression free survival post-12 weeks is the single most important parameter as per the company's deterministic sensitivity analyses. However, it is based on a strong assumption: it was assumed to be the same as in the pre-12 weeks period because *"a PFS HR from week 12 to end of follow-up could not be estimated given the low number of events post week 12 observed in Cheah"* (page 141). This suggests that the HR in the pre-12 week period is potentially an over-estimate of comparative effectiveness of pembrolizumab versus standard of care.

a. Please comment on the potential bias this introduces, given that most events in Cheah et al. (2016)[1] had already happened in the first 12 weeks and that therefore the HR after 12 weeks would be expected to be much less in favour of pembrolizumab – and also in light of choosing different distributions for modelling pre- and post-12 week PFS.

The hazard ratio of PFS for pembrolizumab versus standard of care was derived using the entire follow-up period. This assumes a constant treatment effect across the follow-up period, e.g. that the proportional hazard assumption holds. This is consistent with the assumptions of the Cox regression analysis. MSD agree that this could introduce bias. However due to limited data available of only one publication (Cheah et al) in which the majority of events occurred pre 12 weeks it was not possible to determine a HR post 12 weeks. There is significant uncertainty surrounding the PFS of patients in Cheah et al beyond week 12 given the small numbers of patients at risk. This limits our ability to draw meaningful conclusions on the comparative effect of drug during this period.

MSD agrees that using the HR of 0-end is more favourable to pembrolizumab than that of 0-12 weeks. MSD could have utilized the HR for 0-12 weeks in both time periods to be further conservative however we feel this would not be a true reflection and utilisation of the available KEYNOTE-087 data post 12 weeks. It also would have introduced double counting of events to use the HR for 0-12 weeks followed by the HR for 0-end post 12 weeks. In addition, the fact that the 0-end HR is more favourable implies that the averaged treatment effect improves with longer follow-up. If the effect of drug was to wane over time, then we would expect to see the averaged effect reduce with longer follow-up and whilst the point estimates of the hazard ratios are different, there is clear overlap in the confidence intervals between the two analyses, and hence it could be argued that "Statistically" there is no significant difference in effect across the two periods.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

MSD did apply different distributions in the pre- and post- 12 week time period as described in B6a also. MSD believe this is acceptable due to different decision criteria being followed during model selection, given that the pre-12 weeks data has been observed and the post-12 week data is reliant on extrapolation beyond the observed data.

- Pre-12 week period: parametric survival models were assessed on visual fit and the proportion of patients progression-free/alive at the 12 weeks, as there was no extrapolation being applied within this section of the model.
- Post-12 week period: parametric survival models were assessed following the DSU guidance by considering the statistical and visual fit to the observed data and the clinical plausibility of the extrapolated period.

b. Please state whether the use of any alternatives to using the same HR as in the pre-12 week period were considered. If any alternative were considered, please provide details and an explanation of why they were not used.

Alternatives for calculating a HR for the post 12-week period were explored. However, substantial differences between the treatments in the pre-12 week period meant that to do so would have required either 1) the Cheah PFS to be reset to 1 at 12-weeks, which is unreasonable, or 2) the Cheah PFS was not reset to 1 at 12-weeks, in which case the proportional hazards assumption would have been violated.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

c. Please provide one scenario analysis (both cohorts) using HR = 1 for the post-12 week period and another using the estimated HR from the indirect treatment comparison (ITC), reflecting the uncertainty associated with it.

This does not conform to any known data associated with pembrolizumab to date. However, as requested please find the results as per the new base case model of the scenario analysis requested below using a HR of 1 for the post 12 week period (Table 10).

Table 10 Scenario analysis with HR=1 for the post-12 week period

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
HR=1 post-12 weeks	Cohort 1	£106,908	6.153	4.690	£46,298	4.971	3.777	£60,609	0.913	£66,378
	Cohort 2	£92,100	5.594	4.262	£44,405	4.655	3.535	£47,695	0.727	£65,594

As discussed in question B8b, a HR for post 12 weeks could not be calculated hence this scenario analysis has not been conducted.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B9. *Company submission figures 17 and 21:* In cohort 2, patients progress significantly faster than in cohort 1. Please explain what causes patients in cohort 2 to start progressing at approximately [redacted] weeks (company submission figure 17 page 156) and for all of them to have progressed at approximately [redacted] months (company submission figure 21 page 167), as opposed to much longer estimates for time to progression in cohort 1?

After seeking input from a clinical expert, it is thought that the difference observed between cohorts 1 and 2 with regard to "earlier disease progression" is due to inherent multifactorial difference between the two populations.

Patients within Cohort 1 and 2 will have underlying differences in their RRcHL disease state. Cohort 2 is made up of an older patient population, with 18.5% of the population described as ≥65 years of age versus 0% in Cohort 1. Patients in Cohort 2 by definition were unable to undergo an autoSCT, meaning they are likely to have been refractory to chemotherapy. Clinical opinion suggests that Cohort 2 represents a higher risk group compared with Cohort 1 and thus are likely to progress more quickly compared with Cohort 1. Cohort 2 for this reason is thought to represent the main source of unmet clinical need in patients with RRcHL.

B10. *Company submission page 169:* The calculation of post progression survival was based on data from Cheah et al. (instead of KEYNOTE-087). Please provide a scenario analysis using data from the KEYNOTE-087 study to calculate post progression survival.

Kaplan-Meier plots for time from progression to death in KEYNOTE-087 for cohorts 1 and 2 are presented in

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Figure 4 and

Figure 5, respectively. As of March 2017, there were two post-progression death events in cohort 1, and one post-progression death event in cohort 2.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Figure 4: Kaplan-Meier plot of time from progression to death in KEYNOTE-087 (cohort 1)

Figure 5: Kaplan-Meier plot of time from progression to death in KEYNOTE-087 (cohort 2)

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Exponential survival models were fitted to post-progression survival in cohorts 1 and 2 of KEYNOTE-087 (March 2017 data cut) to estimate a constant transition probability for progression to death after pembrolizumab treatment.

A summary of the estimated constant transition probabilities are provided in Table 11.

Table 11: Estimated constant transition probabilities from post-progression survival in KEYNOTE-087

Cohort	Fitted constant hazard rate (weekly)	Transition probability based on exponential model fitted to PPS in KENOTE-087	Probability of survival in progressed disease state (one minus probability of death/value used in the model)	Predicted Median PPS in years based on exponential (ln(2)/rate)	Predicted Mean PPS in years based on exponential (1/rate)
1	0.002235	0.223%	99.78%	5.96	8.60
2	0.000722	0.072%	99.93%	18.45	26.62
Cheah et al base case	-	-	99.37%	-	-

Based on the exponential model, the predicted mean PPS is 8.60 years for cohort 1 and 26.62 years for cohort 2.

Please find the results as per the new base case model (Table 12) of the scenario analysis requested below using the PPS estimated from KEYNOTE-087 for cohorts 1 and 2 separately and then for cohorts 1 and 2 combined applied to pembrolizumab only.

Table 12: Scenario analysis using PPS calculated from KEYNOTE-087

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
PPS estimated from KN-087 cohort 1: 99.78% & cohort 2: 99.93%;	Cohort 1	£114,627	8.312	6.198	£44,278	4.385	3.322	£70,349	2.876	£24,457
	Cohort 2	£114,387	11.974	8.611	£43,275	4.330	3.280	£71,112	5.331	£13,340

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

pembrolizumab only										
PPS estimated from KN-087 cohort 1&2: 99.87%; pembrolizumab only	Cohort 1	£120,214	9.909	7.275	£44,278	4.385	3.322	£75,937	3.954	£19,206
	Cohort 2	£106,913	9.772	7.150	£43,275	4.330	3.280	£63,638	3.869	£16,447

B11. **Priority question:** Lafferty et al. 2017[2] is an abstract reporting on a single centre experience of 13 patients with Hodgkin Lymphoma who had allogeneic stem cell transplant. According to page 29 of the company submission, there were 23 patients in the KEYNOTE-087 study who had allogeneic stem cell transplant.

a. Please explain why Lafferty et al. (2017) was preferred over the KEYNOTE-087 study to estimate post-allogeneic stem cell transplant overall survival (OS), proportions for complete response (CR)/partial response(PR)/stable disease (SD)/progressive disease (PD) as well as for acute graft vs host disease after allogeneic stem cell transplant.

KEYNOTE-087 was a global study and not set up to look at outcomes associated with alloSCT. The clinical study report (June 2016) for KEYNOTE-087 provides full details relating to study design, study objectives, and exploratory endpoints; these did not include the subsequent investigation of patients treated with pembrolizumab who were then treated with a stem cell transplant. Please refer to question A7 for further information on this.

As explained on page 144 of the manufacturer submission, OS data for the entire study population of KEYNOTE-087 was deemed to be too immature to provide robust extrapolations of survival and hence this is the case also for the patients who had an alloSCT.

In addition, there was also precedent set for the use of Lafferty et al (2017) in this way as it was accepted in the preferred analysis by the committee in TA462 which is in the same advanced cHL patient population as described in this submission.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. Please estimate post-allogeneic stem cell transplant OS, proportions for CR/PR/SD/PD as well as for acute graft vs host disease after allogeneic stem cell transplant based on the KEYNOTE-087 study and use this in a scenario analysis.

As explained in the answer to B11a, the OS data is not available for patients who underwent an alloSCT and so MSD has not been able to do this analysis.

MSD was able to produce a scenario analysis based on the new based case model using the KEYNOTE 087+013 rate of GVHD which can is based on [redacted] alloSCT patients as of September 2016 data cut (Table 13).

Table 13 Scenario analysis using KEYNOTE 087+013 GVHD rate

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
GVHD using KEYNOTE-087 rate	Cohort 1	£106,907	6.153	4.695	£44,277	4.585	3.324	£62,130	1.371	£45,690
	Cohort 2	£92,100	5.594	4.266	£43,275	4.330	3.285	£48,824	0.983	£49,651

c. Please comment on whether expert opinion was consulted to validate the OS estimates resulting from this study.

As detailed in section 5.3.4 on page 181 of the manufacturer submission, we can confirm that expert clinical opinion was consulted on the estimates of OS following alloSCT derived from Lafferty et al 2017. The OS estimates from Lafferty were applied equally across both pembrolizumab and SoC which can be considered conservative compared to the outcomes which might be expected on treatment with pembrolizumab prior to alloSCT.

OS data for KEYNOTE 087 was deemed too immature to provide robust extrapolations of survival and so parametric survival models were fitted to PFS post week 12 in the no alloSCT part of the model. These estimates were also validated with a clinical expert.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B12. In TA 462 investigational agents were excluded from the Cheah data analysis. Referring to Question A.13, please provide a scenario analysis in which patients receiving pembrolizumab, or if this is not possible investigational agents, are excluded.

Please see the response to question A13 a and b.

B13. **Priority question:** Please provide a scenario analysis in which, in line with the marketing authorisation for pembrolizumab, patients can continue treatment after 24 months.

below shows the results as per the new base case model of the scenario analysis in which

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
PEM beyond 24 months	Cohort 1	£120,979	6.153	4.690	£44,278	4.385	3.322	£76,702	1.368	£56,055
	Cohort 2	£95,093	5.594	4.262	£43,275	4.330	3.280	£51,818	0.981	£52,798

patients can continue treatment after 24 months in line with the marketing authorisation for pembrolizumab.

Table 14 Scenario analysis of continued treatment with pembrolizumab after 24 months

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
PEM beyond 24 months	Cohort 1	£120,979	6.153	4.690	£44,278	4.385	3.322	£76,702	1.368	£56,055
	Cohort 2	£95,093	5.594	4.262	£43,275	4.330	3.280	£51,818	0.981	£52,798

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

www.nice.org.uk

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Adverse events

B14. *Company submission page 182 onwards*: Adverse event incidence rates for standard of care were derived from different sources reporting adverse events associated with chemotherapy regimens and bendamustine. Please provide more information on how the weighted average for standard of care was derived.

The weighted average was calculated from the composition of SoC in Cheah (2016) (chemotherapy 38.5%; bendamustine 18.5%; investigational agents 43.1% [Table 88 company submission]) and the following assumptions:

- The adverse event rate for chemotherapy was calculated as the sum of all events divided by the sum of sample sizes of each chemotherapy study listed in Table 72 of the company submission for TA462
- Bendamustine rates were derived directly from a phase II study of bendamustine in relapsed and refractory Hodgkin lymphoma (Moskowitz AJ, Hamlin Jr PA, Perales M-A, Gerecitano J, Horwitz SM, Matasar MJ, et al. Phase II study of bendamustine in relapsed and refractory Hodgkin lymphoma. Journal of Clinical Oncology. 2012;31(4):456-60).
- Investigational agents were assumed to have no adverse events due to a lack of data

Health-related quality of life

B15. **Priority question** *company submission page 185 onwards*: Long term utility values have been estimated from KEYNOTE-087 using observations from week 12 only. These utility values might not reflect long-term utility values.

a. Please use a mixed model based on utility data from KEYNOTE-087 to estimate utility values separately for complete response, partial response, stable disease and progressive disease, using all available observations from the KEYNOTE study for participants before they have allogeneic stem cell transplant.

The results of the mixed model analysis of utility data from KEYNOTE-087 are provided in Table 15. These data have been updated in the new base case model attached and have been implemented in all further scenario analysis requested in the clarification questions. MSD believe that since the measure of utility during KEYNOTE-087 was taken shortly after a

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

progression event and so potentially only reflects the immediate consequences of progression whereas the original utility values utilised in the manufacturer submission from Swinburn et al would be more likely to take the longer term utility decrement due to progression into account hence the results from KENOTE-087 for the progressed disease health state are likely an overestimate of utility.

Table 15: Mixed effects model of utilities by response state using all observed EQ-5D values in KEYNOTE-087

Covariates	Estimated effect	Standard error
Intercept	**********	**********
PD versus CR	**********	**********
PR versus CR	**********	**********
SD versus CR	**********	**********

The predicted absolute utility values are:

- CR: **********
- PR: **********
- SD: **********
- PD: **********

b. Please use a mixed model based on utility data from KEYNOTE-087 to estimate utility values for post-allogeneic stem cell transplant, using all available observations from the KEYNOTE study for participants after they have allogeneic stem cell transplant. If possible also provide an additional analysis distinguishing between <14 weeks and >14 weeks post-allogeneic stem cell transplant, or alternatively distinguishing between patients with and without acute graft vs host disease.

As of March 2017, post-stem cell transplantation utility data were available for 1 patient that had received an alloSCT transplant in KEYNOTE-087 (cohorts 1 and 2 combined). As such, there are insufficient data to perform a mixed effects regression analysis of utilities for post-allogeneic stem cell transplantation.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B16. Please explain why the proportions of responders (complete response and partial response) used to calculate utilities for the progression free health state (that is, used in calculations in cells G18 and H18 of the model; "NonClinicalData" worksheet) are different from the proportion of responders used to calculate the proportions of patients receiving allogeneic stem cell transplant at week 12 (that is, proportions in rows 38-41; "ClinicalData" worksheet). Please provide a corrected model, if necessary.

The model is correct. The response rates used to calculate the utilities for the progression-free health state were adjusted as outlined in the company submission; Table 80 footnote 'patients that had progressed at week 12 were excluded; those who were not assessed were assumed to have stable disease' (i.e. the proportion of patients with CR/PR/SD of those who were progression-free at week 12).

B17. The lack of a post-progression health state in the allogeneic stem cell transplant group potentially causes a bias in the estimation of quality of life.

a. The one-year progression free survival after allogeneic stem cell transplant reported by Lafferty et al. (2017)[2] was 54%. Please indicate how the occurrence of progression is reflected in quality of life in the post allogeneic stem cell transplant health state (from which patients cannot transit to a progressive disease health state). Although, in the model, patients have a disutility during the first 14 weeks post-allogeneic stem cell transplant, after this period patients are assumed to have a utility of 0.865 while a utility of [redacted] is assumed for the progressive disease health state.

A review by Pidala et al (Pidala J, Anasetti C, Jim H, Quality of life after allogeneic hematopoietic cell transplantation. Blood. 2009 Jul 2;114(1):7-19) which aimed to critically evaluate the literature on QoL following an alloSCT states that patients experience a decline in QoL immediately following alloSCT but this largely returns to baseline values or improves after 100 days and remains at this level with the majority of studies suggesting over 60% of patients reporting good to excellent QoL in years 1-4 following an alloSCT. This suggests that progression post-alloSCT is not a largely determinant of chronic quality of life.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. Please explain why the allogeneic stem cell transplant utilities were assumed to be driven by complete response/partial response/stable disease (given that this utility is calculated based on the complete response/partial response/stable disease proportions and the 12-week utility values).

All health state utilities in the model were based on utilities from KEYNOTE-087, to ensure consistency in values applied to patients who entered the model (e.g. enrolled to KEYNOTE-087) and those that went on to receive alloSCT.

With the exception of the first 100 days post-alloSCT where utilities were adjusted for graft versus host disease (GVHD), The alloSCT health state utilities in the model were predicted by response and progression status available at week 12. The response status was determined by treatment with rates varying across patients treated with pembrolizumab and standard of care. This ensures that any differences in utility scores across health states are determined by the effectiveness and safety of treatment, conditional on the target population of persons with RRcHL. As mentioned in B17a, patients who have undergone alloSCT are expected to receive an initial utility decrement followed by a return to their initial utility or greater following 100 days which would be the response related utility seen at 12 weeks in the model.

MSD accepts that using utility from week 12 only may introduce bias however there is limited data available and utilisation of this is in keeping with the model structure chosen (further rationale for which presented in B4a). Additionally this removes the need to obtain utilities from external data sources, which in the case of RRcHL with its limited evidence base would necessitate the use of data from other cancers. The mixing of utilities from different data sources that may be studied on different populations, including those with different cancers, introduces a risk bias, as difference in utilities may be due to differences in population characteristics and not the intervention or health status of the population.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B18. Please explain why only the disutility for acute graft vs host disease was used from Kurosawa et al. (2015)[3] instead of using the utilities from this study for the post allogeneic stem cell transplant health state (that is, utility values of 0.65 and 0.80 for <14 weeks and >14 weeks post allogeneic stem cell transplant respectively).

Kurosawa et al. (2015) was a cross-sectional study of quality of life in patients with acute leukaemia (68% AML and 30% ALL). The majority of patients had a disease status of first complete remission at the time of transplantation (66%), with a median time from treatment to survey of 5-years. None of the patients enrolled in this study had RRcHL as per the enrolment of KEYNOTE-087. Due to the differences between populations in Kurosawa et al. and KEYNOTE-087, the absolute utility scores of 0.80 (alloSCT without graft vs host disease (GVHD)) and 0.65 (alloSCT with GVHD) were not considered suitable for modelling post-alloSCT in the economic analysis.

Further, the utility values from Kurosawa et al. relate to the preferences of patients that had alloSCT with versus without GVHD. These utility values do not correspond to the utilities of the <14 week and > 14 week post alloSCT, and hence were not applied in this manner in the submission. The utilities reported in this study provide the best available estimate of the utility loss from GVHD post-alloSCT. Graft versus host disease was shown to be a strong determinant of health related quality of life post-alloSCT in a systematic literature review by Pidala et al. Rates of Graft versus host disease, and its associated impact on quality of life, was used to model the overall detriment in quality of life expected after alloSCT. No other preference based (e.g. EQ-5D) utility measures for post-alloSCT were identified in the utility review or via other ad hoc searches of the published literature.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Resource use and costs

B19. For the calculation of standard of care costs, it is assumed that all chemotherapy agents contribute an equal proportion of treatment to standard of care. Please justify this assumption.

There is a paucity of evidence on the preferred or standard mix of chemotherapy regimens given to patients in UK clinical practice. In the absence of this data, the SoC composition was modelled using an equal proportion across regimens, following the approach accepted by the committee in the preferred analysis in TA462.

B20. According to the company submission and the model file, no doses of pembrolizumab or standard or care were missed.

a. Please provide a justification for this assumption and explain whether alternative assumptions were considered.

The manufacturer can confirm that missing doses of pembrolizumab or standard of care were not considered in base case economic analyses. This is justified on the basis that there is insufficient information on adherence to standard chemotherapy in RRcHL to support adjustment of costs for missing doses. Applying adjustment to only one arm of the analysis (e.g. pembrolizumab) would have reduced the costs of the main intervention whilst keeping standard of care costs at full cost and hence bias the results of the analysis in favour of the intervention. High rates of adherence to pembrolizumab were reported of over [redacted] in the KEYNOTE-087 September 2016 data cut.

b. Please provide a scenario analysis considering missed doses.

Table 16 shows the results of a scenario analysis as per the new base case model incorporating data for missed doses of pembrolizumab [redacted] adherence rate for pembrolizumab and an assumed 100% for SoC) from KEYNOTE-087 based on the September 2016 data cut.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 16 Scenario analysis incorporating pembrolizumab missed doses

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Base case (company submission)	Cohort 1	£106,907	6.153	4.397	£44,277	4.385	2.757	£62,630	1.639	£38,201
	Cohort 2	£92,100	5.594	3.892	£43,275	4.330	2.711	£48,824	1.181	£41,341
Base case (mixed model utilities)	Cohort 1	£106,908	6.153	4.690	£44,278	4.385	3.322	£62,630	1.368	£45,772
	Cohort 2	£92,100	5.594	4.262	£43,275	4.330	3.280	£48,825	0.981	£49,748
% doses of PEM: [redacted]	Cohort 1	£106,202	6.153	4.690	£44,278	4.385	3.322	£61,924	1.368	£45,255
	Cohort 2	£91,531	5.594	4.262	£43,275	4.330	3.280	£48,256	0.981	£49,169

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B21. In the company submission it is stated that baseline cohort characteristics of the KEYNOTE-087 cohort are used in the model (in table 100). The model however shows a body surface area (BSA) of 1.85 was used which does not match with the mean BSA of the KEYNOTE-087 cohort (as reported in the company submission table 56 page 150). Please explain this discrepancy, and correct the model if necessary.

Apologies, the BSA reported in Table 56 is an error. The value of 1.85 used in the model is the correct value from KEYNOTE-087 cohort 1 and 2 combined.

B22. *Company submission page 214*: In the model, one-off costs are applied upon treatment with allogeneic stem cell transplant. These costs are taken from Radford et al. (2017)[4], a study reporting on costs in 14 relapsed or refractory classical Hodgkin Lymphoma patients treated with allogeneic stem cell transplant, where mean follow-up was 3.44 years (the proportion of patients that died was not stated). Given the uncertainty on what proportion of patients died in Radford et al. (2017), it seems impossible to assess the assumption that terminal care costs are already included in the one-off cost. Furthermore, given the follow-up period of maximum 5 years and the fact that in the model 40% of patients treated with allogeneic stem cell transplant are alive at 10 years, it is questionable if the one-off cost accurately captures monitoring and subsequent treatment costs allogeneic stem cell transplant.

a. Please explain why a one-off cost was deemed appropriate over the use of individual costs for allogeneic stem cell transplant intervention and subsequent treatment, monitoring, adverse events and terminal care costs. Please also explain why these costs are assumed to be reflected by the one-off cost.

The cost of alloSCT has been taken from Radford as this was the preferred source by the committee in TA462. Radford was a retrospective analysis that studied the cost and resource use in 40 cHL patients who had failed after autoSCT. A total of 15 patients subsequently received alloSCT and were followed up to date of death or to most recent follow-up (mean 3.44 years), with a mean total cost of £110,374. Of this total cost 31.5%–39.9% were due to the alloSCT procedure itself ranging between £34,783 and £44,059 per patient. Therefore, a substantial proportion of the cost was associated with additional follow-up cost so no additional disease management costs or AE costs were included in the model for post alloSCT patients.

b. Please provide a scenario where the one-off cost for allogeneic stem cell transplant is replaced by costs accounting for the allogeneic stem cell

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

transplant intervention and post- allogeneic stem cell transplant period separately.

A scenario analysis incorporating this was not feasible within the timeframe for response to clarification questions. However, as detailed in question B22a, the cost of alloSCT used in the manufacturer submission was the preferred source by the committee in TA462 which is in the same advanced cHL patient population. Since the majority of the costs (60.1-68.5%) accounted for within this cost was associated with additional follow-up cost we believe that the Radford cost utilised in the manufacturer submission adequately reflects the post alloSCT period.

c. It is stated in the company submission on page 209 that palliative care costs within best supportive care are set to £0 in line with assumptions made in TA462,[5] please provide an explanation why this assumption is appropriate for this appraisal.

This assumption was made on the basis that there is a paucity of available evidence in this particular patient population. Furthermore, it is assumed based on discussions with UK clinical experts that BSC encompasses palliative care costs. There was also precedent set for this assumption as it was accepted in the preferred analysis by the committee in TA462 which is in the same advanced RRcHL patient population.

d. According to the company submission on page 212, costs of terminal care come from a previous HTA assessment on non-small cell lung cancer that have been updated. Please provide a justification for why data from a different patient population was used and explain how the costs were updated. Furthermore, please clarify whether the inclusion of terminal care costs results in double-counting of costs, given that patients receive treatment until they die.

This cost was used on the basis that there is a paucity of available evidence in this particular patient population. The terminal care cost in non-small cell lung cancer was the most similar patient population cost which could be found and there is also a precedent for its use as the committee preferred terminal care costing in TA462 which is in the same advanced RRcHL patient population. MSD agrees with your point that there may be some double counting within the terminal care cost. However since this cost is applied across both the intervention and SoC we do not believe this to have a large impact on the results of the analysis.

Cost effectiveness results

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B23. *Company submission page 220:* Please provide the UK standard of care clinical outcomes which appear to be missing from Table 102.

Please see updated table below.

		Pembrolizumab		UK SoC	
Outcome		**Base case**	**KEYNOTE-087**	**Base case**	**Cheah et al**
% PFS at 1 Year *	Cohort 1	59.7%	[redacted]	4.1%	~7.5%
	Cohort 2	44.1%	[redacted]	4.9%	
OS at week 12	Cohort 1	98.96%	[redacted]	98.96%	~100%
	Cohort 2	98.76%	[redacted]	98.76%	
OS at 72 Months**	Cohort 1	28%	-	15%	15%
	Cohort 2	22%		15%	
OS after alloSCT 5 years		Base case	KEYNOTE-087	Base case	Lafferty et al
	Cohort 1	54.50	-	54.50	53.47
	Cohort 2				
*using data post week 12 assuming no alloSCT as per KEYNOTE-087 design ** when no alloSCT is assumed as per assumption made about Cheah SoC arm					

B24. Please provide the total LYs for each comparator obtained from the probabilistic sensitivity analysis.

Please see updated table below

Technologies	**Cohort**	**Total LYs**
UK SoC	Cohort 1	4.710
	Cohort 2	4.648
Pembrolizumab	Cohort 1	6.255
	Cohort 2	5.594

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B25. Please explain how the cost and effects of best supportive care were estimated in scenario analyses 1. Please provide a model file in which best supportive care can be selected as a comparator, and both best supportive care and standard of care can be included simultaneously in the probabilistic sensitivity analysis.

The composition, dosages, acquisition and administration costs of best supportive care (BSC) as a primary comparator in scenario 1 are identical to the cost of BSC as a subsequent treatment (outlined in section 5.5.6 of the submission), with the exception of the durations which were informed by the PFS from Cheah (2016) with the maximum number of cycle applied where appropriate.

The efficacy of BSC was assumed equal to SoC (Cheah 2016) in line with TA462.
The model only allows one comparator to be included, thus a significant overhaul of the model would be required to run both SoC and BSC simultaneously in the PSA which is not plausible in the timeframe available. It would also be uninformative given that the comparators use the same efficacy data. Therefore, please find attached the model with BSC as a comparator.

Model validation

B26. Please provide a detailed cross validation with TA462,[5] commenting on all differences in modelling assumptions, data sources used and model predictions, at least for standard of care, using publicly available data on the inputs and results from TA462.

A comparison of many of the model assumptions vs TA462 and the de novo analysis for this appraisal is provided in the main submission document (Table 100).

A detailed comparison with TA462 is complicated by the amount of information redacted in the publically available documents, and the complexity of the modelling and partial descriptions provided in the company submission.

B27. Please provide model files without hidden worksheets, rows and columns.

The attached new base case model has the hidden worksheets, rows and columns unhidden.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B28. Please confirm whether the model files for cohort 1 and cohort 2 are exactly the same.

The model files for cohort 1 and cohort 2 contain identical data. Two models were only submitted for the presentation of the PSA; both cohorts can be run in the same model deterministically.

References

[1] Cheah CY, Chihara D, Horowitz S, Sevin A, Oki Y, Zhou S, et al. Patients with classical Hodgkin lymphoma experiencing disease progression after treatment with brentuximab vedotin have poor outcomes. *Ann Oncol* 2016;27(7):1317-23.

[2] Lafferty N, Anandram S, Lawes N, Usman M, Liebersbach S, Benn K, et al. Allogeneic stem cell transplantaion in patients with Hodgkin Lymphoma: a retrospective single centre case series. *Br J Haematol.* 2017;176(Suppl 1):5-145.

[3] Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplant* 2015;50(9):1241-9.

[4] 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 Transplant* 2017;52(3):452-4.

[5] National Institute for Health and Care Excellence. *Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma: NICE technology appraisal guidance TA462 [Internet]*. London: NICE, 2017 [accessed 18.9.17] Available from: https://www.nice.org.uk/guidance/ta462/resources/nivolumab-for-treating-relapsed-or-refractory-classical-hodgkin-lymphoma-pdf-82604902197445

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)
Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

Thank you for agreeing to make a submission on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your submission, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation: Royal College of Pathologists
Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? - yes
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? - yes
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc)? – yes: chair of the NCRI Hodgkin lymphoma study group and chair of the British Society of Haematology Special Interest Group for Lymphoma
- other? (please specify)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)
Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

80% of classical Hodgkin lymphoma patients are cured with first line treatment. For those that relapse, standard treatment is combination chemotherapy followed by an autologous stem cell transplant (ASCT). This cures approx. 50% of relapsed patients. For those who relapse after ASCT, standard treatment is brentuximab vedotin followed by allogeneic stem cell transplant (alloSCT). Current areas of unmet need are:
- Patients who fail to get in a good enough remission to proceed to ASCT (only patients in a complete remission or near complete remission will be offered ASCT)
- Patients who fail to respond to BV after relapsing post ASCT (70% patients respond to BV but in 40% of these case the remission are not very durable and patients frequently relapse before the alloSCT can be performed
- older patients who are not fit for ASCT or alloSCT represent a major area on unmet need in Hodgkin lymphoma

Nivolumab is licensed for use (and received a positive FAD) for treatment of relapse patients following ASCT and BV. Pembrolizumab has a wider EMEA license, allowing use in patients who have had BV but are not suitable for ASCT. This will include older, frailer patients who will never be fit for transplant but it will also include those who are not in a good enough remission to proceed to transplant (although they may become suitable candidates if a stable remission is obtained).

Pembrolizumab is currently not available for use in the NHS for Hodgkin lymphoma due to absence of funding. If reimbursed, it would be used in a secondary and tertiary care setting. PD1 inhibitors are being used widely for other oncological indications so many centres have quite wide experience of their use now, but little in Hodgkin.

The current ESMO guidelines for Hodgkin lymphoma do not include the use of PD1 inhibitors; the BCSH guidelines also make no mention of them. This is because they were both published over 3 years ago, before any of the trial results were available.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)
Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

1. Patients who are not suitable for a stem cell transplant.
Thinking about the 2 populations within this group:
(i) older / frail patients who will never be fit for ASCT or alloSCT. Current options are brentuximab vedotin which, although associated with high overall response rates, is not a cure. Patients will therefore fail this treatment at some point. There are no good options at all for these patients. Palliative care or single agent, ineffective, chemotherapy are the alternatives. PD1 inhibitors offer a highly effective and well tolerated option. Nivolumab is licensed and has had a recent positive FAD for this indication.
(ii) patients who are fit for ASCT or alloSCT but who cannot get in a stable remission (despite the use of BV) so are not suitable for a transplant at that stage. Other combination chemotherapy regimens can be tried (e.g. GDP, miniBEAM, DHAP) but one patients have failed 2 or more lines of salvage therapy, it is unlikely they will respond stably to further chemotherapy. I would emphasise the intentiaon for these patients is to cure them – they are usually young and fit; frequently with new careers; young families; or in full time education. PD1 inhibitors are shown to be safe and effective in this patient group, offering a 60-70% chance of a PR or CR which is durable and offers a good chance of proceeding to ASCT or alloSCT. ***NB – nivolumab is not licensed and not available for this indication.***

2. Patients relapsing after ASCT and BV. Nivolumab is now available for these patients. I don't see any difference between nivolumab and pembrolizumab – I see

them as interchangeable for this patient group. Patients failing nivolumab would not be expected to respond to pembrolizumab and vice versa. The one advantage of pembrolizumab is that it's given once every 3 weeks (nivolumab is once every 2 weeks) therefore slightly reducing the burden on day treatment units.

PD1 inhibitors work differently from chemotherapy and responses can take longer to develop. Very occasionally a 'flare' response is seen so treating beyond initial 'progression' (when safe to do so) may benefit a patient having a flare response as they may go on and formally respond to the treatment. If a response is going to happen it will have happened by 6 months, so it would be reasonable to stop treatment at 6 months for those people who have failed to achieve a partial or complete metabolic remission.

The main trial of pembrolizumab is highly relevant to UK practise. The cohort that differed from nivolumab was for patients who had not had a transplant and who had failed BV. This represents a group who are very challenging to treat and for whom no other PD1 inhibitor is available.

The major outcomes in the trials were overall response, duration of response and PFS. These are very relevant endpoints and show pembro to be active, leading to responses which are, on the whole, very durable. It is important to appreciate that in the UK most patients will then be moved onto potentially curative treatment in the form of an alloSCT (more rarely ASCT). These patients will not get prolonged exposure to PD1 inhibitors. It is only in the rare, frail patient group that transplants will not performed and treatment with pembro could be prolonged. Yes, in my view high, durable response rates do act as a surrogate for overall survival as they lead to potentially curative transplants to occur.

The main side effect of concern are immune related (hypo/hyperthyroidism, diabetes, adrenalitis, pancreatitis, uveitis etc). Serious reactions occur in < 5% of patients. Due to increasing use of PD1 inhibitors for other oncology indications, experience of these side effects is increasing and local guidelines for management are being written and implemented.

The trials suggest that quality of life is well maintained on PD1 inhibitors (although the trials are not comparative).

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

No – I am not aware of other evidence.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)
Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

Implementation issues

The NHS is required by the Department of Health to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

As PD1 inhibitor use is widespread for other indications, implementation for relapsed Hodgkin (which is rare) will have little implication on health care resources.

Equality
NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts.

I know on no equality issues

NHS England submission on the NICE Technology Appraisal of pembrolizumab in the treatment of relapsed/refractory Hodgkin lymphoma (HL)

1. Pembrolizumab has a different marketing authorisation in HL to that of nivolumab in HL. Both are licensed for the treatment of relapsed/refractory HL after treatment with both autologous stem cell transplantation (SCT) and brentuximab and thus both could act either as a bridge to allogeneic SCT or be the sole treatment in those that are allogeneic SCT-ineligible. Pembrolizumab has an additional license in patients with relapsed/refractory HL who have had brentuximab and are ineligible for a SCT: in this second indication, pembrolizumab could potentially provide a bridge to autologous or allogeneic SCT or be the sole treatment in those that are SCT-ineligible.
2. Nivolumab should be a relevant comparator for cohort 1 as this is currently in baseline commissioning and has been funded ever since being first recommended by NICE in July 2017.
3. The Keynote-087 study had an inclusion criterion that restricted entry to patients with ECOG performance status (PS) 0 or 1. Thus all patients entering K-087 were of good performance status. The median ages of cohorts 1 and 2 were 34 and 42 respectively and thus it is likely that these cohorts were fit enough for further active treatment rather than best supportive care (BSC). In addition, the EPAR states that there were few patients in both cohorts who were potentially ineligible for SCT as a consequence of comorbidities. It is thus reasonable for the company not to have used BSC as a comparator given this high likelihood of further treatment in the populations studied. NHS England notes therefore with concern that the Royal College of Pathologists' submission indicates the potential wish to use pembrolizumab in HL to treat 'frail' patients as there is no evidence of such use in the company's submission or in the SPC/EPAR. The EPAR does however specifically note that even in patients with PS 1 there were higher rates of drug-related adverse events (AEs), grade 3-5 AEs, serious AEs and AEs leading to discontinuation compared to patients of PS 0. NHS England would therefore only wish to potentially commission the use of pembrolizumab in the treatment of HL in patients of PS 0 or 1.
4. NHS England notes that the K-087 trial stipulated that the maximum treatment duration of pembrolizumab was 2 years. As the K-087 trial is the major evidence base for both the licensing and the NICE appraisal of pembrolizumab in HL, NHS England would only commission the use of pembrolizumab in HL for a maximum of 2 years in treatment duration. It has already made such a direction to commissioning in other indications for both pembrolizumab and nivolumab.
5. NHS England notes that the median duration of follow up was 16 months and notes that very few patients were at risk after 15 months in the KM progression-free

survival plots for both cohorts. In addition, at data cut-off, [redacted] of patients were still on treatment in cohort 1 and [redacted] in cohort 2. The overall clinical data for efficacy of pembrolizumab in HL is therefore immature.

6. The K-087 trial did allow patients in complete remission to stop treatment at 24 weeks and potentially restart pembrolizumab at relapse if the trial was still open and if no other treatment had been used since the pembrolizumab discontinuation. The number of patients that did discontinue treatment in this way is small (only [redacted] in both cohorts) and follow-up is very immature in this very small group of patients. The EPAR states that it did not wish to make a treatment policy for such a discontinuation at 24 weeks on the basis of such a small amount of evidence.
7. The economic model assumes that patients proceeding to allogeneic SCTs did so after 12 weeks of treatment. The median times to response in the K-087 trial were [redacted] and [redacted] months in cohorts 1 and 2 and the mean figures were [redacted] and [redacted] months respectively. As response assessment is only likely to be in clinical practice in the 3rd month of treatment and it takes time to set up allogeneic SCTs, the likely time of allogeneic SCT is likely to be in the period of 12-24 weeks after initiation of pembrolizumab therapy. The economic model is therefore unrealistic I making this assumption of SCTs proceeding at 12 weeks and thus in these patients, the treatment duration of pembrolizumab will be longer than the company has modelled.
8. NHS England notes with concern that the economic model does not have a state of progressive disease post allogeneic SCT. Would that this was the case in clinical practice. As the Committee knows, the rate of progression free survival in the Lafferty paper was 54% in patients following allogeneic SCT.
9. The company's submission has a utility in the progressed state of [redacted] versus the preferred utility value of [redacted] of the ERG. Given the young age of the patients in the K-087 trial, the utility value for a considerable part of the progressed state must be significantly higher than [redacted].
10. NHS England notes that the K-087 trial excluded patients who had had an allogeneic SCT within the previous 5 years. This would not be applied to any commissioning of pembrolizumab in NHS England as it is expected that significant numbers of patients with relapsed HL post allogeneic SCT would be fit for treatment with pembrolizumab and there is no biological reason why such patients would not benefit as much as patients in K-087.
11. The HRG tariffs for chemotherapy administration used in the economic model are not the ones in the 2017/18 tariff, nor even the 2016/17 tariff.
12. NHS England notes that the figure for Life Years Gained in the economic model in cohort 1 for the non-pembrolizumab arm was 4.9 years and in cohort 2 was 4.8 years.
13. The license for pembrolizumab is limited to adults. Relapsed/refractory HL is also seen in patients aged less than 18 years and there is no biological reason why any NICE

recommendation as to the clinical and cost effectiveness of pembrolizumab for its indication in HL would not be valid in paediatric and teenager populations. In this situation, NHS England would ensure that the funding of pembrolizumab within baseline commissioning is extended to relevant patients under the age of 18 years.

[redacted]

[redacted] NHS England Chemotherapy Clinical Reference Group and [redacted] for the Cancer Drugs Fund

December 2017

NICE National Institute for Health and Care Excellence

Clinical expert statement

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Graham Collins**
2. Name of organisation	**Royal College of Pathologists/British Society for Haematology**

3. Job title or position	**Consultant Haematologist, Oxford University Hospitals NHS Foundation Trust and Lymphoma Lead for Thames Valley**
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To induce a remission in patients with relapsed / refractory Hodgkin Lymphoma. For most patients this would then allow a bridge to a potentially curative stem cell transplant.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	A partial or a complete remission. This is defined as a reduction in tumour size of 50% or more.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Nivolumab is a very similar drug which we are able to use now in the relapse post ASCT setting. However we are currently NOT able to use this in patients who have not had a stem cell transplant. This is where I see the main use of pembrolizumab as the trial included a cohort of patients who had not had a stem cell transplant. These patients have no good treatment option and pembrolizumab represents a very active option.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	At relapse (or when chemorefractoriness is diagnosed) patients receive 1st line salvage chemotherapy. If this works well (defined by a complete metabolic remission – CMR - on PET scan), patients proceed to autologous stem cell transplant (ASCT) which cures around 60-70% of patients. If however a good remission is not achieved after 1st lines salvage, 2nd line salvage is used (most centres use brentuximab vedotin in this setting although this is currently being evaluated by NICE). If a CMR is obtained with 2nd line salvage most patients then have an autologous stem cell transplant. If however a good remission is still not obtained, options are limited. This is where pembrolizumab may have a defined role in the UK treatment pathway. For those who have had an ASCT and relapse, treatment is usually with brentuximab vedotin (if not had before, if had before with a very good response). The intention is usually to bridge to an allogeneic stem cell transplant. For those who don't respond well, nivolumab is now approved for use. Nivolumab and pembrolizumab are very similar drugs and I wouldn't expect a big difference between them in this very particular setting.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	The BCSH guidelines are used but are rather out of date as they were written before any of the PD1 literature was published.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Generally it's well defined although some differences exist. E.g. - exactly what first and 2nd line salvage treatment is used does vary between centre and there is no randomised data to guide the decisions. - all centres would proceed to ASCT if the patient is PET negative after 1st line salvage - some centres would proceed to alloSCT if in a CMR after 2nd line salvage (although most proceed to ASCT).

• What impact would the technology have on the current pathway of care?	The particular area where pembrolizumab would be hugely helpful, is in those patient who cannot get into a good remission PRIOR to stem cell transplant and who have failed 2 lines (or more) of salvage therapy.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Pembrolizumab is not being used at all for relapsed Hodgkin outside of trials in the UK.
• How does healthcare resource use differ between the technology and current care?	
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary or Tertiary care units.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No special training is needed as nivolumab is being used already and is a very similar drug.
12. Do you expect the technology to provide clinically	Yes – for those failing 2 or more lines of salvage treatment who have not had a stem cell transplant.

meaningful benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	Yes – if it enables bridge to a stem cell transplant.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes – PD1 inhibitors as a single agent are well tolerated compared with further chemotherapy.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	None in addition to those specified above.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare	Easier. Pembrolizumab is very straightforward to use. Short infusion, no pre-med.

professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	It would be reasonable to say that the treatment should be stopped if not response is seen after 12 weeks.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the	The complexity (as always with relapsed Hodgkin) is that many patients will be bridged to a stem cell transplant – most alloSCT. This can be very hard to factor into QoL analyses.

quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes – PD1 inhibitors are VERY active in relapsed Hodgkin. There is little available for patient failing 2 lines of savage chemotherapy. NB – this is demonstrated well by the recent data collection that I helped publish. See Br J Haematol. 2017 Nov;179(3):471-479. doi: 10.1111/bjh.14898. Epub 2017 Aug 31. **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.**

	Eyre TA[1], Phillips EH[2], Linton KM[3], Kassam S[4], Gibb A[3], Allibone S[3], Radford J[3], Peggs K[2], Burton C[5], Stewart G[5], LeDieu R[6], Booth C[6], Osborne WL[7], Miall F[8], Eyre DW[9], Ardeshna KM[2], Collins GP[1].
• Does the use of the technology address any particular unmet need of the patient population?	Yes – see the reference above for the unmet need for those failing brentuximab in the pre-transplant setting.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	PD1 inhibitors can cause serious immune related side effects but these are rare.
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	
• If not, how could the results be extrapolated to the UK setting?	There is a relative lack of patients being bridged to an alloSCT. US practise is going against using alloSCT whereas UK practise is to use more alloSCT.

• What, in your view, are the most important outcomes, and were they measured in the trials?	For an alloSCT, a durable remission needs to be achieved. So ORR, DOR / PFS and toxicity are the most important outcomes.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	See reference quoted above.
21. How do data on real-world experience compare with the trial data?	There is no real world pembro data as it is not being used in the UK currently in Hodgkin.

Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	
23. In up to 5 bullet points, please summarise the key messages of your statement. • Unmet need for pre-transplant patients failing 2 or more lines of salvage • Pembrolizumab trial had a cohort who had NOT had a stem cell transplant • Pembrolizumab would be invaluable in the setting of failing 2 lines of salvage prior to SCT • Safe •	

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Debra Fayter, Systematic Reviewer, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC
Bram Ramaekers, Health Economist, Maastricht UMC
Svenja Petersohn, Health Economist, Maastricht UMC
Rob Riemsma, Reviews Manager, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Xavier Pouwels, Health Economist, Maastricht UMC
Willem Witlox, Health Economist, Maastricht UMC
Caro Noake, Information Specialist, KSR Ltd, UK
Gill Worthy, Statistician, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 14/11/2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 16/119/02.

Declared competing interests of the authors
None.

Acknowledgements
None

[redacted]

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Fayter D, Grimm S, Ramaekers B, Petersohn S, Riemsma R, Armstrong N, Witlox W, Pouwels X, Noake C, Worthy G, Joore M, Kleijnen J. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma; A Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2017.

Contributions of authors
Debra Fayter acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Sabine Grimm acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Bram Ramaekers, Svenja Petersohn, Xavier Pouwels, Willem Witlox and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Rob Riemsma acted as a systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore critiqued the manufacturer's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

ABVD regimen	Doxorubicin, bleomycin, vinblastine and dacarbazine
AE	Adverse Events
AEOSI	Adverse events of special interest
AIC	Akaike Information Criterion
AlloSCT	Allogeneic Stem Cell Transplant
ASaT	All Subjects as Treated
ASCO	American Society of Clinical Oncology
ASHAP	Doxorubicin, methylprednisolone, cytarabine, cisplatin
AUC	Area under the curve
AutoSCT	Autologous Stem Cell Transplant
BCSH	British Committee for Standards in Haematology Guidelines
BEACOPP regimen	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone
BIC	Bayesian information criterion
BICR	Blinded independent central radiologists'
BNF	British National Formulary
BOR	Best overall response
BSA	Body surface area
BSC	Best supportive care
BTD	Breakthrough Therapy Designation
BV	Brentuximab Vedotin
C	Cirrhotic
CADTH	Canadian Agency for Drugs and Technologies in Health
CAA	Commercial access agreement
CDF	Cancer Drugs Fund
CE	Cost Effectiveness
CEA	Cost effectiveness Analysis
CEAC	Cost effectiveness Acceptability Curve
cHL	Classical Hodgkin Lymphoma
CHMP	Committee for Medicinal Products for Human Use
CHOP	cyclophosphamide, doxorubicin, prednisolone, vincristine
CI	Confidence Interval
CMU	Commercial medicines unit
CPS	Combined positive score
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CSR	Clinical study report
CT	Computer Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DAA	Direct-acting antivirals
DAE	Discontinuation due to adverse events
DHAOx	Dexamethasone, cytarabine, oxaliplatin
DHAP	Dexamethasone, cytarabine, cisplatin
DoH	Department of Health
DOR	Duration Of Response
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal Growth Factor Receptor
EMA	European Medicines Agency
EPAR	European public assessment report
EORTC-QLQC30	European Organisation for Research and Treatment Cancer Quality of Life Questionnaire

EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESHAP	Etoposide, methylprednisolone, cytarabine, cisplatin
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
FAS	Full analysis set
FDA	Food and Drug Administration
GDP	Gemcitabine, dexamethasone, cisplatin
GEM-P	Gemcitabine, cisplatin, methylprednisolone
GVD	Gemcitabine, vinorelbine, liposomal doxorubicin
GVHD	Graft Versus Host Disease
HL	Hodgkin Lymphoma
HR	Hazard ratio
HRG	Healthcare Resource Group
HRQoL	Health-related Quality of Life
HTA	Health Technology Assessment
IC	Indirect Comparison
ICE	Ifosfamide, carboplatin, etoposide
ICER	Incremental Cost-effectiveness Ratio
ICTRP	International Clinical Trials Registry Platform
IRG	Independent Review Group
ISPOR	International Society for Pharmacoecomics and Outcomes Research
ITT	Intention to Treat
IV	Intravenous
IVE	Ifosfamide, epirubicin, etoposide
IVOx	Ifosfamide, etoposide, oxaliplatin
IWG	International Working Group
KM	Kaplan-Meier
KSR	Kleijnen Systematic Reviews
LY	Life Year
mAB	Monoclonal antibody
MAIC	Matched Adjusted Indirect treatment comparison
MeSH	Medical Subject Headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MINE	Mitoxantrone, ifosfamide, vinorelbine, etoposide
MK-3475	Pembrolizumab - Keytruda®
MS	Manufacturer's Submission
MSD	Merck Sharp and Dohme
MTC	Mixed Treatment Comparison
NA	Not applicable
NHL	Non Hodgkin Lymphoma
NHS	National Health Services
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not Reported
ORR	Objective Response Rate
OS	Overall survival
PD	Progressive Disease
PD-1	Programmed death 1 protein
PD-L1	Programmed death ligand 1
PET	Positron Emission Tomography
PFR	Progression-free rate
PFS	Progression free survival
PI	Principal Investigator

PIM	Promising Innovative Medicines
PK	Pharmacokinetics
PMitCEBO	Bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient-reported outcome
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
PSSRU	Personal and Personal and Social Services Research Unit
Q3W	Every 3 weeks
QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of life
RCT	Randomised Controlled Trial
RR	Response Rate; Relative Risk; Risk Ratio
RRcHL	Relapsed or refractory classical Hodgkin Lymphoma
RSC	Reed-Sternberg cells
RVIG	Gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine
SAE	Serious Adverse Events
SCT	Stem cell transplant
SD	Stable Disease; Standard deviation
SG	Standard Gamble
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SOC	Standard of Care
STA	Single Technology Appraisal
TA	Technology Appraisal
ToT	Time on Treatment
TTO	Time trade off
UK	United Kingdom of Great Britain and Northern Ireland
UMC	University Medical Centre
VAS	Visual Analogue Scale
VAT	Value-Added Tax

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The population of this appraisal is in line with the NICE scope. The main trial in the company submission (CS) (KEYNOTE-087) covers both cohorts of interest (cohort 1: people with relapsed or refractory classical Hodgkin Lymphoma (cHL) who have received autologous stem cell transplant (autoSCT) and brentuximab vedotin (BV) and, cohort 2: those who have received BV when autoSCT is not a treatment option). However only 14 patients in the trial were from the UK. None of the patients in the comparator study (Cheah et al. 2016) were from the UK. The comparator study in this appraisal was also used in a previous appraisal (TA462). NICE concluded in TA462 that "the comparator data may not fully represent UK clinical practice".

The intervention (pembrolizumab) is in line with the scope. Regulatory approval by the EMA for the indication considered within this submission was granted on the 2nd May 2017.

The description of the comparators in the NICE scope is as follows:

- Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin
- Best supportive care.

The company uses one retrospective USA database study as a comparator. In this study patients received the following types of therapy: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or allogeneic SCT (alloSCT), or other treatment. The company has not provided separate data for comparators; instead a combined data set has been provided for multiple comparators, some of which are within the scope and others not. This combined data set was used as a comparator for both populations, cohort 1 and cohort 2.

The company's submission matches the NICE scope on outcome measures. The primary outcome in the KEYNOTE-087 trial is overall response rate (ORR). Although progression-free survival and overall survival are investigated, as per the NICE scope, the data for these outcomes are not fully mature.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company did not identify any randomised controlled trials of pembrolizumab and its comparators in patients with classical Hodgkin Lymphoma who have either received autoSCT and BV or BV alone due to autoSCT being unsuitable. One ongoing, single arm study of the efficacy and safety of pembrolizumab was identified (KEYNOTE-087) and this formed the basis of the submission. KEYNOTE-087 includes 150 patients (14 UK patients) relevant to this appraisal. It covers both cohorts of interest (cohort 1: people with relapsed or refractory cHL who have received autologous stem cell transplant and brentuximab vedotin and cohort 2: patients who have received brentuximab vedotin when autologous stem cell transplant is not a treatment option). The company presented data based on a median follow up of 15.9 months. The median time on treatment was [redacted] days for cohort 1 and [redacted] days for cohort 2.

The primary outcome of KEYNOTE-087 was overall response rate (ORR) as assessed by independent committee. ORR was 75.4% in cohort 1 and 66.7% in cohort 2 over the course of the trial. Median progression free survival (PFS) in cohort 1 was 16.7 months (11.2 to NR). In cohort 2 it was 11.1 months (7.6 to 13.7). Median overall survival (OS) was[redacted]. At 12 months survival was [redacted] in cohort 1 and [redacted] in cohort 2. In cohort 1 [redacted] of patients had one or more adverse events. In cohort 2 [redacted] of patients had one or more adverse events. The company noted that

most AEs were low grade ([redacted] and [redacted] grades 3 to 5 in cohort 1 and 2 respectively). In cohort 1 [redacted] of AEs were classed as serious and in Cohort 2 [redacted]. The most common adverse events were pyrexia, cough, fatigue, diarrhoea and vomiting. The company conducted post-hoc analyses of response at 12 weeks to use in the comparison of clinical and cost effectiveness. Overall response rates were lower at 12 weeks than over the whole course of the trial ([redacted]

As KEYNOTE-087 did not have a comparator group the company identified a comparative observational study from the literature (Cheah 2016 et al). This is a retrospective USA database study in which patients received the following types of therapy: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or alloSCT, or other treatment. The company has not provided separate data for comparators; instead a combined data set has been provided for multiple comparators.

The company performed two types of analyses: a naïve indirect comparison between KEYNOTE-087 and Cheah and a matched adjusted indirect treatment comparison (MAIC) of the two studies. With the exception of one of the naïve comparisons, all results significantly favoured pembrolizumab over SoC for ORR and PFS.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The CS and response to clarification provided sufficient details for the ERG to appraise the searches for eligible trials and to critique the clinical effectiveness of pembrolizumab. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4 using a good range of databases. Additional searches of conference proceedings, trials database and the NICE website were reported.

The clinical effectiveness of pembrolizumab submitted in this appraisal is based on the KEYNOTE-087 study. The most important methodological aspect to note is that although this trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative study which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention as the role of natural history and baseline characteristics is not taken into account. As treatment is known to participants, clinicians and assessors this can lead to bias in the delivery of the intervention and the reporting of outcomes. Other limitations in applying the results of the trial to UK practice include the fact that median progression free survival data were immature and median overall survival [redacted]. The trial has only 150 relevant participants so the evidence base for this appraisal is small. Patients over 65 are not well represented. Furthermore, a small number of patients were from the UK (14) so the trial may not totally reflect the UK population and setting. It is recognised, however, that the population matching the scope of this appraisal from which to draw participants is in itself small. In clinical practice, for those who are suitable, pembrolizumab represents a bridge to alloSCT, a potentially curative treatment. However, the company submission stated that 'KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2.' The company further stated that 'the use of stem cell transplant would have been at the discretion of the treating physician on a per patient basis.'

The main comparative study is a US observational study with a range of different treatments both within and outside the NICE scope. In the previous appraisal of nivolumab (TA462), the committee considered that "the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision-making, but overall the clinical effectiveness of nivolumab compared with standard of

care was highly uncertain because the comparator data may not fully represent UK clinical practice." However, the ERG is not aware of a more appropriate source of data for the comparator population for this appraisal.

The ERG identified problems with compatibility of the two studies in the CS regarding baseline characteristics and methods of outcomes assessment. In the MAIC the company adjusted for potential confounding variables so that the KEYNOTE-087 study more closely resembled the Cheah study. According to DSU report 18, unanchored indirect comparisons (i.e. those based on single-arm studies) are susceptible to large amounts of systematic error unless all prognostic variables and effect modifiers are accounted for. However, in the current MAIC the company was dependent on the variables reported in Cheah et al. (2016) and these are unlikely to be all relevant prognostic variables and effect modifiers. Therefore, the results are likely to contain systematic error but it is not possible to estimate the size of the potential error. Both the naïve IC and MAIC have major limitations for decision making.

1.4 Summary of cost effectiveness submitted evidence by the company

The company developed a de novo cohort state transition model with health states based on response, uptake of alloSCT, and survival. The model structure consists of a short term component (first 12 weeks), a subsequent decision tree element (at 12 weeks) to determine the proportion of patients transiting to alloSCT (conditional upon response at 12 weeks) and a long-term component (after the first 12 weeks) separately for patients who had alloSCT and patients who did not have alloSCT at 12 weeks. At 12 weeks, patients were allocated to alloSCT based on their response status and probabilities of alloSCT uptake were applied conditional on patients' response status. Any alloSCTs were assumed to happen at this 12-week time point, without any lag. Justifying their approach, the company believed that alloSCT data from KEYNOTE-087 were not reflective of UK clinical practice and that they did not have Kaplan-Meier data for time-to-alloSCT from Cheah et al.

In line with the marketing authorisation and the final scope issued by NICE, two distinct populations were considered in the cost effectiveness model: patients with RRcHL who have failed autoSCT and BV (cohort 1) and patients with RRcHL who are autoSCT ineligible and have failed BV (cohort 2).

Pembrolizumab monotherapy is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration for RRcHL (i.e. administered intravenously at a fixed dose of 200 mg over 30 minutes every three weeks [Q3W]). The company assumed that in the model pembrolizumab monotherapy will be provided for a maximum of 24 months (35 cycles).

The company only considered "standard of care" (SoC) as comparator in its base-case. SoC as considered by the company consists of the following regimens: chemotherapy, bendamustine or investigational agents. The distribution of patients among these regimens was based on the distribution observed in Cheah et al (2016). The company also presented a scenario analysis, in which best supportive care (BSC) was added as a comparator. The company justified this deviation from the scope (i.e. not including BSC in its base-case) with their belief that BSC use would be minimal as eligible patients are likely to receive therapy whenever feasible.

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The cycle length was one week to account for the length of treatment cycles. A half-cycle correction was applied. A time horizon of 40 years was adopted to capture all relevant costs and outcomes. All costs and utilities were discounted at a rate of 3.5% per year.

Treatment effectiveness for pembrolizumab was primarily based on the KEYNOTE-087 study. The primary data source for the SoC comparator was the Cheah et al (2016) study. The naïve indirect

treatment comparison was used to inform relative overall survival (OS), progression-free survival (PFS) and response rates at week 12. The MAIC was only used in scenario analysis. Both KEYNOTE-087 cohorts were compared with the Cheah et al (2016) study cohort.

Due to the company's model structure, treatment effectiveness and time to treatment discontinuation (TTD) were estimated for the pre-12 week period and for the post-12 week period separately. Parametric models were fitted to the entire study data from KEYNOTE-087 to estimate OS and PFS for patients receiving pembrolizumab in the pre-12 week period. To inform the decision tree element at week 12, response rates from KEYNOTE-087 were used, as well as two clinician surveys to inform estimates of alloSCT uptake conditional on response status. For the post-12 week period, treatment effectiveness depended on whether patients received alloSCT or not. Mortality post-alloSCT was based on Lafferty et al (2017) and post-progression mortality for patients who did not receive alloSCT was based on Cheah et al (2016). The company justified the use of different data sources by stating that survival data from KEYNOTE-087 were immature.

TTD for patients treated with pembrolizumab for the pre-12 week period was assumed to be equivalent to PFS. TTD for the post-12 week period was estimated directly from KEYNOTE-087. Furthermore, TTD for SoC was assumed equivalent to PFS in Cheah et al. TTD for pembrolizumab was capped at 24 months.

Health-related quality-of-life (HRQoL) was measured in KEYNOTE-087 at different time points, but only responses from week 12 were used to obtain health state utility values, ignoring observations at other time points. The company calculated utility values stratified by response and response rates at 12 weeks to obtain progression-free health state utilities, and used response rates from Lafferty et al to calculate the post-alloSCT utility. The company did not use the progressed disease utility score from KEYNOTE-087 and instead opted to use a utility decrement from Swinburn et al (2015).

The electronic market information tool (eMit) was used to acquire drug acquisition costs of pembrolizumab and components of SoC. When these were unavailable, costs from the British National Formulary were used. Administration costs were obtained from the NHS reference costs. The list price of 200 mg pembrolizumab was £5,260. Through a Commercial Access Agreement (CAA), [redacted]. The cost for SoC was assumed to consist of acquisition and administration costs for the different chemotherapy regimens (equal use assumed), and bendamustine. Health state costs consisted of monitoring costs and outpatient attendance. For the post-alloSCT health state, a one-off cost was applied.

In the deterministic base-case analysis, total QALYs and LYs gained as well as total costs (with the CAA) were larger in the pembrolizumab treatment arm compared to UK SoC in both cohorts. Incremental costs mainly stemmed from differences in acquisition costs and alloSCT costs between pembrolizumab and SoC. Pembrolizumab treatment resulted in deterministic incremental cost effectiveness ratios (ICERs) of £43,511 and £48,571 per QALY gained for cohort 1 and cohort 2 respectively, as per the company's corrected base-case.

The company performed probabilistic sensitivity analysis (PSA) and deterministic sensitivity analyses (DSA). The PSA with 1,000 iterations resulted in ICERs of £43,653 and £50,894 per QALY gained for cohorts 1 and 2 respectively for pembrolizumab versus SoC. The explored scenarios resulted in significant changes to the ICERs in both cohorts.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

In the absence of cost effectiveness studies performed on the population and intervention of interest from the literature, the ERG agreed that a de novo approach to modelling cost effectiveness of pembrolizumab was necessary. However, it was unclear why the company did not provide a complete overview of the publications included and excluded from their cost effectiveness, cost and resource and utility and HRQoL systematic literature reviews (SLRs). The company prioritised aligning their sources with TA462 over using the results of their SLRs.

No justification was provided for the model structure only allowing patients to have alloSCT at 12 weeks after starting treatment, thereby ignoring responses that can occur at later time points (as acknowledged by the company). The alloSCT at 12 weeks assumption furthermore neglects the time required to identify a donor and schedule the procedure. This entails that alloSCT in the present model is performed earlier than would be expected in clinical practice. Hence, the post-alloSCT benefits are applied earlier, which favours pembrolizumab. The company failed to include a post-alloSCT progressed disease health state in their model, not in line with evidence from Lafferty et al, thereby also favouring pembrolizumab.

The populations described by the company are consistent with the final scope issued by NICE for this appraisal. For KEYNOTE-087, the company was able to distinguish between patients who did and did not receive autoSCT (i.e. cohort 1 and 2 respectively). The company did not have access to the individual patient level data in Cheah et al and hence used the mixed population for comparisons with both cohorts. This likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively.

BSC was not incorporated in the CS base-case (inconsistent with the scope), but only presented in a scenario analysis. Moreover, nivolumab was recently recommended by NICE in part of this population (cohort 1) and may become a relevant comparator in the future.

The assumption that pembrolizumab monotherapy will be stopped after 24 months is inconsistent with the SmPC but in line with the KEYNOTE-087 protocol. It is unclear whether pembrolizumab, in UK clinical practice, would also be provided for a maximum of 24 months. Removing this cap resulted in substantially increased ICERs for both cohorts, showing that the company's base-case might under-estimate the cost incurred with the use of pembrolizumab if a 24-months stopping rule is not enforced in clinical practice.

The ERG considered the adopted perspective and discounting to be appropriate for this appraisal.

Treatment and relative treatment effectiveness used in the model relied on the use of evidence from single-arm studies and a naïve indirect comparison. There is therefore substantial uncertainty about relative treatment effectiveness. The use of the naïve indirect comparison instead of the MAIC favoured SoC.

The combining of survey results to inform alloSCT uptake conditional on response status was viewed as inappropriate considering that the company acknowledged that it was possible for both surveys to include the same clinical experts. The company omitted the result from its survey that patients with progressed disease could still be eligible for alloSCT. Both assumptions favoured pembrolizumab.

Post-12 week mortality data from KEYNOTE-087 was deemed immature by the company and the ERG agreed with this assessment. [redacted] and the ERG considers that these may be informative for the present model. Furthermore, the ERG was

concerned about the use of Lafferty et al, given its small sample size and the questionable generalisability to UK clinical practice. The company's method used for extrapolating OS post-alloSCT was deemed by the ERG to over-estimate OS, which favoured pembrolizumab. There was also significant uncertainty around extrapolating PFS post-12 weeks, which translated into significant increases in the ICERs when alternative parametric survival models were chosen in both cohorts.

The mixed effects model utilities provided in response to the clarification letter, were deemed by the ERG to make better use of the KEYNOTE-087 data. The ERG preferred estimating the progressed disease (PD) utility from KEYNOTE-087, rather than Swinburn et al. The ERG considered the proportion of responders used for calculating utility values as inconsistent.

The ERG was concerned about the assumption that all chemotherapy agents contributed equally to the mix of SoC in calculating costs. This likely favoured pembrolizumab. Resource use and costs associated with alloSCT were likely under-estimated in the model, also favouring pembrolizumab.

Cost effectiveness results were not presented for BSC in the base-case. The number of iterations (1,000) in the PSA was likely too small to achieve stable results.

The ERG also had concerns about model validation, mostly relating to the lack of cross-validation with TA462 and the irreproducibility of model estimates used for external validation.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

Overall, the CS reported searches were well presented and easily reproducible. Searches were carried out on a good range of databases. The clinical effectiveness strategies utilised a recognised study design filter. Supplementary searches of conference proceedings and the NICE website were undertaken by the company, along with a manual search of the WHO ICTRP trial database in order to identify additional on-going trials. The clinical evidence is based on a well conducted, multicentre single-arm trial reflecting both cohorts of patients relevant to the decision problem. Outcomes assessed reflect the scope.

Overall, the model is well built and transparent. The company reflected that pembrolizumab can be considered as a bridging treatment to alloSCT by incorporating alloSCT in the economic model. The company provided alternative data (for example derived from the MAIC) and alternative survival functions to enable exploratory analyses in the model.

1.6.2 Weaknesses and areas of uncertainty

The ERG had some concerns about the language bias of restricting clinical effectiveness searches to English language only as this is not in line with current best practice. However, the main weakness of this appraisal is the lack of relevant randomised controlled trials (RCTs). Outcomes relating to pembrolizumab are based on a single arm trial. Comparisons with the comparators in the scope are problematic due to the availability of only one US study with a mix of different treatments. The naïve and matched adjusted comparisons conducted by the company have a number of limitations and represent a much weaker level of evidence than a RCT. Additionally progression-free survival and overall survival data are not fully mature. KEYNOTE-087 is an ongoing trial so more information will be available in future regarding uncertainties in progression-free and overall survival.

The model structure did not appropriately reflect the timing of the alloSCT decision and the timing of the actual alloSCT procedure. The model therefore under-estimates the time to alloSCT and assumes

that any benefits will be obtained sooner than is likely to occur in clinical practice. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT. These assumptions favour pembrolizumab.

The company informed alloSCT uptake conditional on response status at 12 weeks after treatment start through a UK clinician survey and then combined these survey results with the previously performed BMS survey results (from TA462). The appropriateness of combining both surveys is questionable. The appropriate approach for incorporating alloSCT in the model would have been to use time to alloSCT data directly from the main source of evidence. There remains major uncertainty about the alloSCT uptake estimates. Furthermore, the elicited alloSCT uptake (from the MSD survey) for patients with progressed disease was ignored. Both, the combining of both surveys and ignoring alloSCT uptake in progressed disease patients, were shown in scenario analysis to be major drivers of cost effectiveness.

A major limitation was the use of single-arm evidence to inform treatment effectiveness. There was uncertainty whether the MAIC or the naïve indirect comparison should be used. The company provided both and the ERG, like the company, used the naïve indirect comparison in the base-case and the MAIC in scenario analysis. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation as this necessitated the use of post-alloSCT OS and utility estimates from alternative data sources, one of which was based on 13 patients only. The methods used to extrapolate from this data source were also questionable. [redacted], and the ERG considers that these may be informative for the present analysis.

It is of note that the population used for the comparator was a mixed population of cohorts 1 and 2, that is, it included patients who did and did not receive autoSCT, derived from Cheah et al. The Cheah et al. population is more comparable with KEYNOTE-087 cohort 1 than with cohort 2 in terms of patient characteristics. The use of this mixed comparator population likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively, but this could not be formally explored in scenario analysis.

Of further note, the economic model, and the evidence from KEYNOTE-087, rely on the assumption that treatment with pembrolizumab is capped at 24 months, which is inconsistent with its SmPC. It is unclear whether in UK clinical practice pembrolizumab would also be provided for a maximum of 24 months. This assumption favoured pembrolizumab.

Model extrapolations lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A number of issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in ICERs (probabilistic) of pembrolizumab (with confidential access agreement (CAA)) versus SoC of £64,186 and £78,696 per QALY gained for cohorts 1 and 2 respectively.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. The scenarios with the largest impact were alternative assumptions for extrapolating post-alloSCT, an alternative survival model for extrapolating post-12-week PFS in cohort 2, the use of the MAIC instead of the naïve comparison and removing the cap of 24 months on TTD (Table 1.1).

Table 1.1. ERG base-case and exploratory analyses

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company corrected base-case cohort 1	Pembrolizumab	£107,459	4.497			
	SoC	£52,017	3.223	£55,442	1.274	£43,511
ERG base-case cohort 1	Pembrolizumab	£107,998	4.460			
	SoC	£50,913	3.535	£57,085	0.925	£61,705
Use of MAIC (2) cohort 1	Pembrolizumab	£107,998	4.460			
	SoC	£47,997	3.359	£60,001	1.102	£54,466
No 24-months cap on TTD (3) cohort 1	Pembrolizumab	£123,990	4.460			
	SoC	£50,913	3.535	£73,077	0.925	£78,992
Alternative OS post-alloSCT assumption (5)	Pembrolizumab	£107,030	3.558			
	SoC	£50,157	2.830	£56,873	0.727	£78,204
Company corrected base-case cohort 2	Pembrolizumab	£93,732	4.072			
	SoC	£51,424	3.200	£42,308	0.871	£48,571
ERG base-case cohort 2	Pembrolizumab	£93,095	4.118			
	SoC	£50,609	3.541	£42,486	0.577	£73,594
Alternative distributions (1.b) cohort 2	Pembrolizumab	£92,556	3.995			
	SoC	£50,550	3.529	£42,007	0.466	£90,152
Use of MAIC (2) cohort 2	Pembrolizumab	£93,095	4.118			
	SoC	£45,924	3.337	£47,171	0.781	£60,372
No 24-months cap on TTD (3) cohort 2	Pembrolizumab	£96,380	4.118			
	SoC	£50,609	3.541	£45,771	0.577	£79,284
Alternative OS post-alloSCT assumption (5)	Pembrolizumab	£92,204	3.287			
	SoC	£49,863	2.844	£42,341	0.442	£95,712

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

2. BACKGROUND

In this report the ERG provides a review of the evidence submitted by Merck Sharp & Dohme (MSD) in support of pembrolizumab, trade name KEYTRUDA®, for the treatment of patients with relapsed or refractory Classical Hodgkin lymphoma (cHL). In this section we outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken from Chapter 3 of the company submission (CS)[1] with sections referenced as appropriate.

2.1 Critique of company's description of underlying health problem

The underlying health problem of this appraisal is Classical Hodgkin Lymphoma which the company describes as 'a rare, localised or disseminated, malignant proliferation of cells of the lymphoreticular system, occurring mostly in lymph node tissues, spleen, liver and bone marrow.'[1]

The CS clarifies that Classical Hodgkin Lymphoma is the predominant subgroup of Hodgkin Lymphoma and accounts for 95% of cases of the disease. The presence of Reed-Sternberg cells in Hodgkin Lymphoma is highlighted.

There are four subtypes of Classical Hodgkin Lymphoma: nodular sclerosing (60%) which is usually identified early due to swelling of the lymph nodes in the neck; lymphocyte rich (20%), mixed cellularity (15%) and lymphocyte depleted (very rare).[2] Patients may present with bulky disease. This is defined as a lymph node that is 10cm or more or a lymphoma in the centre of the chest (mediastinum) which is at least one third of the width of the chest.[2]

The company highlights the symptomatic burden of cHL and that patients with B symptoms (presence of fever, weight loss and drenching night sweats) are associated with worse outcomes.

The CS cites Cancer Research UK data that states that in 2014 there were 2,106 new cases of Hodgkin Lymphoma in the UK. The CS also states that according to Cancer Research UK data incidence rates may increase by 5% in the UK population overall between 2014 and 2035.[1]

The company highlights that incidence of Hodgkin Lymphoma peaks in young adults (20 to 24 years of age) and older males and females (75 to 79 years of age) with approximately half of diagnoses reported in people aged 45 and over.

The company describes the survival rates for HL as promising with rates of 91.4% at one year, 85.0% at five years and 80.4% at 10 years. However, they caution that the relapsed/refractory population under consideration for this appraisal are likely to have a poorer prognosis compared with patients who respond to therapy. The company mention a retrospective trial of 81 patients showing a five year survival of less than 20%.[3]

The CS refers to the burden of costs affecting patients, caregivers and society. It is noted that there is a relatively high proportion of patients with Hodgkin Lymphoma who are of working age.

ERG comment:

- The company provides a good overview of the underlying health problem. The ERG checked the references provided to support the statements in the company submission. In general, these were found to be appropriate.
- The population in this appraisal is specifically people with relapsed or refractory cHL who have received autologous stem cell transplant (autoSCT) and brentuximab (BV) or BV when autoSCT is not a treatment option.

2.2 Critique of company's overview of current service provision

The company correctly reports that there is no NICE guideline on relapsed/refractory CHL.

For first line therapy chemotherapy alone or chemotherapy combined with radiotherapy is used in practice. Between 15 and 30% of patients with HL do not achieve remission with these treatments.[4] The CS outlines that those patients who do not achieve remission may be offered chemotherapy and/or radiotherapy to enable autoSCT. AutoSCT is potentially curative and effective in about 50% of people.[4] However autoSCT may not be an option for some patients if their disease does not respond adequately to treatment or the patient's age or comorbidities prevent offering it as an option.

The CS highlights the recent approval of brentuximab vedotin (BV) for patients with relapsed or refractory disease after autoSCT or those who have had at least two prior therapies if the patient cannot have (autoSCT) or multi-agent chemotherapy.[5]

The company state that 'for those who do not respond to BV the prognosis remains poor with little / no treatment options.'[1] There is no standard therapy after autoSCT and BV.[4] BV can be used as retreatment according to its licence but no specific recommendations have been made by NICE regarding retreatment.[5] Single or combination treatments including different chemotherapy regimens (some outside their marketing authorisation) may be used. This is the point in the clinical pathway at which pembrolizumab is aimed.

Pembrolizumab is therefore at least a third line treatment for people with relapsed or refractory cHL who have received autologous stem cell transplant (autoSCT) and brentuximab (BV) or BV when autoSCT is not a treatment option. For those who are suitable, pembrolizumab represents a bridge to allogeneic SCT (alloSCT), a potentially curative treatment.

ERG comment:

- The company's overview of current service provision is appropriate and relevant to the decision problem under consideration.
- Although not listed as a comparator in the NICE scope and not referenced in the CS, nivolumab has recently received approval from NICE for this condition. It is recommended 'as an option for treating relapsed or refractory classical Hodgkin lymphoma in adults after autologous stem cell transplant and treatment with brentuximab vedotin.'[6] Nivolumab is, however, not recommended for one of the populations in this appraisal (those who have received BV but who have not received an autoSCT).

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
Population	People with relapsed or refractory classical Hodgkin Lymphoma who have received: • autologous stem cell transplant and brentuximab vedotin • brentuximab vedotin when autologous stem cell transplant is not a treatment option.	As per final scope	This is in accordance with the scope.
Intervention	Pembrolizumab	As per final scope	This is in accordance with the scope.
Comparator(s)	• Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin • Best supportive care.	Standard of care as per Cheah et al. 2016) including: • Investigational agent • Gemcitabine • Bendamustine • Other alkylatory • BV retreatment • Platinum based • autoSCT • Other Cheah et al. 2016 reported outcome data for a mix of chemotherapy regimens and was preferred by the ERG in TA462. To separate individual regimens survival outcome data would not have been possible in the absence of individual patient level data and hence conservatively MSD have included all survival outcomes reported here.	Not in line with the final scope. The company has not provided separate data for comparators; instead a combined data set has been provided for multiple comparators, some of which are within the scope and others not. In TA462[6],"The committee concluded that the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision-making, but overall the clinical effectiveness of nivolumab compared with standard of care was highly uncertain because the comparator data may not fully represent UK clinical practice."

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life	As per final scope, with the exception of long term overall survival data. The model structure utilised OS data from week 0-12 from KEYNOTE-087, response rates at week 12, PFS from week 12 onward and external literature OS sources for post alloSCT survival. At follow up (15.9 month), there were insufficient mortality events and median OS [redacted]. Hence all available data from KEYNOTE-087 has been utilised where possible.	Mostly in line with the final scope. However, survival data (OS and PFS) are immature. In addition, only two outcomes have been included in the indirect comparison: PFS and ORR.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.	As per final scope	In line with the scope. However, a minor remark, the time horizon of 40 years was too short to capture the lifetime of all patients. A time horizon of 50 years, which was sufficiently long, was used in scenario analysis. Furthermore, Best Supportive Care was not presented as a comparator, with the exception of a scenario analysis. The company justified this citing a lack of data.
Subgroups to be considered	If the evidence allows, a scenario analysis including allogeneic stem cell transplant as a subsequent treatment after pembrolizumab or its comparators will be considered. This should reflect the	No response.	Mostly in line with the scope. Allogeneic stem cell transplant was incorporated into the company's base-case model as a subsequent treatment, reflecting the proportion of people who proceed to it

	Final scope issued by NICE	Decision problem addressed in the company submission and rationale	ERG comments
	proportion of people who proceed to allogeneic stem cell transplant after each treatment, as well as the costs and quality-adjusted life year benefits of the procedure.		after each treatment, as well as costs and quality-adjusted life year benefits of the procedure. A model without this option was not provided.
Special considerations including issues related to equity or equality	Not applicable.	Not applicable.	

Source: Table 1, Section B.1.1 of the CS.[1] and NICE FAD for TA462[6]

alloSCT = Allogeneic Stem Cell Transplant; ERG = Evidence Review Group; MSD = Merck Sharp and Dohme Ltd; NHS = National Health Service; OS = Overall Survival; TA = Technology Assessment.

3.1 Population

The population of this appraisal is in line with the scope. The main trial in the CS covers both cohorts of interest (people with relapsed or refractory cHL who have received autologous stem cell transplant and brentuximab and those who have received BV when autoSCT is not a treatment option).

However, only four out of 69 patients in cohort 1 and 10 out of 81 patients in cohort 2 of the intervention study (KEYNOTE-087) were from the UK. None of the patients in the comparator study (Cheah et al. 2016[7]) were from the UK.

The comparator study in this appraisal was also used in a previous appraisal (TA462[6]). NICE concluded in TA462 that "the comparator data may not fully represent UK clinical practice". If that is the case, then the results of the Matching-Adjusted Indirect Comparison (MAIC) in this appraisal are also not representative for UK clinical practice. This is because, the MAIC aims to generate the effect of pembrolizumab that would be observed in the Cheah trial population.[8]

3.2 Intervention

The intervention (pembrolizumab) is in line with the scope. Regulatory approval by the EMA for the indication considered within this submission was granted on the 2 May 2017. This stated that pembrolizumab as monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and BV, or who are transplant-ineligible and have failed BV.

Pembrolizumab is a highly selective humanised monoclonal antibody against programmed death-1 (PD-1) receptor, which exerts dual ligand blockade of the PD-1 pathway, including PD-L1 and PD-L2, on antigen presenting tumour cells. By inhibiting the PD-1 receptor from binding to its ligands, pembrolizumab activates tumour-specific cytotoxic T lymphocytes in the tumour microenvironment and reactivates anti-tumour immunity.

The route of administration for pembrolizumab is IV infusion, over a 30-minute period. The anticipated licensed dosing regimen for patients with relapsed or refractory classical Hodgkin Lymphoma who have failed autoSCT and Brentuximab Vedotin (BV), or who are transplant ineligible and have failed BV is 200 mg every three weeks. Treatment with pembrolizumab continues until disease progression or unacceptable toxicity, whichever occurs first. The list price of pembrolizumab is £2,630 per 100 mg vial ([redacted]).[1]

3.3 Comparators

The description of the comparators in the NICE scope is as follows:

- Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin
- Best supportive care.

The company provides one study for the comparator. This is a retrospective USA database study published in 2016 by Cheah and colleagues in which patients received the following types of therapy: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or alloSCT, or other treatment.[7] This is referred to in the CS as standard of care (SoC).

This comparator broadly matches the comparator described in the NICE scope: "Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin." However the ERG notes that there is some uncertainty about how well the Cheah study[7] which drew on data from patients treated

in the USA and which provides the base case comparator data, reflects the experience of patients treated in the UK. There is a lack of detail in the Cheah and colleagues' publication about the precise composition of the treatment regimens received by patients who had received ASCT and brentuximab vedotin. Many patients for whom outcome evaluations were available (28/67; 42%) were enrolled onto trial protocols and received what is described as 'Investigational agent', but there is no further detail about which therapies may have been classified under this heading. To find out whether treatments such as pembrolizumab were included among the 'Investigational agent' treatments, the ERG asked the company to clarify this (Clarification letter, Question A13).[9] The company replied that investigational agents did not included pembrolizumab although 'a couple of patients in the study received a PD-1 inhibitor.'[10] The company provided response rates results data excluding investigational agents.[10] The next most common regimens received by patients in the Cheah and colleagues study were gemcitabine-based (12/67; 18%) or bendamustine-based (11/67; 16%).

As reported in the ERG report for TA462, "gemcitabine regimens such as GDP (gemcitabine, dexamethasone, cisplatin) are commonly used in this patient population in the UK but platinum-containing regimens such as ESHAP (etoposide, methylprednisolone, cytarabine, cisplatin) and DHAP (dexamethasone, cytarabine, cisplatin) are also in common use. In the Cheah study 12/67 (18%) of patients with outcome evaluations received gemcitabine and just 4/67 (6%) of patients received platinum-based regimens."[11]

However, despite the uncertainty about how closely the experience of patients from the USA may match that of patients in the UK, the ERG is not aware of a more appropriate source of data for the comparator population.

Evidence for the clinical efficacy of best supportive care (BSC) is not presented within the clinical effectiveness section of the CS and in section 5 (cost effectiveness) of the CS, the company states that "Based on BCSH guidelines and clinician opinion, it is believed that use of BSC is minimal at this stage in the treatment pathway, as eligible patients are likely to receive therapy where feasible. As such, BSC has been applied within the model as a subsequent therapy in the base case analysis, with the composition derived from a recent NHL NICE Technology Appraisal (TA306).[12]" (CS, page 148)

In the economic model a scenario analysis was provided assessing the impact of BSC as a comparator. Due to lack of data informing the efficacy of BSC, in this scenario analysis, efficacy of BSC was assumed equivalent to that of Standard of Care (SoC).

3.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall survival (OS)
- progression free survival (PFS)
- response rate (RR)
- adverse effects of treatment (AE)
- health-related quality of life (HRQoL)

These outcomes are reported in the CS. However, survival data (OS and PFS) are immature, and only two outcomes have been included in the indirect comparison: PFS and ORR.

3.5 Other relevant factors

According to the company a commercial access agreement (CAA) is in place with the Department of Health [redacted]) of the list price of pembrolizumab (CS, Table 4, page 31).

In addition, the company states that "no additional tests or investigations are required further to the usual tests undertaken in current clinical practice. No diagnostic test is required to identify the population for whom pembrolizumab is indicated and no particular administration for the technology is required." (CS, section 2.4, page 32).

Regarding the innovative nature of pembrolizumab, the company states that the US Food and Drug Administration (FDA) granted pembrolizumab Orphan Drug Designation for the treatment of HL, and Breakthrough Therapy Designation; in addition, the application received priority review status and accelerated approval.[13]

No equity or equality issues were specified in the final scope or identified by the company. The ERG is not aware of any issues related to equity or equality in the use of pembrolizumab in patients with relapsed or refractory classical Hodgkin lymphoma following (autoSCT and) brentuximab vedotin.

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company updated an existing systematic review to identify evidence on the use of pembrolizumab in classical Hodgkin Lymphoma. The review was designed to identify both clinical trials and observational studies and to inform both direct and indirect comparisons between the interventions relevant to the NICE scope. This section critiques the methods of the review including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

4.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the company submission. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[14] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[15] The ERG has presented only the major limitations of each search strategy in the report.

The company submission stated that systematic review searches were undertaken in October and December 2016, with an update in June 2017. Search strategies were reported in Appendix 2 of the CS for the following databases: Embase, MEDLINE, MEDLINE in-Process, Cochrane's CENTRAL database.

Additional searches of the following conference proceedings using the Northern Light database were reported: American Society of Clinical Oncology (ASCO) (2015-2016) and the American Society of Haematology (ASH) (2014-2016), as well as a manual search of the WHO International Clinical Trials Registry (WHO ICTRP) to identify ongoing trials.

Searches utilised study design filters based on the Scottish Intercollegiate Guidelines Network (SIGN) filters for RCTs and Observational Studies.[16]

ERG comment:

- The database searches were clearly structured and documented.
- The ERG was concerned that limiting the clinical effectiveness searches reported in Appendix 2 to English language only may have introduced language bias. Current best practice states that *"Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication"*.[17]
- Best practice outlined in the Cochrane handbook states that *"Reference lists in other reviews, guidelines, included (and excluded) studies and other related articles should be searched for additional studies"*.[18] However the ERG found no mention of reference checking within the report. It was unclear whether this was due to a reporting error or an omission within the SR process.
- Free text terms were used to search for relapsed/refractory in the search strategies for observational studies on the Embase, MEDLINE and MEDLINE in-Process databases. This facet could have been extended to a broader range of search terms i.e. resist$ or persist$ or return$ or reocur$ or reoccur$ or recurren$ or recidiv$ or regenerat$ and the inclusion of MeSH/Emtree terms such as relapse/. Given the low number of hits retrieved due to the addition of a facet for brentuximab vedotin, the inclusion of the line for relapsed/refractory terms may have been overly restrictive. However, this is unlikely to have greatly affected the overall recall of results.

4.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for randomised controlled trials (RCTs) and non-RCTs is presented in Table 4.1.

The original review by the company was conducted in 2016 with an update in June 2017. The original inclusion criteria for the 2016 search strategy included a wider population and a longer list of interventions than the update. For the 2017 update search, the population was restricted to a population that was more in line with the final NICE scope (those who had disease progression during or after BV), and the interventions were defined in the same terms as those in the final scope: "Single or combination chemotherapy including drugs such as cisplatin, gemcitabine and vinblastine, and best supportive care". The updated review was designed to identify studies to inform both direct and indirect comparisons between interventions relevant to the NICE scope. The CS stated that two reviewers were involved in study selection with a third consulted in case of discrepancies.

Table 4.1: Eligibility criteria used in search strategy

	Description	
	Original SLR (Oct.19 and Dec. 2, 2016)	**Updated SLR (June 15 2017)**
Population	Adult cHL patients who either: failed to achieve a response to any line of therapy (refractory patients) or who have relapsed after ≥ 3 prior lines of therapy	Additional criteria added to restrict patients to those with disease progression during or after treatment with BV
Interventions	The following targeted drugs alone or as combinations with systemic chemotherapies: • Pembrolizumab • Brentuximab vedotin • Everolimus • Lenalidomide • Lucatumumab • Nivolumab • Ofatumumab • Panobinostat • Rituximab • Vorinostat The following systemic chemotherapies alone or in combinations: • Adriamycin • Bendamustine • Bleomycin • Carmustine • Cisplatin • Cyclophosphamide • Cytarabine • Dacarbazine • Etoposide • Gemcitabine • Ifosfamide • Mecholrethamine (Nitrogen mustard) • Melphalan • Mitoxantrone • Oxaliplatin • Procarbazine • Vinblastine • Vincristine • Vinorelbine Other treatments in combination with chemotherapies: • Dexamethasone • Methylprednisolone • Prednisone	Additional criteria were added to reflect only those interventions considered relevant to UK clinical practice: • Single or combination chemotherapy including drugs such as: o Cisplatin o Gemcitabine o Vinblastine • Best supportive care
Comparators	Any	Any

	Description	
	Original SLR (Oct.19 and Dec. 2, 2016)	**Updated SLR (June 15 2017)**
Outcomes	• Overall survival • Progression-free survival • Objective response • Complete response • Partial response • Treatment discontinuation due to AEs • Serious (grade 3 and above) AEs (not used for study selection)	No change
Study design	• Randomised controlled trials • Non-randomised controlled trials • Single arm trials • Retrospective and prospective controlled observational studies • Single group observational studies	No change
Source: CS, Table 6, page 43[1] AEs = adverse events; BV = brentuximab vedotin; cHL = classical Hodgkin Lymphoma		

ERG comment:

- The restriction of the updated systematic review to a population more in line with the NICE scope was appropriate.
- The original criteria for the 2016 systematic review included a longer list of interventions. For the 2017 update the interventions were defined in the same terms as those in the final scope: "Single or combination chemotherapy including drugs such as cisplatin, gemcitabine and vinblastine, and best supportive care". However, the phrase 'drugs such as' is rather vague and studies were excluded because the treatment 'did not reflect UK practice'; therefore, we asked the company to specify which interventions were included (Clarification letter, Question A2). The company responded by repeating the NICE scope: "MSD included comparators listed in the NICE final scope (March 2017) considered to represent UK clinical practice. This comprises: single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin and best supportive care."[10] The impact of this is discussed in Section 4.2.1 of this report.

4.1.3 Critique of data extraction

The CS stated that two investigators extracted data independently from the included studies. Any discrepancies between data extractions were resolved by involving a third reviewer and coming to a consensus.

ERG comment:

- Data extraction appears to have been conducted appropriately.

4.1.4 Quality assessment

It appears that two investigators assessed study quality independently. Any discrepancies between assessments were resolved by involving a third reviewer and coming to a consensus. The tool used was the Newcastle-Ottawa Scale covering issues related to selection bias and assessment of outcomes.

ERG comment:

- Study quality was assessed appropriately. Results of the quality assessment by the company and the ERG of the KEYNOTE-087 trial are outlined in Section 4.2.2.4 of this report. The limitations of single-arm studies are also outlined in Section 4.2 of this report.

4.1.5 Evidence synthesis

No trials directly comparing pembrolizumab with a comparator of interest were identified therefore a meta-analysis of the direct evidence could not be performed. The company described the results of the KEYNOTE-087 single arm trial. A retrospective observational study (Cheah et al. 2016[7]) was identified from searches of the literature and used as a comparator in naïve comparisons and matched adjusted indirect comparison (MAIC). This analysis and its results are described more fully in Section 4.4 of this report.

ERG comment:

- The ERG agrees that no direct meta-analysis was possible given that only one single arm study of pembrolizumab was identified (KEYNOTE-087).

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Overview of the evidence in the submission

No relevant randomised controlled trials (RCTs) of pembrolizumab were identified by the company. The CS was based on one ongoing single arm phase II trial (KEYNOTE-087). KEYNOTE-087 will be discussed in detail in this section of the report.

The submission briefly mentions a phase 1b trial of pembrolizumab (KEYNOTE-013). However, this study did not correspond to the EMA licensing for the dosing of pembrolizumab and was used as supporting evidence for safety only, therefore will only be briefly mentioned in Section 4.2.3 of this report. The company also provides details of a clinician survey to support understanding of UK clinical practice. This survey is also briefly discussed in Section 4.2.3 of this report.

Two further trials were mentioned as being ongoing: KEYNOTE-204 and NCT03077828. These studies are discussed in Section 4.2.4 of this report.

ERG comment:

- The ERG was provided with a list of excluded studies. The company stated in response to clarification that 'MSD included comparators listed in the NICE final scope (March 2017) considered to represent UK clinical practice. This comprises: single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin and best supportive care.'[10] The ERG checked the list of studies excluded based on intervention and concluded that no comparative studies had been inappropriately excluded.
- A small number of studies of nivolumab were identified and excluded. Although not listed as a comparator in the NICE scope and not referenced in the CS, nivolumab has recently received approval from NICE for this condition. It is recommended 'as an option for treating relapsed or refractory classical Hodgkin lymphoma in adults after autologous stem cell transplant and treatment with brentuximab vedotin.'[6] Nivolumab is, however, not recommended for one of the populations in this appraisal (those who have received BV but who have not received an autoSCT).

- Bendamustine has also been investigated in small observational studies which were excluded from the review. The ERG did not believe these studies were suitable comparator studies and agreed that they should be excluded.[19, 20]
- The ERG notes that the evidence for pembrolizumab is based on one single arm, ongoing trial.

4.2.2 KEYNOTE-087

4.2.2.1 Methodology of KEYNOTE-087

KEYNOTE-087 is a phase II, multicentre, single arm trial of pembrolizumab in adult patients with RRcHL. See Table 4.2.

Table 4.2: Methodology of the KEYNOTE-087 trial

PICOS	Details		
Population	Patients ≥ 18 with relapsed[a] or refractory[b] de novo classical Hodgkin Lymphoma		
	Measurable disease defined as ≥ 1 lesion accurately measured in ≥ 2 dimensions with spiral CT. Minimum measurement > 15mm in the longest diameter or > 10mm in the short axis.		
	ECOG Performance Scale 0 or 1		
	Cohort 1 (n = 69)	*Cohort 2 (n = 81)*	*Cohort 3[c]*
	Have failed to achieve a response or have progressed after autoSCT.	Were unable to achieve a CR or a PR to salvage chemotherapy and did not receive autoSCT.	Have failed to achieve a response or have progressed after autoSCT
	Patients must have relapsed after treatment with or failed to respond to BV post autoSCT.	Patients must have relapsed after treatment with or failed to respond to BV	Patients have not received BV post autoSCT.
Setting	Three study sites in the UK[d], 23 elsewhere in Europe, 11 in the USA, seven in Japan, four in Israel, two in Australia and one in Canada		
Intervention	200mg pembrolizumab as 30 min IV infusion every three weeks in the outpatient setting		
Outcomes[e]	*Primary*	*Secondary*	
	Overall response rate (ORR) defined as the proportion of patients who have complete remission (CR) or partial remission (PR) using IWG response criteria assessed by CT / PET at any time during the study as determined by blinded, independent central review (BICR).	ORR using IWG criteria at any time during the study as determined by investigator	
	Safety and tolerability (including adverse events and serious adverse events)	ORR using 5-point scale according to the Lugano classification as determined by BICR	
		Progression-free survival (PFS) and duration of response (DOR) by BICR and by investigator according to the IWG criteria	

PICOS	Details	
		Overall survival
Study design	Phase II single arm, open label trial	

Source: Section 4.3.1 of the CS
Footnote: a) Disease progression after most recent therapy; b) failure to achieve CR or PR to most recent therapy; c) Not relevant to this appraisal; d) 14 patients were from the UK (Cohort 1, n = 4; Cohort 2, n = 10); e) The trial also listed exploratory outcomes including an assessment of ORR, CRR, PFS and DOR for patients who continue treatment with pembrolizumab beyond documented progression and an assessment of health-related quality of life.
autoSCT = Autologous Stem Cell Transplant; BICR = Blinded independent central radiologists; BV = Brentuximab Vedotin; CR = Complete response; ECOG = Eastern Cooperative Oncology Group; IV = intravenous; IWG = International Working Group; ORR = Objective Response Rate; PR = Partial response

The trial has three cohorts. Cohort 1 includes patients who have failed to achieve a response or who have progressed after autoSCT and have relapsed after treatment with or have failed to respond to BV post autoSCT. Cohort 2 comprises patients, most of whom were unable to achieve CR or PR to salvage chemotherapy and did not receive autoSCT, and have relapsed after treatment with or failed to respond to BV. Cohort 3 includes patients who have failed to respond to, or have progressed after autoSCT and have not received BV post autoSCT (see Table 4.2). Cohort 3 is not relevant to this submission so effectiveness results are not presented for this cohort in this report.

A number of patient exclusion criteria were outlined in the CS. Most relevant are that patients who had undergone prior alloSCT within the last five years were excluded. Patients who had a transplant greater than five years ago were eligible provided there were no symptoms of graft vs. host disease. A further exclusion criterion was that patients should not have received prior therapy with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD-137 or anticytotoxic T-lymphocyte associated antigen-4 antibody (including ipilimumab or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways).

KEYNOTE-087 has a total of 210 participants of whom 150 are relevant to this submission (69 cohort 1, 81 cohort 2). It is a multinational trial including three sites in the UK. The CS further detailed that four patients in cohort 1 and 10 in cohort 2 were from the UK.

As a single arm, open label trial, treatment was known to both investigators and patients. Patients received 200 mg pembrolizumab as 30 min IV infusion every three weeks in the outpatient setting. Neither dose escalation nor dose reduction of pembrolizumab was permitted in the trial. Dose modification due to adverse events (both serious and non-serious) was permitted. All concomitant permitted medications received within 28 days before the first dose of trial treatment and 30 days after the last dose of trial treatment were recorded.

Disease response assessments were planned for every 12 weeks until documented disease progression, the start of a new anti-cancer treatment, withdrawal of consent, death or the end of the study, whichever occurred first. Bone marrow biopsies were collected to confirm complete remission (in patients who had bone marrow involvement) or if clinically indicated. Where a patient showed progressive disease pembrolizumab could be continued at the discretion of the principal investigator (PI) until the next disease response assessment provided their clinical condition was stable. Imaging should have occurred at any time where there was clinical suspicion of progression. Patients who experienced a complete or partial response or had stable disease were able to remain on treatment for up to two years (approximately 37 administrations) or until unacceptable toxicity or progression. Patients who attained

a complete response could stop pembrolizumab after a minimum of 24 weeks of treatment with at least two doses since initial confirmation of CR. Patients who later experienced disease progression could be retreated with pembrolizumab at the same dose and schedule as at the time of initial discontinuation if no cancer treatment had been administered since the last dose of pembrolizumab. [redacted]

The primary outcome was best overall response rate (best ORR or BOR); ORR is defined as the proportion of patients who have complete remission (CR) or partial remission (PR) using International Working Group (IWG) response criteria assessed by CT/PET at any time during the study as determined by blinded, independent central review (BICR), and the Best Overall Response (BOR) is the best response recorded from the start of the study treatment until the disease progression/recurrence. Progression free survival and overall survival were assessed as secondary outcomes.

Health-related quality of life was also evaluated as an exploratory outcome. Assessments were made from baseline using the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life (QoL) Questionnaire C30 (QLQ-C30) and European Quality of Life Five Dimensions Questionnaire (EuroQoL, EQ-5D).

ERG comments:

- The most important methodological aspects to note are that although the trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative study which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention as the role of natural history and baseline characteristics is not taken into account. This is in contrast to a well-conducted randomised trial where bias is minimised and we can be confident that outcomes we observe are due to differences between the interventions evaluated.
- As a single-arm, open-label trial the intervention is known to participants, clinicians and assessors. Knowledge of interventions can lead to bias in delivering interventions and reporting outcomes.
- The trial gives a maximum two years of outcome data on patients. Outcomes are relevant but the primary outcome is objective response rate rather than the longer-term outcomes of PFS and OS which are evaluated as secondary outcomes.
- Cohorts 1 and 2 of the trial are relevant to the decision problem in the NICE scope. Inclusion and exclusion criteria in terms of population appear to be appropriate.
- Although the trial is multinational it only has 150 relevant participants so the evidence base for this appraisal is small. However, the population matching the scope of this appraisal is in itself small so conducting a larger trial would be challenging.
- A small number of patients were from the UK (14) so the trial may not totally reflect the UK population and setting. However, once again the population from which to draw participants is small.
- In clinical practice, for those who are suitable, pembrolizumab represents a bridge to alloSCT, a potentially curative treatment. However, the company submission stated that 'KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2.'[1] The company further stated that 'the use of stem cell transplant would have been at the discretion of the treating physician on a per patient basis.'[10] The company clarified that [redacted] of the KEYNOTE-087 population had an alloSCT following pembrolizumab ([redacted] from cohort 1 and [redacted] from cohort 2).[10] However it was noted

that [redacted] patients in cohort 1 and 2 in the UK were transplanted with allogeneic stem cells respectively.[1]

4.2.2.2 Statistical analysis of KEYNOTE-087

The primary hypothesis of this study was that i.v. administration of pembrolizumab would reach an ORR of greater than 20% in each of the three cohorts using IWG criteria by independent review committee. The selection of 20% as a control rate was based partly on the published literature prior to the approval of BV and downgraded to take account that this patient group have failed treatment with BV. Enrolment of 60 patients per cohort was required to have 93% power at a one-sided 2.5% α level to detect a 40% or higher ORR for pembrolizumab compared to a fixed control rate of 20% using an exact binomial test. Each cohort was analysed separately and also as a pooled group. However, only the results for cohorts 1 and 2 are presented in this report. The company stated that '*No additional multiplicity adjustment was required because each cohort was evaluated independently*.'[1]

Final analysis was to be conducted for each cohort when the last participant reached the Week 12 response assessment or discontinued study therapy. Results are presented as a percentage with the exact 95% two-sided CI (Clopper-Pearson method). An exact binomial test was used to obtain a one-sided p-value for comparing the observed ORR to the control value of 20% (null hypothesis $p \leq 0.20$ vs. alternative hypothesis $p > 0.20$) for each cohort. The analysis of the primary endpoints used the All Subjects as Treated (ASaT) population (those who had received at least one dose of medication). Supportive analyses were also conducted using the full analysis set (FAS) which was all patients who received at least one dose of study medication, had a baseline disease assessment and either a post-baseline disease assessment or who discontinued the trial due to progressive disease/drug related AE.

Time to event outcomes (response duration, PFS and OS) were summarised by the median time to event with 95% CI using the Kaplan-Meier method. The percentage surviving at different time points (3, 6, 9 and 12 months for PFS) and for OS (6, 9, 12 and 15 months) were also obtained using the Kaplan-Meier method,

The ASaT population was also used for the analysis of safety. Additionally, at least one laboratory or vital sign measurement obtained subsequent to at least one dose of trial treatment was required for inclusion in the analysis of each specific safety parameter.

Table 4.3 gives an overview of the main analyses undertaken in KEYNOTE-087.

Table 4.3: Efficacy analysis of primary and secondary endpoints in KEYNOTE-087

Endpoint / Variable	Statistical method	Analysis population	Missing data approach
Primary outcome			
Overall response rate IWG criteria (2007) • Central review	Exact test of binomial parameter; 2-sided 95% exact CI	ASaT / FAS	Participants with missing data are considered non-responders
Secondary outcomes			
Overall response rate IWG criteria (2007) • Study site Lugano criteria (2014) • Central review	Point estimate; 2-sided 95% exact CI	ASaT / FAS	Participants with missing data are considered non-responders

Endpoint / Variable	Statistical method	Analysis population	Missing data approach
Complete remission rate IWG criteria (2007) • Central review • Study site Lugano criteria (2014) • Central review	Point estimate; 2-sided 95% exact CI	ASaT / FAS	Participants with missing data are considered non-responders
Progression-free survival IWG criteria (2007) • Central review • Study site	Summary statistics using Kaplan-Meier method	ASaT / FAS	Censored at last assessment
Duration of response IWG criteria (2007) • Central review • Study site	Summary statistics using Kaplan-Meier method	All responders	Non-responders are excluded in analysis
Overall survival	Summary statistics using Kaplan-Meier method	ASaT / FAS	Censored at last assessment
Source: CS, Table 9, page 62. ASaT = All Subjects as Treated; CI = Confidence Interval; FAS = Full Analysis Set; IWG = International Working Group.			

ERG comment:

- The ERG has no concerns about the design or statistical analyses of the KEYNOTE-087 trial. It was a non-comparative single-arm trial and the sample size calculation and analysis methods are appropriate.

4.2.2.3 Participants in the KEYNOTE-087 trial

A total of 210 patients were enrolled in the KEYNOTE-087 trial of which 69 formed cohort 1 (who had failed to achieve a response or had progressed after autoSCT) and 81 formed cohort 2. In response to clarification the company stated that the majority of cohort 2 did not qualify for an autoSCT [redacted]. This information was not routinely gathered but the company stated that a small number of participants did not receive autoSCT for a variety of reasons [redacted].[10]

All patients in both cohorts had relapsed or refractory disease and all had used BV as per the inclusion criteria for the trial. Cohort 3 (the remaining 60 patients) are not relevant to this appraisal. Patient characteristics for cohorts 1 and 2 are reported in Table 4.4.

Table 4.4: Patient characteristics in the KEYNOTE-087 trial

	Cohort 1 (n = 69) n (%)	Cohort 2 (n = 81) n (%)
Gender		
Male	36 (52.2)	43 (53.1)
Female	33 (47.8)	38 (46.9)
Age Years		
<65	69 (100)	66 (81.5)

	Cohort 1 (n = 69) **n (%)**	**Cohort 2 (n = 81)** **n (%)**
≥ 65	0	15 (18.5)
Mean (SD)	37.0 (10.9)	42.3 (17.4)
Median	34.0	40
Range	19 to 64	20 to 76
Race		
American Indian or Alaska native	0	1 (1.2)
Asian	7 (10.1)	4 (4.9)
Black or African American	2 (2.9)	2 (2.5)
Missing	1 (1.4)	1 (1.2)
Multi-racial	2 (2.9)	0
White	57 (82.6)	73 (90.1)
Disease subtype		
CHL – nodular sclerosis	55 (79.7)	65 (80.2)
CHL – mixed cellularity	9 (13.0)	10 (12.3)
CHL – lymphocyte rich	4 (5.8)	1 (1.2)
CHL – lymphocyte depleted	0	4 (4.9)
Missing	1 (1.4)	1 (1.2)
ECOG performance status		
0	29 (42.0)	44 (54.3)
1	39 (56.5)	37 (45.7)
2	1 (1.4)	0
Prior lines of therapy		
≥ 3	68 (98.6)	78 (96.3)
< 3	1 (1.4)	3 (3.7)
Mean (SD)	4.5 (1.7)	4.0 (1.7)
Median	4.0	4.0
Range	2 to 12	1 to 11
Time of relapse since SCT failure (months)		
≥ 12	37 (53.6)	0
< 12	32 (46.4)	0
Mean (SD)	60.2 (39.6)	NA
Median	12.6	NA
Range	2.5 to 247.9	NA
Prior radiation		
Yes	31 (44.9)	21 (25.9)
No	38 (55.1)	60 (74.1)
Bulky lymphadenopathy		
Yes	5 (7.2)	12 (14.8)

	Cohort 1 (n = 69) n (%)	Cohort 2 (n = 81) n (%)
No	64 (92.8)	69 (85.2)
Baseline B symptoms		
Yes	22 (31.9)	26 (32.1)
No	47 (68.1)	55 (67.9)
Baseline bone marrow involvement		
Yes	3 (4.3)	5 (6.2)
No	66 (95.7)	75 (92.6)
Missing	0	1 (1.2)
Source: Table 11 of the CS (abbreviated) CHL = Classical Hodgkin Lymphoma; ECOG = Eastern Cooperative Oncology Group; SCT = stem cell transplant; SD = standard deviation		

Both cohorts had slightly more male than female participants (52.2% in cohort 1 and 53.1% in cohort 2). Most participants across the cohorts were white (82.6% in cohort 1 and 90.1% in cohort 2). Although both cohorts had a wide age range (cohort 1: 19 to 64, cohort 2: 20 to 76) all of the participants in cohort 1 and 85.1% of the participants in cohort 2 were under 65 years of age. The most common disease subtype was cHL – nodular sclerosis (cohort 1: 79.7%, cohort 2: 80.2%). All patients except one in cohort 1 had an ECOG score of 0 or 1. Approximately a third of patients across the cohorts had B symptoms (cohort 1: 31.9%, cohort 2: 32.1%). A small number had bone marrow involvement (cohort 1: 4.3%, cohort 2: 6.2%).

The mean time of relapse since autoSCT was 60.2 months. Both cohorts were heavily pre-treated. In cohort 1 [redacted] of patients had received at least three lines of therapy (range 2 to 12). In cohort 2 [redacted] had received at least three lines of therapy (range 1 to 11). Participants in cohort 1 had a median of [redacted] days since the last dose of BV (Range [redacted] days) whilst cohort 2 had a median of [redacted] days since their last dose (range [redacted] days).

ERG comment:

- There is a peak in incidence of cHL in older males and females (75 to 79 for men and 70 to 74 for women) but no patients in cohort 1 are 65 or over 65. In cohort 2 18.5% of patients are 65 or over. At least in cohort 1 older patients are underrepresented in KEYNOTE-087.
- Advisers to the company stated that typically patients within the UK would have received between three and four prior lines of therapy, including BV, before starting treatment with a PD-L1. In the trial only [redacted] patient in cohort 1[redacted] and [redacted] in cohort 2 ([redacted] had received fewer than three therapies so in this respect is applicable to UK practice. However, it should be noted that [redacted] of cohort 1 ([redacted]) had received five or more therapies. In cohort 2 [redacted] ([redacted]) had received five or more therapies. The population of the trial could, therefore, be more heavily treated than in UK practice.

4.2.2.4 Quality assessment of the KEYNOTE-087 trial

The results of the company's and the ERG's assessment of KEYNOTE-087 are shown in Table 4.5. It should be noted that not all of the questions in the tool are applicable to a single-arm study.

Table 4.5: Quality assessment of the KEYNOTE-087 trial

	CS evaluation	ERG evaluation	ERG comment
Selection bias			
Representativeness of cohort	*	*	Representative of the cHL population but may not be representative of the UK population
Selection of non-exposed cohort	NA	NA	
Ascertainment of exposure	*	*	Assessment was made of number of patients who received at least one dose of treatment
Outcome of interest	*	*	Presence of the outcome of interest was assessed before exposure to the intervention.
Comparability of cohorts	NA	NA	
Outcome bias			
Outcome assessment	*	*	Outcomes were evaluated by an independent review committee (IRC).
Adequate duration of follow-up			Median follow up time was 15.9 months. This was adequate for ORR but not for PFS and OS.
Adequate follow-up of cohort	*	*	Explanations were provided regarding missing data or loss to follow up.
Source: CS, Table 12, page 68 CS = company submission; ERG = evidence review group; NA = non-applicable			

ERG comments:

- The most important methodological aspect to note is that although the trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative trial which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention. The role of natural history and baseline characteristics is not taken into account.
- The study had an adequate follow-up (median 15.9 months) for the main outcome evaluated (ORR defined as the proportion of patients who have complete remission (CR) or partial remission (PR)). However median progression free survival was immature and [redacted]

4.2.2.5 Main efficacy results of the KEYNOTE-087 trial

At the 21 March 2017 data cut off [redacted] of cohort 1 patients and [redacted] of cohort 2 patients remained on treatment. Table 4.6 gives the current status of the patients in the KEYNOTE-087 trial.

Table 4.6: Patient status in the KEYNOTE-087 trial

Patient Status	Cohort 1 (n = 69) n (%)	Cohort 2 (n = 81) n (%)
Started	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]
Bone marrow transplant	[redacted]	[redacted]
Clinical progression	[redacted]	[redacted]
Complete response	[redacted]	[redacted]
Death	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]
Physicians Decision	[redacted]	[redacted]
Pregnancy	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]
Withdrawal by subject	[redacted]	[redacted]
Treatment on-going	[redacted]	[redacted]
Source: CS, Table 10, page 64		

Table 4.7 presents the main efficacy data for the trial. The company confirmed that these were the latest efficacy data available.

Table 4.7: Summary efficacy results of the KEYNOTE-087 trial

Outcome[a]	Results[b]	
	Cohort 1 N = 69	**Cohort 2 N = 81**
Overall survival		
Death n (%)	[redacted]	[redacted]
Median (95% CI) months[c]	[redacted]	[redacted]
OS at 12 months % (95% CI)[c]	[redacted]	[redacted]
Progression-free survival		
Median (95% CI) months[c]	16.7 (11.2 to NR)	11.1 (7.6 to 13.7)
PFS at 12 months % (95% CI) [c]	[redacted]	[redacted]
Response rates		
ORR n (%)[d]	52 (75.4)	54 (66.7)
CR n (%)	19 (27.5)	20 (24.7)
PR n (%)	33 (47.8)	34 (42)
SD n (%)	[redacted]	[redacted]
PD n (%)	[redacted]	[redacted]
No assessment n (%)	[redacted]	[redacted]
Time to response Median (range) months[c]	[redacted]	[redacted]
Duration of response Median (range) months[c]	[redacted]	[redacted]

Outcome[a]	Results[b]	
	Cohort 1 N = 69	Cohort 2 N = 81
Source: CS, Section 4.7, tables 14 and 15 Footnote: a) as per the NICE scope; b) 21 March 2017 unless otherwise stated. Median follow-up 15.9 months (range 1.0 to 20.9 months); c) From product-limit (Kaplan-Meier) method for censored data; d) assessed by BICR using IWG criteria CI = confidence interval; ORR = overall response rate; OS = overall survival; PD = progressive disease; PFS = progression free survival; PR = partial response; SD = stable disease		

Overall response rate (the primary outcome as assessed by the independent committee using IWG criteria) was 75.4% in cohort 1 and 66.7% in cohort 2. In Cohort 1 27.5% of patients had a complete response and in cohort 2 this figure was 24.7%. Median time to response was [redacted] and [redacted] respectively. However median duration of response was [redacted] in cohort 1 and was [redacted] months in cohort 2.

Median PFS in cohort 1 as assessed by independent committee was 16.7 months (11.2 to NR). In cohort 2 it was 11.1 months (7.6 to 13.7).

Median OS was [redacted]. At 12 months survival was [redacted] in cohort 1 and [redacted] in cohort 2.

ERG comment:

- As stated above, the trial was long enough to show the benefit of pembrolizumab on overall response rates including both CR and PR. However, PFS and OS data are not fully mature.

4.2.2.6 Post-hoc analyses of the KEYNOTE-087 trial

The company conducted post-hoc analyses of response to inform the naïve indirect treatment comparison and the Matched Adjusted Indirect Comparison (MAIC) (discussed in Section 4.3). The main difference between this post-hoc analysis of response and the primary analysis of response, referred to as 'best' response rate, is that response was determined at a single time point for each patient i.e. 12 weeks as opposed to any time point up to the point of progression. Data in the form of the proportion who respond by a specific time point was required in order to apportion patients that were progression-free into CR, PR or SD in the cost-effectiveness model. The table below shows the response rates at week 12.

Table 4.8: 12-week response results in the KEYNOTE-087 trial

Outcome[a]	Results[b]	
	Cohort 1 N = 69	Cohort 2 N = 81
Response rates		
ORR n (%)[d]	[redacted]	[redacted]
CR n (%)	[redacted]	[redacted]
PR n (%)	[redacted]	[redacted]
Stable disease (SD) n (%)	[redacted]	[redacted]
Progressive disease (PD) n (%)	[redacted]	[redacted]

Outcome[a]	Results[b]	
	Cohort 1 N = 69	Cohort 2 N = 81
No assessment n (%)	[redacted]	[redacted]

Source: Section 4.8 of the CS
Footnote: a) as per the NICE scope; b) 21 March 2017 unless otherwise stated. Median follow-up 15.9 months (range 1.0 to 20.9 months); c) From product-limit (Kaplan-Meier) method for censored data; d) assessed by BICR using IWG criteria
ORR = overall response rate; CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease

ERG comment:

- The ERG noted that overall response rates were lower at 12 weeks than over the course of the trial ([redacted]

4.2.2.7 Safety results of the KEYNOTE-087 trial

Safety results from KEYNOTE-087 are presented from the data cut-off of 25 September 2016. The company confirmed that these were the most recent safety data available. All enrolled patients received at least one dose of study treatment. The median time on treatment was [redacted] days for cohort 1 and [redacted] days for cohort 2. Cohort 1 had a mean number of [redacted] administrations whilst cohort 2 had a mean of [redacted].

Table 4.9 gives an overview of the numbers and percentages of patients who had an AE up to 30 days and serious AEs up to 90 days after the last dose of study medication.

Table 4.9: Overview of adverse events in the KEYNOTE-087 trial

Database Cut-off Date: 25 Sep 2016	Cohort 1 (n = 69)	Cohort 2 (n = 81)
Patients in population	**n (%)**	**n (%)**
with ≥ 1 AE	[redacted]	[redacted]
with no AE	[redacted]	[redacted]
with drug-related[a] adverse events	[redacted]	[redacted]
with toxicity grade 3-5[b] adverse events	[redacted]	[redacted]
with toxicity grade 3-5 drug-related adverse events	[redacted]	[redacted]
with non-serious adverse events	[redacted]	[redacted]
with serious adverse events	[redacted]	[redacted]
with serious drug-related adverse events	[redacted]	[redacted]
who died	[redacted]	[redacted]
who died due to a drug-related adverse event	[redacted]	[redacted]
Discontinued[d] due to an adverse event	[redacted]	[redacted]
discontinued due to a drug-related adverse event	[redacted]	[redacted]
discontinued due to a serious adverse event	[redacted]	[redacted]
discontinued due to a serious drug-related adverse event	[redacted]	[redacted]

Database Cut-off Date: 25 Sep 2016	Cohort 1 (n = 69)	Cohort 2 (n = 81)
Source: Table 39 of the CS AE = adverse event		

In cohort 1 [redacted] of patients had one or more adverse events. In cohort 2 [redacted] of patients had one or more adverse events. The company noted that most AEs were low grade ([redacted] and [redacted] grades 3 to 5 in cohort 1 and 2 respectively). In cohort 1 [redacted] of AEs were classed as serious and with approximately half of these drug-related. Similarly [redacted]% of cohort 2 experienced serious AEs of which approximately a quarter were serious. [redacted]. A small number of patients in both cohorts discontinued due to an adverse event ([redacted] in cohort 1, [redacted] in cohort 2).

The most common adverse events were pyrexia ([redacted]), cough ([redacted]) fatigue ([redacted]), diarrhoea ([redacted]) and vomiting ([redacted]).

[redacted] of AEs were deemed to be drug-related in cohort 1, and [redacted] of AEs were deemed to be drug-related in cohort 2. The most common drug-related AEs were hypothyroidism [redacted]) pyrexia ([redacted]), fatigue ([redacted]), rash ([redacted]) diarrhoea ([redacted]) and headache [redacted]).

Table 4.10 lists the drug-related serious adverse events by category in KEYNOTE-087. A SAE was defined as any AE that occurred during the use of pembrolizumab that resulted in: death, was life threatening, resulted in persistent or significant disability/incapacity, resulted in, or prolonged, an existing in-patient hospitalisation, was a congenital anomaly/birth defect, or was considered as another important medical event.

Table 4.10: Drug-related serious adverse events in the KEYNOTE-087 trial

Database Cut-off Date: 25 Sep 2016 up to 90 days after last dose	Cohort 1 (n = 69)	Cohort 2 (n = 81)
Patients in population[a]	**n (%)**	**n (%)**
One or more serious AE	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]
Myocarditis	[redacted]	[redacted]
Pericarditis	[redacted]	[redacted]
Immune system disorders	[redacted]	[redacted]
Cytokine release syndrome	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
Herpes simplex	[redacted]	[redacted]
Herpes zoster	[redacted]	[redacted]
Myelitis	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]
Infusion-related reaction	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]
Myositis	[redacted]	[redacted]

Database Cut-off Date: 25 Sep 2016 up to 90 days after last dose	Cohort 1 (n = 69)	Cohort 2 (n = 81)
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]
Pneumonitis	[redacted]	[redacted]
Source: Table 47 of the CS Footnote: a) Adverse events appear in this table if > 0 in Cohort 1 or 2. AE = adverse event		

ERG comment:

- Patients will need to be informed of the adverse events to make an informed decision on treatment. The percentage of drug-related events is high ([redacted] and [redacted] in cohorts 1 and 2 respectively). [redacted] of adverse events were grades 3 to 5 and in cohort 1 [redacted] were serious, in cohort 2 [redacted]. Given that nivolumab is now available for patients in cohort 1 it will be important to compare their adverse event profile.

4.2.3 Supporting evidence

Lafferty et al.[21]

The company used a study by Lafferty et al to provide data for the economic model of this appraisal.[21] Exact details of which data were used is discussed in the cost effectiveness section of this report. The study was not described in full in the CS and is only available as an abstract.

Briefly, the retrospective study evaluates 13 patients with HL who underwent alloSCT between 2008 and 2015. The population is described as being heavily pre-treated and all patients had received at least three lines of chemotherapy. Eight of 13 (62%) had undergone autoSCT prior to alloSCT. It was not stated if patients had received BV. Median age of the participants was 33. At the time of transplant 11 patients were in partial remission and two in complete remission. Donors were matched sibling (six patients), matched unrelated volunteer (six patients) and double cord stem cell transplant (one).

Median length of follow up in survivors was 424 days. At one year OS was 69% and PFS 54%. The four deaths in the first year were due to respiratory syncytial virus (RSV) infection, air embolism, acute graft versus host disease (GVHD). Relapse or progression post-transplant occurred in three patients (23%) all within one year. Acute GVHD developed in eight (62%) of patients and was grade II to IV in five (38%).

ERG comment:

- This study is relevant to the UK and was used in a previous appraisal (TA462).
- The study was available in abstract form only so could not be fully quality assessed. However, it is clear that as a source of data for the model, the study is very limited. It is a small, retrospective case series from a single centre in the UK. The care provided may not be typical of the general UK setting. The 13 patients may not fully reflect the characteristics of patients seen elsewhere in clinical practice. Older patients are not represented in this sample. It is unclear if all patients had received BV as per the population in this appraisal. There is no comparison of the outcome between those receiving alloSCT and those not. The role of natural history cannot be ascertained. The small numbers of patients mean that these results cannot be extrapolated to larger samples.

Clinician Survey

Due to the paucity of data available on standard of care for this patient group, the company commissioned a clinician survey to support understanding of UK clinical practice. Specifically, the survey aimed to determine UK clinical practice for the treatment of patients with RRcHL, to consider the treatment pathway and eligibility of patients with RRcHL following standard of care and to assess the validity of the Cheah et al.[7] study in relation to UK practice and the Lafferty et al.[21] study in relation to rates of alloSCT and outcomes after alloSCT in patients who have received standard of care in the relapsed/refractory setting.

The questionnaire was made available via a website and was completed by 16 clinicians (12 from England, one from Wales and three from Scotland). Respondents were either haematologists or haematological oncologists. The average number of patients seen by a clinician matching cohort 1 (failed autoSCT and BV) was four patients annually. The average number seen matching cohort 2 (ineligible for autoSCT and failed BV) was three patients annually. Three of 16 clinicians had experience of using PD-1s in cHL.

Clinicians noted that both cohorts of patients would receive standard of care for approximately 12 weeks. They considered that only a minority of patients on standard care would proceed to allogenic SCT when ineligible for autologous SCT and having failed BV (cohort 2 equivalent). This was estimated as 17% of those gaining a CR and 13% of those gaining a partial response. A CR to standard care was estimated as 12% of patients and a partial response to standard care was estimated from 19% of patients. This was in contrast to a cohort 1 equivalent where response to standard care was similar (14% CR, 21% PR) but 57% of those with a complete response would go on to alloSCT and 44% of those with a PR would receive alloSCT. However, it was noted that individual clinicians have small numbers of these patients so these percentages are estimates only.

The CS noted that clinicians surveyed were largely in agreement with the data in Cheah[7] and Lafferty[21] compared to clinical practice including a PFS of 3.5 months and OS of 25.2 months reported in Cheah. However, three clinicians suggested an OS of 12 months based on their practice. Furthermore, although the clinicians accepted the findings of this study, many reported no access to investigational agents which are included in Cheah et al.[7]

ERG comment:

- The company made efforts to apply the appraisal to a UK context with the use of the clinician survey. However due to the rarity of the disease at this stage clinicians did not see many patients per year (most commonly 3 or 4). Hence duration of treatment and percentage processing to alloSCT are estimates based on sparse data.

KEYNOTE-013

The company provided an overview of the safety results of a phase 1b (single arm) trial of pembrolizumab in patients with relapsed or refractory disease.[22] In this trial (KEYNOTE-013) patients had relapsed after, were considered ineligible for, or had refused autoSCT. All 31 patients had progressed on or after treatment with BV. The CS stated that the dosing of pembrolizumab does not support the EMA recommendation so it was excluded from the decision problem. In this trial, pembrolizumab was administered intravenously at a dose of 10 mg/kg every two weeks.

The company stated that AEs of any grade and attribution were reported in 30 0f 31 patients (97%). Overall, 68% of patients experienced one or more AEs that were deemed related to treatment. There were no grade 4 treatment-related AEs and no deaths related to study treatment. The publication

associated with this trial also provided further details on efficacy. The CR rate was 16% (90% CI, 7% to 31%). In addition, 48% of patients achieved a partial remission, for an overall response rate of 65% (90% CI, 48% to 79%). (70% of the responses lasted longer than 24 weeks (range, 0.14+ to 74+ weeks), with a median follow-up of 17 months. The progression-free survival rate was 69% at 24 weeks and 46% at 52 weeks.

ERG comment:

- As the dosing regimen of KEYNOTE-013 did not reflect the EMA recommendation for pembrolizumab, the company appropriately provided details of this trial as supplementary information only and did not use it to inform modelling.

4.2.4 Ongoing trials

KEYNOTE-087 is an ongoing trial but the company stated that all available data had been included in the submission. They further stated in response to clarification that [redacted].[10]

Two further trials were mentioned as being ongoing: KEYNOTE-204 and NCT03077828. KEYNOTE-204 is an ongoing, randomised, non-blinded study of pembrolizumab versus BV in patients with relapsed or refractory cHL. The company stated that KEYNOTE-204 was not within the indication/ license in the submission. NCT03077828 is a single arm, open-label phase II study of pembrolizumab together with a chemotherapy regimen "ICE" (ifosfamide, carboplatin, and etoposide) for the treatment of relapsed/refractory cHL. The company indicates that those patients in this trial who have received BV prior to enrolment may be relevant to the current submission. However the estimated study completion date is February 2020.[1]

ERG comment:

- The ERG believes that none of the ongoing studies could have informed the current submission.
- Further analysis of the KEYNOTE-087 trial may be informative particularly for assessing longer-term OS.

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

The original search (CS, page 45) resulted in one relevant citation: the CSR for KEYNOTE-087. No comparator studies were found. The updated search (CS, page 46) found three more citations for the KEYNOTE-087 trial, but again, no comparator studies. Finally, a separate search for observational studies (CS, page 46) retrieved one relevant study: Cheah et al. 2016.[7]

Cheah et al. 2016 was also used as the comparator study in TA462 (ID972 - Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma). In this appraisal the population of interest was patients who had had previous autologous stem cell transplant and brentuximab vedotin (i.e. cohort 1 in the current appraisal). In TA462, the committee concluded (FAD, point 4.7):

> "The committee considered whether the population and composition of treatments in the Cheah study reflected clinical practice in the UK. The committee noted that the study population partially matched the population of interest because around 70% of patients had previous autologous stem cell transplant and brentuximab vedotin. The committee noted a lack of detail on the precise combinations of chemotherapies given as standard of care in the study, and the inclusion of platinum-based therapies and 'other alkylators'. It considered that the study may not reflect UK practice, particularly regarding subsequent rates of allogeneic stem cell transplant. The committee noted that in response to

consultation, the company had explored UK standard-of-care data from the Haematological Malignancy Research Network and surveyed clinicians actively treating relapsed or refractory classical Hodgkin lymphoma in the UK. The committee considered that both the network data and the clinician survey somewhat supported the Cheah study as reflecting UK practice, but it recognised that the data were limited. The committee concluded that the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision-making, but overall the clinical effectiveness of nivolumab compared with standard of care was highly uncertain because the comparator data may not fully represent UK clinical practice."[6]

ERG comments:

- This means the committee accepted that Cheah[7] is appropriate as a comparator study for people with relapsed or refractory classical Hodgkin lymphoma who have received autologous stem cell transplant and brentuximab vedotin, i.e. cohort 1 in the current appraisal.
- As mentioned above, in Cheah et al. 2016, 70% of patients had received previous autologous stem cell transplant and brentuximab vedotin. In Table 1 of their publication, Cheah et al. 2016[7] report baseline characteristics of 97 included patients before commencement of brentuximab vedotin. Of these 97 patients, 70 had previous stem cell transplantation (SCT), 66 had autoSCT and 4 had alloSCT. The remaining 27 patients did not undergo consolidative transplant; for these, the primary reason was failure to respond to therapy (n = 21, 75%), age or co-morbidities (n = 1, 4%), failed mobilization (n = 1, 4%), patient decision (n = 1, 4%), financial reasons (n = 1, 4%) or reason unknown (n=2, 7%).[7] The CS reports ITT data from Cheah et al., i.e. data for the whole population, with and without transplant. We asked the company to provide separate cohort analyses, using separate data from the two cohorts (with and without transplant) in the Cheah et al. study (Clarification Letter, question A16). The company responded that "MSD do not have access to individual patient level data for Cheah et al. and therefore it was not possible to determine cohorts using the same inclusion criteria as were applied to cohorts 1 and 2 in KEYNOTE-087".[10]
- Using the full Cheah et al. 2016 population as a comparator for cohort 2 is problematic. First of all, only 27 out of 97 patients (28%) did not undergo consolidative transplant. In addition, as shown in Table 4.11 there are differences between the population in cohort 2 and Cheah regarding age, ECOG score, Baseline B symptoms, Haemoglobin, Lymphocytes, Albumin, White cell count and Bulky Lymphadenopathy.
- In summary, the Cheah population is a mixture of both cohort 1 and 2 (as defined in the scope) and is probably most comparable to cohort 1 in the KEYNOTE-087 trial. Because KEYNOTE-087 shows that results are more favourable in cohort 1 compared to cohort 2, using the total population from Cheah as the comparator means results of the naïve comparison will probably overestimate the effect of pembrolizumab in cohort 1 and underestimate the effect of pembrolizumab in cohort 2.

Table 4.11: Baseline characteristics of patients in the included studies

Characteristic		KEYNOTE-087, Cohort 1	KEYNOTE-087, Cohort 2	Cheah et al. (2016)
Treatment		Pembrolizumab 200mg		Mix of therapies including chemotherapy, and investigational agents
Number of patients		69	81	97[α] or 89[β]
Age (median)		34.0	40.0	32[β]
Age >45 (%)		25%[γ]	42%[γ]	14 (14%)[β]
Female (%)		33 (47.8%)	38 (46.9%)	46 (47%)[α]
ECOG	0	29 (42.0%)	44 (54.3%)	33 (41%)[β]
	1	39 (56.5%)	37 (45.7%)	44 (54%)
	2	1 (1.4%)	0 (0.0%)	3 (4%)
Stage	1	NR	NR	2 (3%)[β]
	2	NR	NR	25 (30%)
	3	NR	NR	18 (21%)
	4	NR	NR	39 (46%)
Baseline B symptoms		22 (31.9%)	26 (32.1%)	7 (8%)[β]
Haemoglobin <105 g/l		35%[γ]	27%[γ]	18 of 51 (35%)[β]
Lymphocytes <0.6 × 10^9/l		19%[γ]	15%[γ]	19 of 46 (41%)[β]
White cell count >15 × 10^9/l		9%[γ]	17%[γ]	4 of 82 (5%)[β]
Albumin <40 g/l		48%[γ]	49%[γ]	23 of 82 (28%)[β]
Any extranodal site		56%[γ]	41%[γ]	31 of 88 (35%)[β]
Maximum tumour diameter ≥4 cm		49%[γ]	42%[γ]	18 of 69 (26%)
Bulky Lymphadenopathy		5 (7.2%)	11 (13.6%)	15 (37%)[α]
Bone marrow involvement		3 (4.3%)	5 (6.2%)	NR
Disease status - relapse		46 (66.7%)	24 (29.6%)	NR
Disease status – refractory		23 (33.3%)	57 (70.4%)	NR
Previous BV therapy		69 (100.0%)	81 (100.0%)	89 (100%)[β]
Prior autoSCT		69 (100.0%)	0 (0.0%)	66 of 97 (68%)[α]
Prior radiation		31 (44.9%)	21 (25.9%)	NR
Median no. of prior line of therapy		4	4	3[α]

Source: CS, Table 53, page 137-139[1], CS Appendix 8, and Tables 1 and 2 in Cheah et al. 2016[7].
[α] Sample before commencement of BV (Table 1); [β] Sample at the time of documented progression following therapy with BV (Table 2) – not all characteristics were available from the same sample; [γ] From Appendix 8.
BV = Brentuximab vedotin; ECOG = Eastern Cooperative Oncology Group; N/A = Not applicable; NR = Not reported; SCT = Stem cell transplant

Another issue regarding the population in the Cheah et al. 2016 study is that patients received a wide variety of treatments (see Table 4.12).

Cheah et al. (2016)[7] conducted a retrospective review of their institutional database (at the MD Anderson Cancer Center, Texas) to identify patients who had been treated with BV between June 2007 and January 2015. To be included in the study patients had to meet the following criteria:

- A histologically confirmed diagnosis of classical Hodgkin lymphoma
- Treatment with BV for relapsed Hodgkin lymphoma
- Disease progression at any time after treatment with BV

The aim of the study was to determine PFS and OS following disease relapse after BV therapy. Secondary outcomes were to analyse the efficacy of subsequent therapeutic strategies and to explore candidate prognostic factors for PFS and OS.

Cheah et al. (2016) report that 66/97 (68%) had prior ASCT and 4 (4%) had prior alloSCT conducted at the time of second remission. Data were available on subsequent therapy for 83 patients with disease progression following BV therapy and these data are reproduced below in Table 4.12. The proportion of patients who had prior ASCT among the 83 patients with disease progression is not reported.

Table 4.12: Therapies received by patients in the Cheah et al. study who had disease progression following BV therapy

Treatment	n	Evaluated	CR (%)	PR (%)	ORR (%)	mPFS	mOS
Investigational agent	28	28	4 (14)	3 (11)	7 (25)	2.4 (months)	47.7 (months)
Gemcitabine	15	12	4 (27)	4 (27)	8 (53)	2.1	NR
Bendamustine	12	11	2 (17)	4 (33)	6 (50)	3.7	34.0
Other alkylator	6	4	1 (17)	1 (17)	2 (33)	5.0	9.5
BV retreatment	6	4	0 (0)	2 (33)	2 (33)	3.5	10.4
Platinum based	4	4	0 (0)	1 (25)	1 (25)	0.9	25.2
ASCT	3	3	1 (33)	0 (0)	1 (33)	-	11.9
Other	5	1	0 (0)	0 (0)	0 (0)	-	24.9
Overall	79	67 (85%)	12 (15)	15 (16)	27 (34)	3.5	25.2
No treatment received	4 due to poor performance status and/or patient decision						
TOTAL	83						

ASCT = autologous stem cell transplant; BV = brentuximab vedotin; CR = complete response; mOS = median overall survival; mPFS = median progression-free survival; ORR = objective response rate; PR = partial response

In TA462,[6] the company performed two analyses: using the overall Cheah population (i.e. including efficacy from all the treatments listed above) and using the Cheah population but excluding efficacy data for the n=28 patients who received investigational agents. This was because the 'Investigational Agent' group could have included nivolumab. According to the ERG report for TA462 "only a couple of patients in the study received PD-1 inhibitors (although numerical data to support this statement were not provided)."[11] Results of these analyses showed that excluding data for patients who received investigational agents, improved effectiveness results for Cheah. Results for pembrolizumab for the current STA excluding investigational agents are presented below in Section 4.4.2.

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

4.4.1 Methodology of the indirect comparison

The company presents indirect comparisons for the following outcomes: response rate ORR (CR+PR) and survival (PFS).

The company states that it was not possible to consider OS within the long-term model structure in those who do not receive an alloSCT due to a lack of events during the follow-up period. In addition, no formal method of data analysis was proposed by the company for AEs or HRQoL, as these data are not available within the comparator study Cheah et al. (2016).

For each outcome, the company performed two types of analyses: a naïve indirect comparison (IC) and a matched adjusted indirect treatment comparison (MAIC).

A naïve IC was used to compare pembrolizumab using data from KEYNOTE-087 with standard of care (SoC) using data from Cheah et al. (2016). This was a comparison of two single arms, due to the lack of a randomised comparison. PFS was compared using a Cox proportional hazards (PH) model to obtain a naïve unadjusted hazard ratio (HR) for two scenarios:

1. From study initiation to most recent observation
2. From study initiation to week 12

ORR was compared between pembrolizumab and SoC using a chi-squared test for the same time periods as the PFS analysis.

The MAIC used weighting to match the IPD from KEYNOTE-087 to the summary data from Cheah et al. (2016). The methods provided in NICE DSU report 18 (Methods for population-adjusted indirect comparisons in submissions to NICE)[8] were used and the weights applied to the KEYNOTE-087 data were derived from the inverse odds of being in pembrolizumab compared to SoC.

The initial matching used all variables for which data were available in both KEYNOTE-087 and Cheah et al. (2016). In cases where the algorithm used to estimate the weights did not converge using the full set of baseline characteristics, variables were removed in stepwise fashion in a predetermined order until convergence was achieved.

Weights from the propensity model were then applied to a Cox regression model with the same structure as used for the naïve IC to obtain population-adjusted HRs for the same two scenarios:

1. From study initiation to most recent observation;
2. From study initiation to week 12.

For ORR, the same method was used to estimate weights for each separate comparison. Weighted contingency tables and chi-squared test for difference between pembrolizumab and SoC were used to estimate odds ratios.

The company states they conducted a feasibility assessment that focused on two areas: the compatibility of included studies and the data published on potential confounders i.e. the extent to which adjustment could be made to ensure exchangeability.[23, 24] Compatibility was assessed by comparing study design characteristics such as inclusion and exclusion criteria, study endpoints and methods for outcomes assessments (CS, Section 4.10.12). However, the results of this assessment were not fully reported apart from in tables summarising baseline characteristics for each study. The compatibility assessment by the ERG is as follows:

- *Study design characteristics such as inclusion and exclusion criteria*

Cheah et al. (2016)[7] was a retrospective study including patients with (i) a histologically confirmed diagnosis of cHL, (ii) treatment with BV for relapsed HL and (iii) subsequent disease progression at any time after treatment with BV. Patients' treated with BV as part of frontline HL therapy was excluded.

KEYNOTE-087 was a prospective study including patients with relapsed (disease progression after most recent therapy) or refractory (failure to achieve CR or PR to most recent therapy) de novo cHL and

(1) Have failed to achieve a response or progressed after autoSCT. Patients must have relapsed after treatment with or failed to respond to BV post autoSCT (cohort 1); or

(2) Were unable to achieve a CR or a PR to salvage chemotherapy and did not receive autoSCT. Patients must have relapsed after treatment with or failed to respond to BV (cohort 2).

As can be seen from Table 4.11, there are differences in baseline characteristics between the KEYNOTE-087 cohorts and the Cheah et al. (2016)[7] study, regarding age, ECOG score, baseline B symptoms, Lymphocytes, White cell count, Albumin level, extranodal site, tumour diameter, and Bulky Lymphadenopathy.

- *Study endpoints*
  Cheah et al. (2016)[7] reports PFS, OS, and ORR, CR and PR. OS and PFS were reported as median survival times in months and the CR rate, PR rate and ORR as percentages. The same outcomes are reported in the CS for pembrolizumab. However, OS was not included in the indirect comparisons due to due to a lack of events during the follow-up period.

- *Methods for outcomes assessments*
  The ERG notes that there were differences in how PFS was defined between the pembrolizumab study and Cheah et al. (2016)[7]. In the pembrolizumab study (KEYNOTE-087), PFS was defined as the time from first treatment to disease progression, as assessed by BICR per IWG response criteria for malignant lymphoma and by site review or death due to any cause, whichever occurred first (CSR, page 3). In contrast, the PFS definition in Cheah et al. (2016)[7] was the time in months measured from date of confirmed disease relapse following BV to disease progression or death.

  Regarding treatment response, there are also differences in definitions. Cheah et al. (2016) only state that 'treatment responses were determined according to the 2014 Lugano Classification.[25]'. In the pembrolizumab study (KEYNOTE-087), best Overall Response Rate (ORR) was defined as the proportion of patients in the analysis population who have complete remission (CR) or partial remission (PR) using IWG criteria (Cheson 2007[26]) at any time during the study. In KEYNOTE-087 response at 12 weeks follow-up was also assessed.

In summary, the compatibility assessment shows that there are some differences between the two studies regarding baseline characteristics and methods of outcomes assessment. However, the differences in baseline characteristics are more of a concern for the results of the naïve comparison as this is a comparison of two different studies. The MAIC is less affected as the two studies have been matched as part of the analysis method to try and make them comparable at baseline. Differences in methods of outcome assessment are due to the two different study designs (single-arm prospective study vs. retrospective study) and these are a concern as the individual patient data were not available for Cheah et al. (2016) so the OS and PFS outcomes could not be recalculated to match the methods used in KEYNOTE-087.

The second part of the feasibility assessment conducted by the company focused on the data published on potential confounders i.e. the extent to which adjustment could be made to ensure exchangeability.[23, 24] The baseline characteristics from all patients with available data at the time of documented progression following treatment with BV were applied and used for matching (the number of patients with data varied from 89 for age to 46 for lymphocyte count). According to the company, matching was conducted using all variables for which data were available in both KEYNOTE-087 and Cheah et al. (2016). Appendix 8 of the CS lists the variables included in the matching exercise, these are: ECOG >0 (%), B symptoms (%), Age >45 (%), Albumin <40 g/l (%), Haemoglobin <105 g/l (%), Lymphocytes $<0.6 \times 10^9$ (%), White blood cells $>15 \times 10^9$ (%), Max Tumour Diameter >4 cm (%), Any extranodal site (%), Female (%), and Prior lines (mean/median). As can be seen in Table 4.11, apart from Bulky Lymphadenopathy and prior autoSCT, these are indeed the only available variables for adjustment. The company does not explain why Bulky Lymphadenopathy and prior autoSCT were not included in the matching but did state that in cases where the weighting algorithm did not converge then variables were removed in a stepwise fashion in a predetermined order until convergence was achieved.

The naïve IC results should be treated with caution due to the differences in patient populations and study design between KEYNOTE-087 and Cheah et al. (2016) and the fact that they are a comparison of single-arms and not based on randomised trials. There was no attempt to match the populations used in this analysis so the results are from comparisons of different treatment groups from two single-arm studies of different designs (one prospective and one retrospective).

The MAIC used recommended methods and appears to have been conducted correctly. Initially all baseline variables which were available in both studies were included in the matching algorithm. Variables were only excluded from the matching if there were problems with model convergence. Most analyses only excluded one variable 'median prior lines', but the analysis of ORR for cohort 1 in the 12-week scenario only included four variables in the matching model. The reason for this reduced model was not provided. The baseline characteristics pre- and post-matching for each study and outcome are presented in Appendix 8 and show that a satisfactory match was obtained between KEYNOTE-087 and Cheah et al. (2016) for all eligible variables.

The ERG could not reproduce the MAIC for checking as only the IPD for KEYNOTE-087 were provided by the company. The data for Cheah et al. (2016) were not provided even though it was used in the analysis and the ERG had requested all data and the corresponding R code in the clarification letter.

According to DSU report 18 (Methods for population-adjusted indirect comparisons in submissions to nice)[8] "companies deploying MAIC or STC are not only arguing that the treatment effect is dependent on the population, but they are further assuming that the target population is closer to that represented in the comparator trial than in their own trial." In this case, this means that the MAIC analysis is based on the population characteristics as in the Cheah et al. (2016) study. As stated above, in TA462, the committee "considered that the study may not reflect UK practice, particularly regarding subsequent rates of allogeneic stem cell transplant" and "The committee concluded that the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision-making, but overall the clinical effectiveness of nivolumab compared with standard of care was highly uncertain because the comparator data may not fully represent UK clinical practice."[6].

According to DSU report 18,[8] unanchored indirect comparisons (i.e. those based on single-arm studies) are susceptible to large amounts of systematic error unless all prognostic variables and effect modifiers are accounted for in the propensity score weighting model. However, in the current MAIC the company was dependent on the variables reported in Cheah et al. (2016) and these are unlikely to be all relevant

prognostic variables and effect modifiers. In addition, some variables had to be dropped from some models to enable the models to converge. DSU report 18 recommended that information should be provided on the level of bias likely to be introduced as a result of any covariates that are unaccounted for. However, the company did not provide this due to a "lack of studies in the patient population relevant to this analysis". They did not comment on the degree of systematic error within the MAIC estimates. Therefore, the results are likely to contain systematic error but it is not possible to estimate the size of the potential error.

Summary regarding indirect comparison with Cheah et al. 2016[7]:

- There are problems with compatibility of the two studies (KEYNOTE-087 and Cheah et al (2016)) regarding baseline characteristics and methods of outcomes assessment, although this has a greater impact on the results of the naïve IC as the MAIC does try to match the two groups prior to analysis.
- Using the full Cheah et al. (2016) population as a comparator for cohort 1 is probably acceptable given the committee's discussion in TA462.
- Using the full Cheah et al. 2016 population as a comparator for cohort 2 is problematic, because only 27 out of 97 patients (28%) did not undergo consolidative transplant and there are differences between the population in cohort 2 and Cheah regarding age, ECOG score, Baseline B symptoms, Haemoglobin, Lymphocytes, Albumin, White cell count and Bulky Lymphadenopathy (see Table 4.10).
- The MAIC analysis is based on the population characteristics as in the Cheah et al. (2016) study. These characteristics may not fully represent UK clinical practice.
- The naïve IC results are from two different patient populations and study designs and are likely to be biased as they are not based on data from RCTs.
- The results of the MAIC are likely to include systematic error and the relative treatment effects are only estimated for the target population in the comparator trial (Cheah et al. (2016)).
- Both the naïve IC and MAIC have major limitations and neither are fully reliable for decision making. In the company model and in the ERG analysis the naïve IC is used in the base case and the MAIC in sensitivity analyses.

4.4.2 Results of the indirect comparison

Results for PFS and ORR are reported in Tables 4.13 and 4.14, respectively. The company also presented results for CR and PR (CS, Tables 31-34, pages 98-100).

Almost all results for PFS show a significant benefit for pembrolizumab versus SoC. The only exception is the naïve comparison in cohort 1 in the 12-week scenario, this shows a non-significant difference favouring pembrolizumab but the upper 95% confidence limit only just crosses one. Likewise, all results for ORR significantly favour pembrolizumab over SoC. Results of the naïve comparison are similar to MAIC.

This analysis excluded baseline data from median prior lines in the matching.

Table 4.13: Summary of comparisons of progression-free survival for pembrolizumab versus SoC

Cohort	Comparison	Sample size, n	Pembrolizumab		Hazard ratio (95% CI)
			Events, n	Censored, n	
Entire study scenario					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]

	MAIC		[redacted]	[redacted]	[redacted]
12-Week scenario					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC		[redacted]	[redacted]	[redacted]
Source: CS, Tables 27 and 28, page 96. CI = confidence interval; SoC = standard of care					

Table 4.14: Summary of comparisons of objective response rates for pembrolizumab versus SoC

Cohort	Comparison	Sample size, n	ORR with pem	ORR with SOC	Odds ratio (95% CI)
Best overall response					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
12 Weeks					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
Source: CS, Tables 29 and 30, page 97. CI = confidence interval; ORR = objective response rate; pem = pembrolizumab; SOC = standard of care					

In the clarification letter we asked the company to perform an analysis using the Cheah population but excluding efficacy data for the n=28 patients who received investigational agents (as in TA462) (Clarification question A13). The company was not able to provide such an analysis for PFS, but was able to provide this analyses for response (ORR, CR and PR).

Results for ORR are presented in Table 4.15 below, these results still show a significant benefit for pembrolizumab versus SoC although on the whole less favourable. The analysis of the entire study period excluded baseline data from median prior lines in the matching. For the analysis up to 12 weeks the results for cohort 1 only included four variables in the model (ECOG, B symptoms, age and albumin) but all variables except median prior lines, were included in the model for cohort 2.

Table 4.15: Summary of comparisons of objective response rates for pembrolizumab versus SoC after removing investigational agents

Cohort	Comparison	Sample size, n	ORR with pem	ORR with SOC	Odds ratio (95% CI)
Best overall response					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
12 Weeks					
1	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
2	Naïve	[redacted]	[redacted]	[redacted]	[redacted]
	MAIC	[redacted]	[redacted]	[redacted]	[redacted]
Source: Response to clarification, Question A13, Tables 1 and 2. CI = confidence interval; ORR = objective response rate; pem = pembrolizumab; SOC = standard of care					

4.5 Additional work on clinical effectiveness undertaken by the ERG

No further additional work was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The CS includes a systematic review of the evidence for pembrolizumab and its comparators in patients with classical Hodgkin Lymphoma who have either received autoSCT and BV or BV alone due to autoSCT being unsuitable. No relevant randomised trials were identified.

One study of the efficacy and safety of pembrolizumab was identified (KEYNOTE-087) and this formed the basis of the submission. KEYNOTE-087 is a well-conducted single arm trial including 150 patients relevant to this appraisal. This ongoing multicentre trial includes three UK centres (14 UK patients). The main trial in the CS covers both cohorts of interest (cohort 1: people with relapsed or refractory cHL who have received autologous stem cell transplant and brentuximab vedotin and cohort 2: patients who have received brentuximab vedotin when autologous stem cell transplant is not a treatment option). The company presented data based on a median follow up of 15.9 months. The median time on treatment was [redacted] days for cohort 1 and [redacted] days for cohort 2.

The primary outcome was overall response rate (ORR) as assessed by independent committee. ORR was [redacted] in cohort 1 and [redacted] in cohort 2. In cohort 1 [redacted] of patients had a complete response and in cohort 2 this figure was [redacted] Median progression free survival (PFS) in cohort 1 was [redacted]. In cohort 2 it was [redacted]. Median overall survival (OS) was [redacted]. At 12 months survival was [redacted] in cohort 1 and [redacted] in cohort 2. In cohort 1 [redacted] of patients had one or more adverse events. In Cohort 2 [redacted] of patients had one or more adverse events. The company noted that most AEs were low grade ([redacted] and [redacted] grades 3 to 5 in cohort 1 and 2 respectively). In cohort 1 [redacted] of AEs were classed as serious and in cohort 2 [redacted]. The most common adverse events were pyrexia, cough, fatigue, diarrhoea and vomiting. The company conducted post-hoc analyses of response. The main difference between this post-hoc analysis and the primary analysis of response, referred to as 'best' response rate, is that response was determined at a single time point for each patient i.e. 12 weeks as opposed to any time point up to the point of progression. Data in the form of the proportion who respond by a specific time point was required in order to apportion patients that were progression-free into the cost effectiveness model. The ERG noted that overall response rates were lower at 12 weeks than over the course of the trial ([redacted]

The most important methodological aspect to note are that although the trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative study which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention as the role of natural history and baseline characteristics is not taken into account. As treatment is known to participants, clinicians and assessors this can lead to bias in the delivery of the intervention and the reporting of outcomes. Other limitations in applying the results of the trial to UK practice include:

- Although the study had an adequate follow-up (median 15.9 months) for the primary outcome (ORR), median progression free survival was immature and median overall survival [redacted].
- The trial has only 150 relevant participants so the evidence base for this appraisal is small. Patients over 65 are not well represented. Furthermore, a small number of patients were from the UK (14) so the trial may not totally reflect the UK population and setting. It is recognised,

however, that the population matching the scope of this appraisal from which to draw participants is in itself small.

- In clinical practice, for those who are suitable, pembrolizumab represents a bridge to alloSCT, a potentially curative treatment. However the company submission stated that 'KEYNOTE-087 was not designed as a 'bridging' study, therefore the uptake of alloSCT was very low overall across cohorts 1 and 2.'[1] The company further stated that 'the use of stem cell transplant would have been at the discretion of the treating physician on a per patient basis.'[10]

As KEYNOTE-087 did not have a comparator group the company identified a comparative observational study from the literature (Cheah et al 2016[7]). This is a retrospective USA database study in which patients received the following types of therapy: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or alloSCT, or other treatment. The company has not provided separate data for comparators; instead a combined data set has been provided for multiple comparators, some of which are within the scope and others not. In the previous appraisal of nivolumab (TA462)[6], the committee concluded that "the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision-making, but overall the clinical effectiveness of nivolumab compared with standard of care was highly uncertain because the comparator data may not fully represent UK clinical practice." However, the ERG is not aware of a more appropriate source of data for the comparator population for this appraisal.

The company performed two types of analyses: a naïve indirect comparison between KEYNOTE-087 and Cheah and a matched adjusted indirect treatment comparison (MAIC) of the two studies. The ERG identified problems with compatibility of the two studies regarding baseline characteristics and methods of outcomes assessment. In the MAIC the company adjusted for potential confounding variables so that the KEYNOTE-087 study more closely resembled the Cheah study.

Almost all results for PFS show a significant benefit for pembrolizumab versus SoC. The only exception is the naïve comparison in cohort 1 in the 12-week scenario, this shows a non-significant difference favouring pembrolizumab but the upper 95% confidence limit only just crosses one. Likewise, all results for ORR significantly favour pembrolizumab over SoC. Results of the naïve comparison are similar to MAIC. However, the results of the naïve comparison and MAIC are not reliable because they are likely to contain systematic error but it is not possible to estimate the size of the potential error. Both the naïve IC and MAIC have major limitations when used for decision making.

Although not listed as a comparator in the NICE scope and not referenced in the CS, nivolumab has recently received approval from NICE for this condition. It is recommended 'as an option for treating relapsed or refractory classical Hodgkin lymphoma in adults after autologous stem cell transplant and treatment with brentuximab vedotin.'[6] This represents cohort 1 of this appraisal. It will be important to compare the efficacy and safety of nivolumab and pembrolizumab for this cohort.

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

This section refers to the review of cost effectiveness analysis studies comparing pembrolizumab to comparator therapies in the treatment of RRcHL, as well as the review of studies on health-related quality of life and resource requirements and costs associated with treatment of the patient population, as presented in the CS chapter 5.1[1] and Appendix 12 of the CS.[27]

5.1.1 Objective of cost effectiveness review

All searches presented in the CS relating to cost effectiveness will be summarised and commented on in the following paragraphs.

Objective of cost effectiveness analysis search and review

Three SLRs were performed by the company with the aim of identifying all literature supporting the development and population of a model of patients with relapsed or refractory classic Hodgkin Lymphoma, treated with pembrolizumab. Within the SLRs, the company executed a single set of searches to address the following areas: (1) cost-effectiveness studies of comparator therapies vs. pembrolizumab, (2) health-related quality of life (HRQoL) in the patient population, and (3) resource requirements and costs associated with treatment.

The CS reported that searches were carried out in July 2017. Searches were limited to studies published from 2001-2017, but were not limited by language. Searches were carried out on the following databases: Embase, MEDLINE, MEDLINE in-Process (searched via Pubmed), HTA and NHS EED via the Cochrane library and EconLit. Searches contained facets to identify relevant studies regarding the costs, HRQoL and resource use identification of classical Hodgkin Lymphoma. Searches were carried out in line with the NICE 2013 guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4.[28] Supplementary searches of the following conference proceedings were reported for the previous two years: American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR). The CS also reported that the NICE website was searched in order to identify relevant information from previous submissions not otherwise captured.

ERG comment: The ERG comments are in relation to (a) well reported and reproducible searches and (b) limitation around simultaneous search of two databases.

(a) The majority of searches in Appendix 12 were well reported and easily reproducible. In the original submission, strategies for EconLit, the ASCO, ESMO and ISPOR conference proceedings and a search of the NICE website were not included in Appendix 12, these were provided by the company in their response to clarification.[10]

(b) The ERG asked the company to clarify whether the MEDLINE/Embase strategy reported in Appendix 12, was a single search conducted simultaneously over both the Embase and MEDLINE individual databases or a single search of Embase conducted on the understanding that it now contains all records from MEDLINE. The company responded that "The first search strategy covers evidence from both Embase and MEDLINE using the embase.com interface".[10] The ERG took this as confirmation that a simultaneous search of the two databases had taken place. This approach has limitations when using subject heading terms. It appeared that only Embase subject heading terms (Emtree) were used in the search strategy, and although simultaneous searching of embase.com should automatically identify and search for equivalent MEDLINE subject heading terms (MeSH), as

the ERG does not have access to Embase.com for testing it is not clear if this is the case for all potentially useful MeSH terms. Given the potential limitations of this approach, the ERG considered it preferable to search each database separately, or at least to ensure inclusion of both Emtree and MeSH terms in the search strategy. However, given the additional searches, this is unlikely to have affected the overall recall of results.

5.1.2 Inclusion/exclusion criteria used in the study selection

Complete lists of inclusion and exclusion criteria are provided in the CS (CS Table 52)[1] and in CS Appendices 13 and 15.[27] Below a summary of the inclusion criteria is provided:

Population: adult patients with relapsed/refractory cHL, irrespective of age or gender (mixed populations were excluded unless subgroup data on the population of interest was provided).

Intervention and comparator: No restriction, all pharmacological interventions to be captured.

Outcomes:

1) Studies including a comparison of benefits and costs between the intervention and comparator arms. Results expressed in incremental costs, incremental cost effectiveness ratio (ICER), quality-adjusted life-years (QALYs), life-years gained (LYG) or other measures of effectiveness additional to costs.
2) Studies reporting health state utilities of interest
3) Studies reporting costs

Study type:

1) Cost effectiveness analysis, cost utility analysis, cost benefit analysis, cost minimisation analysis, budget impact analysis, cost consequence analysis,
2) Studies using European Quality of Life-5 Dimensions (EQ-5D), European Organisation for Research and Treatment Cancer Quality of Life Questionnaire (EORTC QLQ-C30), short form 36 health survey (SF36), health utility index (HUI), Visual Analogue Scale (VAS), Time trade off (TTO) or Standard Gamble (SG),
3) Cost studies, surveys, burden of disease and resource use studies

Other: studies published from 2001 onwards, full text in English language and reporting UK specific data (cost data from other countries allowed).

5.1.3 Included/excluded studies in the cost effectiveness review

A total of 2,051 references were identified in the SLRs.

(1) Of 848 identified cost effectiveness references, 52 duplicates were removed and 796 abstracts were screened which led to the exclusion of 694 articles. Consequently, 102 full texts were screened, all of which had to be excluded (see CS p. 135 for the PRISMA diagram[1]). No cost effectiveness studies were included through other searches.

(2) Of 1,236 references identified on HRQoL and utilities, 95 duplicates were removed and 1,141 abstracts were screened which led to the exclusion of 1,091 articles. Subsequently, 50 full-texts were screened of which 46 were excluded and two studies from four publications were included (see CS p. 187 for the PRISMA diagram[1]).

(3) Of 882 identified cost articles, 52 duplicates were removed and 830 abstracts were screened which led to the exclusion of 728 articles and the full-text screening of 102 articles. After the exclusion of 86

more articles and the inclusion of one article through conference searching, a total of 14 studies from 17 publications were included, one of them reporting UK-specific costs and resource use (see CS p. 187 for the PRISMA diagram[1] and Appendix 14 of the CS[27] for a list of studies included).

5.1.4 Conclusions of the cost effectiveness review

No cost effectiveness studies in patients with RRcHL were identified that met the inclusion criteria, therefore the company conducted a de novo health economic analysis. The majority of relevant utility studies identified did not use EQ-5D data and were thus inconsistent with the NICE reference case, or reported utilities not stratified by response. Disutilities of grade 3+ adverse events (AEs) were sourced from previous TAs (see Table 78 of the CS[1]). Fourteen cost studies were found to meet the inclusion criteria. As the updated publication of one of the identified cost studies (Radford 2017,[29]) was the preferred source of cost data in TA462[6] (Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma), this reference was selected to inform the economic analysis. AE costs were computed from a weighted average of Healthcare Resource Group (HRG) code prices.

ERG comments:

The ERG agrees that in the absence of cost effectiveness studies performed on the population and intervention of interest from the literature, a de novo approach was necessary. It was, however, unclear why the company did not provide a complete overview of the publications included and excluded from their cost effectiveness, cost and resource and utility and HRQoL SLRs. Furthermore, the number of references found on EconLit was reported inconsistently in CS Appendix 12 [27] and PRISMA diagrams (CS pages 187 and 198). In their response to clarification question B2, the company explained that the PRISMA diagrams contain the correct number of publications. The ERG wishes to point out that the company prioritised aligning their sources with TA462 over using the results of their SLRs.

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

Table 5.1: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS)
Model	A state transition model with a decision tree element to predict response and alloSCT uptake	To provide an estimate of the lifetime costs and effectiveness of a "bridging" treatment to alloSCT or continued treatment with pembrolizumab.	Chapter 5.2.2
States and events	Health states week 0-12 (short-term component): - Progression free, consisting of complete response, partial response and stable disease - Progressed disease - Death Health states after 12 weeks (long-term component): - Alive post-alloSCT	The short-term health states capture initial treatment response determining alloSCT uptake. The long-term health states describe progression free survival, overall survival and response conditional on alloSCT uptake or continued use of pembrolizumab or SoC. A post-alloSCT progressed disease health state is missing with the justification that the	Chapter 5.2.2

	Approach	Source/Justification	Signpost (location in CS)
	- Progression free non-alloSCT - Progressed disease non-alloSCT - Death	implications of progression post-alloSCT are included in the post-alloSCT alive health state utilities and costs.	
Comparators	- SoC, consisting of chemotherapies (38.5%), treatment with investigational agents (43.1%) and bendamustine (18.5%) - BSC (only in scenario analysis)	SoC was included as a comparator as it contained combination chemotherapy such as gemcitabine, vinblastine and cisplatin, as defined in the scope [4]. Although also identified as a comparator in the scope [4], BSC was only used as a comparator in a scenario analysis. According to expert opinion, the use of BSC in UK practice is minimal.	Chapter 5.2.4
Population	Adult patients with RRcHL who have failed autoSCT and BV (cohort 1), or who are autoSCT ineligible and have failed BV (cohort 2).	This is consistent with the final scope issued by NICE and the population of the KEYNOTE-087 trial	Chapter 5.2.1
Treatment effectiveness	Due to the characteristic of pembrolizumab as a "bridging" treatment to alloSCT, treatment effectiveness was driven by the proportion of patients responsive to treatment at 12 weeks, allowing for alloSCT uptake. Pre-12 week effectiveness was informed by OS and PFS curves. Post-12 weeks, in the non-alloSCT pathway, OS and PFS curves were fitted. Post-12 weeks in the alloSCT pathway, OS was independent of prior treatment.	Proportional hazards were assumed to hold for all estimates. Comparative effectiveness and response were estimated by a naïve comparison of single-arm studies Cheah et al [7] and KEYNOTE-087. A matched indirect comparison was performed in a scenario analysis in order to avoid data loss in the base-case. Probabilities for alloSCT conditional on response states were elicited from UK clinical experts via two clinician surveys because the KEYNOTE-087 study was deemed non-generalizable to the UK setting. Post-alloSCT OS estimates were derived from Lafferty et al [21], because there was insufficient long-term data in the KEYNOTE-087 study. In the non-alloSCT post-progression	Chapter 5.3 Chapter 5.3.1

	Approach	Source/Justification	Signpost (location in CS)
		health state, mortality was based on Cheah et al.[7]	
Adverse events	Resource use, costs and utility decrement (one-off) were considered for AEs grade 3+	AEs with an incidence of >0% in either treatment arm, in line with TA462[6] were selected. Disutilities stemmed from literature sources used in previous TAs.[30, 31, 32, 33, 34, 35, 36]	Chapter 5.3.5
Health related QoL	Utilities are based on the KEYNOTE-087 study (using 12 week EQ-5D data only) in combination with (treatment specific) response rates. Moreover, utility values obtained from the literature were used.	Response-specific values elicited consistently with committee preference in TA462.[6]	Chapter 5.4.7 Chapter 5.4.8 Chapter 5.4.6
Resource utilisation and costs	Resource use and costs accounted for in the model are drug acquisition costs, administration costs, monitoring costs, adverse events costs, costs of subsequent treatment, and terminal care costs.	KEYNOTE-087 and Cheah[7] studies and published sources were used when they provided estimates of resource use and costs. This approach had been validated by expert opinion in previous submissions. Sources used are the eMIT,[37] BNF, the KEYNOTE-087 and Cheah[7] studies and studies used in TA462.	Chapter 5.5
Discount rates	Discount of 3.5% for utilities and costs	As per NICE reference case	Table 54
Sub groups	Not applicable		
Sensitivity analysis	Both DSA and PSA were performed as well as scenario analyses	As per NICE reference case	Chapter 5.8

Source: CS[1]

AE = adverse events; alloSCT = Allogeneic Stem Cell Transplant; autoSCT = Autologous Stem Cell Transplant; BNF = British National Formulary; BV = Brentuximab Vedotin; BSC = best supportive care; CS = company submission; DSA = deterministic sensitivity analysis; eMIT = electronic market information tool; OS = overall survival; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; RRcHL = relapsed or refractory classical Hodgkin Lymphoma; SLR = systematic literature review; SoC = standard of care.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.2: Summary of the company's economic evaluation (with signposts to CS): NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope[4]	Yes	
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Partly	BSC is only used in a scenario
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Partly	Time horizon of 40 years, used in the base-case, does not capture all relevant costs and effects (illustrated in CS scenario analysis 5[1])
Synthesis of evidence in outcomes	Systematic review	Yes	SLR and naïve treatment comparison.
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Yes	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	Yes	

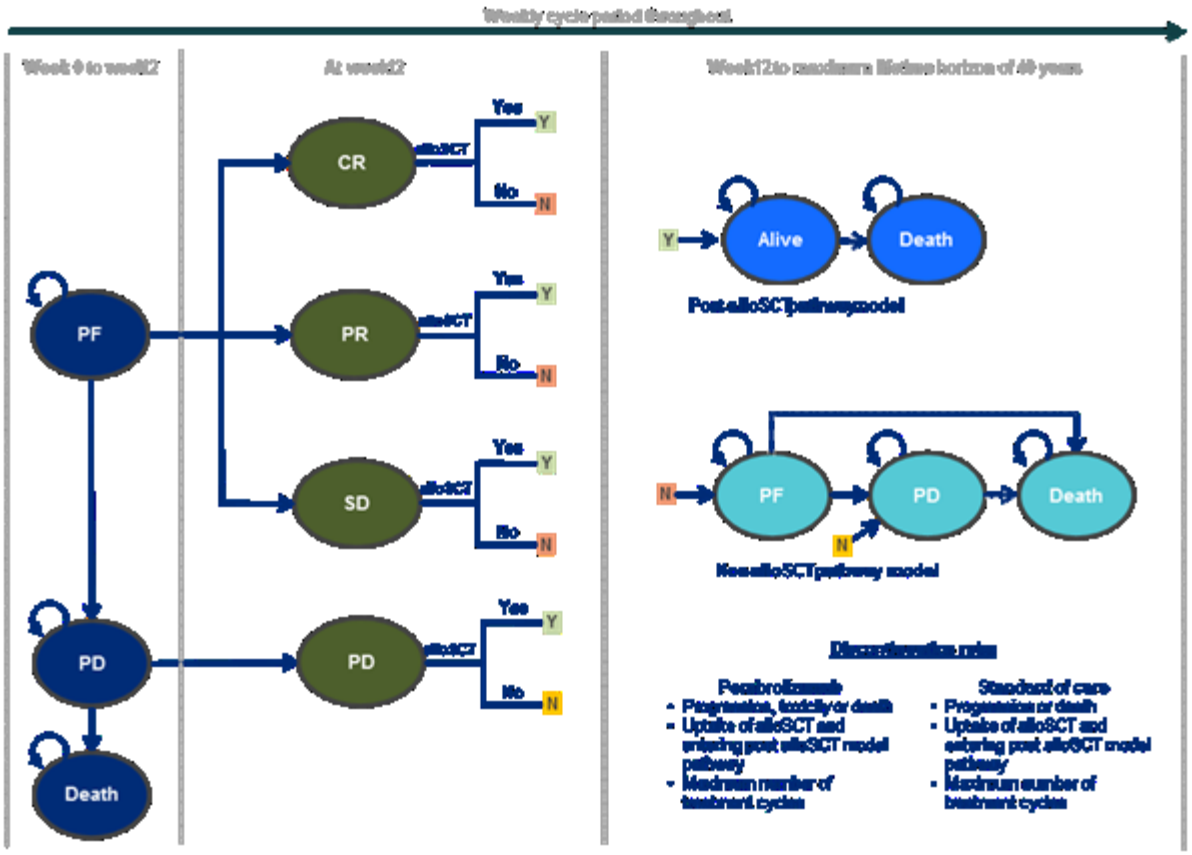
Source: CS [1]
NHS = National Health Service; PSS = Personal Social Services; QALY = quality-adjusted life year

5.2.2 Model structure

A de novo cohort state transition model was developed with health states based on response, uptake of alloSCT, and survival. This approach was adopted as it is expected that pembrolizumab monotherapy will result in higher response rates than SoC and hence will be used as a "bridge" to alloSCT. More specifically, pembrolizumab aims to control the disease, and if possible, elicit a disease response that enables alloSCT. The model has a time horizon of 40 years, weekly cycle length and applies a half-cycle correction.

The model structure consists of a short-term component (first 12 weeks), a subsequent decision tree element (at 12 weeks) to determine the proportion of patients transiting to alloSCT (conditional upon response at 12 weeks) and a long-term component (after the first 12 weeks) separately for patients who had alloSCT and patients who did not have alloSCT at 12 weeks (See Figure 5.1).

Figure 5.1: Model structure



Source: CS Figure 13[1]

Short term model structure (pre-12 weeks)

A partitioned survival approach is used for the first 12 weeks with three health states:

1. Progression free;
2. Progressed disease;
3. Death

Decision tree element (at 12 weeks)
After 12 weeks, the progression free proportion is subdivided into proportions of patients with complete response, partial response and stable disease. Here, patients with non-evaluable response status are categorised as having stable disease. Subsequently, depending on this response status, the proportion of patients continuing to alloSCT is calculated (i.e. patients with complete response have a higher probability of receiving alloSCT than patients with partial disease or stable disease). The company assumed that none of the patients with progressed disease will continue to alloSCT (see Section 5.2.6 for more details).

Long-term model structure (post-12 weeks) separately for alloSCT and non-alloSCT treatment
After the 12-week decision tree element, the cohort is split into patients who did and did not receive alloSCT.

Patients who did not receive alloSCT at 12 weeks will not be able to receive alloSCT for the remainder of the model time horizon. Further, the long-term costs and effects for this group are modelled using three health states consistent with the short-term model structure (for the first 12 weeks):

1. Progression free (patients who did not have alloSCT, progression, or died in the first 12 weeks);
2. Progressed disease;
3. Death

The long-term costs and effects for patients who did receive alloSCT at 12 weeks is modelled using two health states:

1. Alive (patients who did not have progression in the first 12 weeks and did receive alloSCT at week 12);
2. Death

Post-alloSCT survival was assumed to be independent of prior therapy (i.e. equal for patients who initially received pembrolizumab monotherapy and SoC). Moreover, the company justified not considering post-alloSCT progression in the model structure by claiming that the consequences of post-alloSCT progression are incorporated in the post-alloSCT utilities and costs.

ERG comment: The ERG notes the following issues regarding the model structure used by the company: (a) in the model it is only possible to have alloSCT 12 weeks after treatment start, (b) the assumption that alloSCT would be performed immediately after response; (c) neglecting a progression health state after alloSCT.

(a) The model structure only allows patients to have alloSCT at 12 weeks after starting pembrolizumab or SoC. No justification was provided for why this simplifying approach was adopted. This is of particular concern given that one of the main goals of pembrolizumab is to enable alloSCT and hence this should be reflected in the model as accurately as possible. Therefore, the ERG requested an analysis removing this assumption (i.e. incorporating a continuous alloSCT probability). However, in response to clarification question B4d, the company stated that they could not perform such an analysis given that 1) they believed that alloSCT data from KEYNOTE-087 are not reflective of UK clinical practice and; 2) they did not have Kaplan-Meier data for time-to-alloSCT from Cheah et al.[7]

Furthermore, the 12-week timepoint is questionable. It was selected based on a UK clinician survey and the company stated (clarification question B4a) that this timepoint is an accurate representation of the timing of the decision to transplant. The company recognised that response might be obtained later than week 12, but believed the assumption that these 'later responders' would not be considered for alloSCT

to be conservative. The ERG is not convinced that this statement is correct given that this was not appropriately explored by the company and it is unclear how many 'later responders' exist for both pembrolizumab and SoC.

The company's approach is furthermore inconsistent with the approach taken in TA462.[6] The company refers to TA462[6] on multiple occasions to highlight the similarities. This includes similarities regarding the mechanism of action of pembrolizumab and nivolumab stating that both may act as therapy to enable alloSCT. Therefore, it is questionable why the company opted to use a different model structure than in TA462.[6] In TA462,[6] alloSCT is assumed to be performed at six months.

(b) Another related concern is that the company assumes an immediate procedure at the 12-week time point. The company's model structure estimates the proportion of patients undergoing alloSCT based on response at week 12 after starting pembrolizumab or SoC and alloSCT would be performed immediately. This, however, neglects the time required to identify a donor and schedule the procedure. The lag is estimated to be on average [redacted] weeks from eligibility decision to the actual performing of alloSCT (given the company stated treatment is stopped on average [redacted] weeks prior to alloSCT). Hence, the decision to perform alloSCT might be made at 12 weeks, the actual procedure might be performed between 12 and 24 weeks (response to clarification question B4a). This would also be more consistent with TA462[6] wherein it is stated that "*Based on CheckMate 205 and the published literature, it has been assumed that a proportion of eligible patients with an adequate response will receive alloSCT at six months.*" This entails that alloSCT in the present model is performed earlier than would be expected in clinical practice. Hence, the post-alloSCT benefits (e.g. lower mortality probability and higher quality of life) are applied earlier. Given that the proportion of patients proceeding to alloSCT is higher for pembrolizumab than for SoC, this is most likely not a conservative assumption.

(c) As highlighted by the company (CS section 5.2.2), one of the main criticisms on partitioned survival models (recent Decision Support Unit report[38]), is that OS, a key driver of QALY gains in advanced oncology, is modelled independently of an underlying disease model. Hence, partitioned survival models might result in inappropriate extrapolations. This critique is applicable to the long-term post-alloSCT component of the model in which disease progression is not considered despite Lafferty et al[21] reporting a progression free survival one-year post-alloSCT of only 54%. Given that post-alloSCT survival is modelled independently of an underlying disease model, this likely biases the long-term extrapolations, in favour of pembrolizumab. This is also inconsistent with TA462 in which post-alloSCT progression was incorporated.

5.2.3 Population

According to its marketing authorisation, pembrolizumab monotherapy is indicated for the treatment of adult patients with relapsed or refractory classical Hodgkin lymphoma (RRcHL) who have failed autoSCT and brentuximab vedotin (BV) (cohort 1), or who are transplant ineligible and have failed BV (cohort 2). In line with this marketing authorisation and the final scope issued by NICE,[4] two distinct populations are considered in the cost effectiveness model:

- Cohort 1: RRcHL who have failed autoSCT and BV
- Cohort 2: RRcHL who are autoSCT ineligible and have failed BV

See Table 5.3 for the baseline characteristics for cohorts 1 and 2 (from the main evidence sources considered in the model).

Table 5.3: Baseline characteristics from the main evidence sources considered in the model

Characteristic		KEYNOTE-087, Cohort 1	KEYNOTE-087, Cohort 2	Cheah et al. (2016),[7] Cohorts 1 and 2
Treatment		Pembrolizumab 200mg		Mix of therapies including chemotherapy, and investigational agents
Number of patients		69	81	97 or 89
Age (median)		34.0	40.0	32
Age >45 (%)		25%[γ]	42%[γ]	14 (14%)
Female (%)		33 (47.8%)	38 (46.9%)	46 (47%)
ECOG	0	29 (42.0%)	44 (54.3%)	33 (41%)
	1	39 (56.5%)	37 (45.7%)	44 (54%)
	2	1 (1.4%)	0 (0.0%)	3 (4%)
Stage	1	NR	NR	2 (3%)
	2	NR	NR	25 (30%)
	3	NR	NR	18 (21%)
	4	NR	NR	39 (46%)
Baseline B symptoms		22 (31.9%)	26 (32.1%)	7 (8%)
Haemoglobin <105 g/l		35%	27%	18 of 51 (35%)
Lymphocytes <0.6 × 10^9/l		19%	15%	19 of 46 (41%)
White cell count >15 × 10^9/l		9%	17%	4 of 82 (5%)
Albumin <40 g/l		48%	49%	23 of 82 (28%)
Any extranodal site		56%	41%	31 of 88 (35%)
Maximum tumour diameter ≥4 cm		49%	42%	18 of 69 (26%)
Bulky Lymphadenopathy		5 (7.2%)	11 (13.6%)	15 (37%)
Bone marrow involvement		3 (4.3%)	5 (6.2%)	NR
Disease status - relapse		46 (66.7%)	24 (29.6%)	NR
Disease status – refractory		23 (33.3%)	57 (70.4%)	NR
Previous BV therapy		69 (100.0%)	81 (100.0%)	89 (100%)
Prior autoSCT		69 (100.0%)	0 (0.0%)	66 of 97 (68%)
Prior radiation		31 (44.9%)	21 (25.9%)	NR
Median no. of prior line of therapy		4	4	3

Source: CS, Table 53, CS Appendix 8, and Tables 1 and 2 in Cheah et al. 2016
BV = Brentuximab vedotin; ECOG = Eastern Cooperative Oncology Group; N/A = Not applicable; NR = Not reported; SCT = Stem cell transplant

ERG comment: The populations described by the company are consistent with the final scope issued by NICE for this appraisal,[4] but one concern relates to the use of a mixed population comparator.

For KEYNOTE-087, the company was able to distinguish between patients who did and did not receive autoSCT (i.e. cohort 1 and 2 respectively). For the study by Cheah et al.,[7] the company did not have

access to the individual patient level data and hence was unable to make this distinction. Hence, the mixed population from Cheah et al.,[7] including both patient groups that did and did not receive autoSCT, was used for both cohort 1 and 2. Given that the majority of patients (68%) in the study by Cheah et al.[7] did receive autoSCT, this mixed population is more reflective of cohort 1. Additionally, the Cheah et al.[7] population is more comparable with KEYNOTE-087 cohort 1 than with cohort 2 in terms of patient characteristics (see for instance baseline age, ECOG, haemoglobin and white cell count). For other baseline variables (e.g. baseline B symptoms, lymphocytes, albumin, extranodal sites, tumour diameter >=4cm, and bulky lymphadenopathy) the Cheah et al.[7] population differs from both KEYNOTE-087 cohorts.

In response to clarification question B9 the company states that, based on clinical opinion, cohort 2 represents a higher risk group that is likely to progress more quickly compared with cohort 1. If this is the case, using the mixed population from Cheah et al.[7] in the naive comparison likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively.

5.2.4 Interventions and comparators

Pembrolizumab monotherapy is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration for RRcHL (i.e. administered intravenously at a fixed dose of 200mg over 30 minutes every 3 weeks [Q3W]). The company assumed that in the model pembrolizumab monotherapy will be provided for a maximum of 24 months (35 cycles).

The NICE scope specifies the following comparators:

- Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin;
- Best supportive care (BSC).

The company only considered "standard of care" (SoC) as comparator in its base-case. SoC as considered by the company consists of the following regimens:

- chemotherapy (see CS Table 88 for the included treatments);
- bendamustine or;
- investigational agents.

The distribution of patients among these regimens was based on the distribution observed in Cheah et al (2016)[7] (see CS Table 88).

The company also presented a scenario analysis, in which BSC was added as a comparator. The company justified this deviation from the scope (i.e. not including BSC in its base-case) by stating they believed BSC use to be minimal as eligible patients are likely to receive therapy whenever feasible.

ERG comment: The ERG has concerns regarding (a) the exclusion of BSC from the base-case, (b) the recent recommendation of nivolumab in part of this population, which is not reflected in the analysis, (c) the assumption that pembrolizumab treatment stops at 24 months, and (d) the inclusion of investigational agents in the comparator.

(a) Regarding the inclusion of comparators, the ERG wishes to highlight that BSC is not incorporated in the CS base-case (inconsistent with the scope), but only presented in a scenario analysis.

(b) Moreover, nivolumab was recommended by NICE in part of this population (cohort 1). Nevertheless, NICE (personal communication with [redacted]) suggested that it would be

inappropriate to include nivolumab as a new comparator given it is still within the 90-day implementation period and hence is not considered established practice.

(c) The assumption that pembrolizumab monotherapy will be stopped after 24 months (35 cycles) is inconsistent with the SmPC but in line with the KEYNOTE-087 protocol. It is unclear whether pembrolizumab, in UK clinical practice, would also be provided for a maximum of 24 months. The company explored the impact of this assumption in a scenario in which patients continue treatment after 24 months. This scenario increased the CS base-case ICER for both cohorts (response to clarification question B13).

(d) Finally, the company uses the total population from Cheah et al (2016),[7] including patients that received investigational agents. Given that excluding patients that received investigational agents might result in a selected patient sample, the ERG believes this approach is reasonable. Moreover, the appropriateness of using the patients that used investigational agents in the Cheah et al (2016)[7] study was discussed in the final appraisal determination (FAD) of TA462. The committee preferred to use the overall population from Cheah et al (2016)[7] given that it considered that *"selectively excluding potentially the fittest patients from the Cheah dataset could bias the results of the indirect treatment comparison more than including some treatments that may not be used in UK current practice"*.[6]

5.2.5 Perspective, time horizon and discounting

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The cycle length is one week to account for the length of treatment cycles. A time horizon of 40 year was adopted to capture all relevant costs and outcomes. All costs and utilities were discounted at a rate of 3.5% per year.

ERG comment: The ERG considers the adopted perspective and discounting to be appropriate for this appraisal. The time horizon of 40 year might be considered suboptimal given that CS scenario 5 (CS section 5.8.3) suggests that this time horizon is insufficient to capture all costs and outcomes. Therefore, the ERG preferred to use a 50-year time horizon in its base-case.

5.2.6 Treatment effectiveness and extrapolation

Treatment effectiveness for pembrolizumab was primarily based on the KEYNOTE-087 study.[10] The only comparator in the company's base-case was SoC. The primary data source for the SoC comparator was the Cheah et al (2016) study.[7] The company performed a naïve indirect treatment comparison to derive hazard ratios for OS and PFS and response rates at week 12. A MAIC was also performed and results are shown in the company's scenario analysis. Both KEYNOTE-087 cohorts were compared with the Cheah et al (2016) study.[7] In a scenario analysis, the company explored BSC as a comparator. Because no data were available to inform this comparator, the efficacy of SoC was used (CS p. 149).

Due to the company's model structure, treatment effectiveness and time to treatment discontinuation (TTD) were estimated for the pre-12-week period and for the post-12-week period separately. Parametric models were fitted to data from KEYNOTE-087 to estimate OS and PFS for patients receiving pembrolizumab in the pre-12-week period. To inform the decision tree element at week 12, response rates from KEYNOTE-087 were used, as well as two clinician surveys to inform estimates of probability of alloSCT conditional on response status (i.e. complete response, partial response, stable disease). For the post-12-week period, treatment effectiveness depended on whether patients received alloSCT or not. Mortality post-alloSCT was based on Lafferty et al[21] and mortality for patients who did not receive alloSCT was based on Cheah et al.[7] PFS for patients who did not receive alloSCT was

estimated from KEYNOTE-087. The company justified this inconsistency by stating that survival data from KEYNOTE-087 were immature.

TTD for the pre-12-week period was assumed to be equivalent to PFS. TTD for the post-12-week period was estimated directly from KEYNOTE-087. Furthermore, TTD for SoC was assumed equivalent to PFS in Cheah et al for pre- and post-12 weeks.[7]

Table 5.4 presents an overview of use and justification of all parametric models for PFS and OS extrapolations in the two periods, with more detail provided in the following sections.

Table 5.4: Overview of parametric models used for extrapolating OS and PFS in company model

	Parametric model used in company base-case	Best statistical fit? (if No: which one?)	Other justification provided?	Alternative explored in company scenario analysis?	Source used for pembrolizumab
Cohort 1					
Pre-12 weeks PFS	Log-logistic	Yes	None	No	KEYNOTE-087
Pre-12 weeks OS	Lognormal	No (exponential)	Predicted highest mortality	No	KEYNOTE-087
Post-12 weeks PFS	Exponential	Yes	None	No	KEYNOTE-087
Post-12 weeks (non-alloSCT) OS	Constant transition probability estimated from median OS		No KM estimates available from Cheah, KEYNOTE-087 data too immature	Yes, KEYNOTE-087 data were explored in scenario analysis	Cheah et al (2016)[7]
Post-12 weeks (-alloSCT) OS	Weibull	No (gen gamma)	Gen gamma predicted an infinite hazard beyond 150 months and had to be adjusted, thereby under-estimating the survival benefit; AIC/BIC scores were relatively similar; ERG in TA462 considered lognormal and Weibull most clinically plausible	Lognormal	Lafferty et al (2017)[21]
Pre-12 weeks TTD	Same as pre-12 weeks PFS				KEYNOTE-087
Post-12 weeks TTD	Exponential	Yes	Maintained consistency with post-12 week PFS	No	KEYNOTE-087

	Parametric model used in company base-case	Best statistical fit? (if No: which one?)	Other justification provided?	Alternative explored in company scenario analysis?	Source used for pembrolizumab
Cohort 2					
Pre-12 weeks PFS	Generalised gamma	Yes	None	Weibull	KEYNOTE-087
Pre-12 weeks OS	Exponential	No (lognormal)	None	No	KEYNOTE-087
Post-12 weeks PFS	Exponential	No (gen gamma)	The last drop in PFS was not considered informative, due to small patient numbers at risk	Gompertz	KEYNOTE-087
Post-12 weeks (non-alloSCT) OS	Constant transition probability estimated from median OS		No KM estimates available from Cheah, KEYNOTE-087 data too immature	Yes, KEYNOTE-087 data were explored in scenario analysis	Cheah et al (2016)[7]
Post-12 weeks (-alloSCT) OS	Same as for cohort 1				Lafferty et al (2017)[21]
Pre-12 weeks TTD	Same as pre-12 weeks PFS				KEYNOTE-087
Post-12 weeks TTD	Same as for cohort 1				KEYNOTE-087
OS = overall survival; PFS = progression free survival; TTD = time to treatment discontinuation					

ERG comment: The ERG's general comments on treatment and relative treatment effectiveness used in the model relate to (a) inconsistency in the choice of data sources prompted by the immaturity of OS data in KEYNOTE-087, (b) the lack of BSC as a comparator, (c) the use of a naïve indirect comparison and (d) the use of differential parametric models for the pre- and post-12-week periods.

(a) For the post-12 weeks period, the company deviated from their main data source and used the Cheah and Lafferty et al studies to inform mortality for patients without and with alloSCT respectively. This was justified by the company by stating that KEYNOTE-087 OS data were too immature to be used. [redacted] and the ERG considers that these may be informative for the present analysis.

(b) The lack of BSC as a comparator is non-compliant with the scope. The company justified this stating that there were no data to inform this comparison, and provided a conservative scenario analysis in which the effectiveness of BSC was assumed equivalent to that of SoC.

(c) The company's argument for preferring the naïve treatment comparison to minimise data loss (see CS p 149) is plausible in the context of small sample sizes. The MAIC is deemed to introduce systematic error, due the limited availability of prognostic variables. The ERG therefore maintains the naïve

comparison in its base-case and the MAIC is explored in scenario analysis. The naïve comparison favours SoC.

(d) The artificial 12-week time point necessitated the fitting of differential curves to the pre- and post-12-week periods. This leads to loss of data introducing further uncertainty in the extrapolation.

5.2.6.1 Pre-12 weeks: PFS and OS

PFS pre-12 weeks

PFS pre-12 weeks was modelled based on the entire observed data set from KEYNOTE-087 beginning in week 0 to the end of study follow-up. The company justified this by stating that there was only a small number of events occurring in the first 12 weeks.[10] The log-logistic model was deemed to best represent PFS for cohort 1 and the generalised gamma for cohort 2 (based on best statistical fit). The company stated that in cohort 2, the generalised gamma over-predicted the number of patients in the progression-free survival health state and the company explored the Weibull in a scenario analysis, claiming that it would result in fewer patients in the progression-free health state at 12 weeks.

Relative effectiveness was based on the naïve treatment comparison.

OS pre-12 weeks

OS pre-12 weeks was also based on KEYNOTE-087. With very few events, there was no meaningful difference between the different parametric models in terms of statistical fit and the company selected the lognormal model for cohort 1, which predicted the highest mortality but did not have the best statistical fit. For cohort 2, [redacted]. The company chose the exponential model, without providing appropriate justification.

The company assumed that patients treated with SoC would follow the same OS curve as patients receiving pembrolizumab.

ERG comment: The ERG wishes to highlight a few caveats with the company's pre-12 weeks analysis, including (a) the fitting of parametric models for the pre-12-week period using the entire study data, and (b) the poor fit of models for OS in both cohorts, which produces artificially lowered LYs and counter-intuitive results.

(a) Only very few events occurred in the first 12 weeks of the KEYNOTE-087 study. For example, for PFS, more than [redacted] of patients in cohort 1 and approximately [redacted] of patients in cohort 2 were still progression-free at 12 weeks. The fitted models were estimated using the entire study data from week 0 to end of study follow-up, which may have led to the fitted curves being more influenced by the post-12-week period than the pre-12-week period. This is exacerbated for PFS in cohort 2. This is because the KM estimates show that there is a significantly increased rate of progression starting at 11 weeks. This sudden drop, as well as having the parametric models fitted to the entire study data, results in most of the curves not providing a good fit. Furthermore, the scenario analysis using the Weibull over-predicts patients in the progression-free health state even more than the base-case generalised gamma, contrary to the claims of the company. This analysis is therefore disregarded by the ERG, as the only rationale for scenario analysis using the Weibull for PFS in cohort 2 was that it over-predicted PFS to a lesser extent than the generalised gamma. The ERG therefore considers the company's adopted approach of deriving pre-12 weeks PFS and OS estimates from the entire study data as questionable.

(b) For both cohorts, the company chose the pre-12-week OS models that predicted the highest mortality at 12 weeks, disregarding statistical fit (lognormal for cohort 1 and exponential for cohort 2). This likely

produces an artificially lowered number of life-years (LYs) gained, however, it may be worth noting that the company's economic model overall predicts LYs that were considered by the company to be high[1] due to the inclusion of investigational agents in Cheah et al.[7] The combination of using the generalised gamma for PFS and the exponential for OS in cohort 2 also resulted in the crossing of PFS and OS curves in the model (first PFS > OS, then PFS < OS). The company remedied this by choosing whichever was smaller in the simulation of PFS. The ERG preferred to use the model with the best statistical fit in their base-case. This, however, did not solve the problem of crossing OS and PFS.

5.2.6.2 At 12 weeks: response rates and alloSCT probabilities

Response rates at 12 weeks

The distribution across the response states of complete response (CR), partial response (PR), stable disease (SD) and progressed disease (PD) was based on the observations from the KEYNOTE-087 study. The company only presented the patient numbers for cohort 1 in Table 62 of the CS,[1] but corrected this in their response to the clarification letter (see Table 5.5).[10] The company furthermore highlighted in response to the clarification letter that all patients with a non-evaluable response status were assumed to have SD, and presented response rates in comparison with model predictions (Table 5.6).

Response rates at 12 weeks for SoC were based on odds ratios for response derived from the naïve treatment comparison.

Table 5.5: Response rates derived from KEYNOTE-087

Response	n	N
Cohort 1		
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]
Cohort 2		
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]
Source: Response to clarification letter[10]		

Table 5.6: Response rates and model predictions

Status	Table 19 of submission (cohort 1) N (%)	Model predictions (cohort 1) %	Table 19 of submission (cohort 2)	Model predictions (cohort 2)
Complete response	[redacted]	15.94%	[redacted]	8.6%
Partial response	[redacted]	42.0%	[redacted]	43.2%
Stable disease	[redacted]	36.9% *(~27.5%+8.7%)*	[redacted]	38.9% (~18.5% + 8.6%)
Non-evaluable	[redacted]	Not reported (combined in stable disease)	[redacted]	Not reported (combined in stable disease)
Progressed disease	[redacted]	4.10%	[redacted]	7.9%

Status	Table 19 of submission (cohort 1) N (%)	Model predictions (cohort 1) %	Table 19 of submission (cohort 2)	Model predictions (cohort 2)
Death	Not reported	1.04%	Not reported	1.22%
Source: Response to clarification letter[10]				

AlloSCT rates conditional on response

The probabilities of having an alloSCT conditional on response status were elicited through two clinician surveys, one performed by the company (referred to here as the MSD survey) and one performed within the course of previous TA462 (referred to here as the BMS survey).[6] The company stated that it was necessary to use the intermediate step of applying a probability of alloSCT based on response status, because it was not appropriate to use the KEYNOTE-087 study data on time to alloSCT directly. This was justified by a smaller proportion of patients ([redacted]) in the KEYNOTE-087 study receiving alloSCT,[1] compared with UK practice, although no data for UK practice, apart from the survey data, were presented. The KEYNOTE-087 study data on alloSCT were therefore not used directly to inform the present model. The company also stated that [redacted] of UK patients in cohort 1 and [redacted] in cohort 2 received alloSCT (CS p160)[1] but corrected this in their response to the clarification letter to be [redacted] for cohort 1 and [redacted] for cohort 2.[10]

The MSD clinician survey drew on opinions from 16 clinicians from the UK who were asked the proportion of patients they would expect to proceed to alloSCT conditional on response to treatment, which could be CR, PR, SD or PD. The results of this survey were combined with the results from the BMS survey by taking a simple, unweighted average of the means (Table 5.7). The company stated that it disregarded clinicians' responses indicating that some patients in the progressed disease health state could be eligible for alloSCT (a mean of [redacted] according to the company's slides in REF pack 1)[39] following further discussions with UK clinicians on this topic and stating that this was not thought to be standard UK clinical practice. However, in the KEYNOTE-087 study, [redacted] patients were in the progressed disease health state when they received alloSCT, albeit none of them from the UK.

For cohort 2, the same rates were assumed as for cohort 1, but some clinicians suggested that alloSCT rates in that population might be even higher than in cohort 1 due to the unmet need in this population.

The same alloSCT probabilities conditional on response status were adopted for both pembrolizumab and SoC.

Table 5.7: AlloSCT rates conditional on response

	MSD Mean[40]	**Alternative Mean**	**Overall Mean**	**SE**
CR	56.79%	[redacted]	[redacted]	[redacted]
PR	43.93%	[redacted]	[redacted]	[redacted]
SD	18.36%	[redacted]	[redacted]	[redacted]
Source: CS Table 64				

ERG comment: The ERG's comments include that (a) patients with a non-evaluable response status being considered to have SD inflates the proportion of patients in this health state, (b) the omission of

the survey result that patients with progressed disease could still be eligible for alloSCT is non-conservative, and (c) the combination of the MSD and BMS survey results may introduce bias.

(a) The proportions in the SD state in both cohorts in the model are significantly larger than those observed in the KEYNOTE-087 study, as reported in Table 19 of the CS. This is a result of patients with non-evaluable response status being moved into the SD state. In response to the clarification letter,[10] the company provided an overview of model predictions of response status compared with the KEYNOTE-087 data (shown in Table 5.8). It can be seen that the model may over-predict the proportions in the SD state, but this is likely a conservative assumption.

Table 5.8: Comparison of response status in model and KEYNOTE-087

Status	Table 19 of submission (cohort 1) N (%)	Model predictions (cohort 1) %	Table 19 of submission (cohort 2)	Model predictions (cohort 2)
Complete response	[redacted]	15.94%	[redacted]	8.6%
Partial response	[redacted]	42.0%	[redacted]	43.2%
Stable disease	[redacted]	36.9% *(~27.5%+8.7%)*	[redacted]	38.9% (~18.5% + 8.6%)
Non-evaluable	[redacted]	Not reported (combined in stable disease)	[redacted]	Not reported (combined in stable disease)
Progressed disease	[redacted]	4.10%	[redacted]	7.9%
Death	Not reported	1.04%	Not reported	1.22%
Source: Response to clarification letter				

(b) Patients with PD were assumed to not get alloSCT, despite the MSD survey results indicating otherwise ([redacted] of patients with PD would get alloSCT). In response to the clarification letter,[10] the company explained that based on feedback from UK clinicians, it is not UK standard practice that patients in the PD state would receive alloSCT. The company furthermore provided data from KEYNOTE-087, where none of the [redacted] UK patients who underwent alloSCT (in cohorts 1 and 2) had PD prior to alloSCT, but [redacted] was non-evaluable. The ERG was concerned about this argumentation. First, the MSD survey was performed with UK expert clinicians only and it was not explained why the company considered it appropriate that discussions with a number of UK clinicians overrode the survey results. Furthermore, the [redacted] UK patients from KEYNOTE-087 who underwent alloSCT are likely too few to be representative. The company did provide a scenario analysis enabling alloSCT in PD patients, which resulted in increases in the ICER. The ERG adopted the MSD survey results with probabilities for alloSCT in PD patients in its base-case.

(c) The ERG considers the combination of the MSD and BMS surveys as problematic: for one, the company stated that the TA462 committee had deemed the Cheah et al[7] estimates of 66% of responders receiving alloSCT as too high for the UK. However, it can be seen from Table 5.7 that estimated proportions of patients receiving alloSCT from the MSD survey were lower than those from the BMS survey. Hence, when both surveys are combined, according to Table 5.7, the alloSCT rates used in the CS for the CR status are even higher than the Cheah et al[7] estimates, and even when the mean for PR and CR is taken, the resulting alloSCT rates for responders ([redacted]) are not significantly lower than those in Cheah et al.[7] Given that the company's estimation of alloSCT rates based on their own MSD survey would have resulted in lower alloSCT rates for partial and complete responders ([redacted]) compared to

Cheah et al[7], the ERG considers the use of the MSD survey data alone to be more in line with the TA462 committee preferences. The committee conclusion on the BMS survey also entailed the following comment: "the committee also heard that recent NHS referrals for allogeneic stem cell transplant were lower than those reported in the [BMS] survey." It is therefore not clear to the ERG why the company opted to combine the MSD and BMS surveys. This is of particular concern given that the company accepts that *"it is possible for both surveys to have included the same clinical experts"*.[10] It is the ERG's view that bias induced by double-counting of certain experts' opinions cannot be ruled out. The company, in response to the clarification letter, provided a scenario analysis using alloSCT rates from the MSD survey only, which indicated that the ICERs increased. For reasons mentioned above, the ERG preferred to use the MSD survey only, instead of combining them with the BMS survey, in its base-case.

5.2.6.3 Post-12 weeks: patients not receiving alloSCT – PFS and mortality

PFS post-12 weeks

PFS post-12 weeks was estimated using only the observed data from KEYNOTE-087 beginning in week 12 to end of study follow-up.[10] The company stated that alloSCT events were not censored from the survival analysis of KEYNOTE-087 because it was not possible to censor them from the Cheah study either. The exponential distribution was used to estimate PFS post-12 weeks in cohorts 1 and 2. For cohort 1, this represented the model with the best statistical fit. For cohort 2, the exponential did not make the best statistical fit (the generalised gamma did) and it over-estimated PFS at the end of follow-up, however the company argued that the last drop in PFS was not considered particularly informative given the small patient numbers at risk (n=3) (see Figure 5.2). The Gompertz was considered in a scenario analysis.

The company assumed that the post-12 week HR and pre-12 week HR for PFS were equal for cohorts 1 and 2 estimated at [redacted] and [redacted], respectively, for the entire study period. The company justified this by stating that *"a large number of progression events occurred during the first 12 weeks of the SoC study."* and that *"Therefore, it was not possible to estimate a HR between the two treatments after 12 weeks."* (CS p 151) [1] The company concluded that *"a PFS HR from week 12 to end of follow-up could not be estimated given the low number of events post week 12 observed in Cheah"* (CS p 141). [1] In response to the clarification letter,[10] the company furthermore clarified that the HR for pembrolizumab versus SoC was derived using the entire follow-up period.

Figure 5.2: PFS (BIRC) cohort 2 from week 12 extrapolations

CS Figure 21[1]

Mortality pre-progression post-12 weeks

Mortality in the pre-progression health state post-12 weeks when patients did not receive alloSCT was assumed to be equal to general mortality estimates derived from UK life tables for both pembrolizumab and SoC.[41]

Mortality post-progression post-12 weeks

Because the number of patients was considered too small to support robust analysis of post-progression survival, the company used Cheah et al (2016)[7] to estimate post-progression mortality for both pembrolizumab and SoC. The weekly transition probabilities were obtained by converting median OS in Cheah assuming a constant hazard rate based on the exponential distribution. The obtained transition probability of 0.63% (per week) was replaced by background mortality when general mortality estimates obtained from UK life tables exceeded this probability.

The company implicitly assumed a HR = 1 for estimating mortality in the pre- and post-progression health states by using general mortality estimates for pembrolizumab and SoC for the pre-progression health state, and Cheah et al.[7] to inform transition probabilities from the post-progression state to the dead state.

ERG comment: The ERG's concerns relate to (a) uncertainty around extrapolating PFS post-12 weeks, (b) the assumption that patients in the pre-progression health state can only die from all-cause mortality, (c) the assumption that pre- and post-12 week HRs for PFS were equal, and (d) the immature OS data from KEYNOTE-087.

(a) For post-12 week PFS in cohort 1, the choice of the exponential distribution was based on best statistical fit. The Gompertz distribution had a statistical fit within two AIC points and the ERG therefore considered it informative to explore the use of this model in scenario analysis.

In cohort 2, the choice of the exponential distribution for post-12 week PFS is unclear. The generalised gamma distribution has the best statistical fit, followed by the Gompertz and exponential distributions (based on AIC and BIC respectively). Despite this, the company chose the exponential distribution, with the rationale that the small patient numbers at risk at the end of follow-up make the last drop less informative. The ERG considers clinical plausibility important but remains unconvinced that there was sufficient justification for ruling out the generalised gamma. Clinical expert opinion should have been used to validate this assumption. The ERG considers that the model with the best fit (generalised gamma) and second best fit (Gompertz) should be explored in scenario analysis. Results show that the choice of post-12 week PFS model in cohort 2 is very influential and that the company's choice of exponential favoured pembrolizumab.

(b) In the pre-progression health state, patients are assumed to die only from all-cause mortality. There was no indication provided for why this was clinically plausible and the ERG is uncertain about the impact of this assumption on model outcomes.

(c) The ERG considers the assumption that post-12 week HR and pre-12 week HR for PFS were equal to be questionable. The use of a constant HR lacks face validity because different parametric models pre- (log-logistic and generalised gamma in cohorts 1 and 2) and post-12 weeks (exponential in cohorts 1 and 2) were used. The company, upon request, provided a scenario analysis using a HR=1 for the post-12 week period, which increased the ICERs significantly. This should be viewed as a worst-case scenario. Given that the HRs were estimated based on the entire study data, the ERG maintains the HRs used by the company in its base-case.

(d) OS data for the entire study population of KEYNOTE-087 was deemed by the company to be too immature to provide robust extrapolations of survival.[7] Upon request, the company provided scenario analysis with post-12 weeks post-progression survival estimated based on KEYNOTE-087 instead of Cheah et al.[7], which decreased the ICERs. Because of the small number of post-progression events in KEYNOTE-087 ([redacted] in cohort 1, [redacted] in cohort 2),[10] the ERG agrees that these data are too immature to be used in the present analysis.

5.2.6.4 Post-12 weeks: patients receiving alloSCT - OS

OS estimates were obtained from a UK study consisting of 13 patients with classical Hodgkin Lymphoma who received alloSCT after three previous therapies (Lafferty et al, 2017).[21] The company stated that this was in line with previous TA462 on nivolumab for treating relapsed or refractory classical Hodgkin Lymphoma.[6] The company attempted to digitise the KM provided in Appendix 17 of the CS,[27] but resorted to manually adjusting the data because the unknown rate of censoring in the tail of the curve and the limited number of events prevented the company from reproducing patient level data. However, the company used the point estimates and AIC/BIC from TA462, and only used their own digitised version of the Lafferty KM data for the PSA.

The Weibull distribution was used to extrapolate OS beyond the available Lafferty et al data. The Weibull did not have the best statistical fit and, in fact, only came fifth according to the AIC/BIC criteria. However, the company justified their choice by stating that (1) the generalised gamma predicted an infinite hazard beyond 150 months and therefore had to be adjusted, thereby likely under-estimating the survival benefit expected in this population, (2) AIC/BIC scores were relatively similar (for example, AIC score of Weibull <3 points away from the AIC of the generalised gamma, which ranked

first in terms of AIC/BIC, (3) the ERG in TA462 considered the use of log-normal and Weibull models as more clinically plausible as they did not predict infinite survival, and (4) the company considered the Weibull more conservative than the lognormal. The lognormal was explored in the company's scenario analysis. Model predictions of the different models are shown in Table 5.9.

Table 5.9: Summary of the survival models (OS after alloSCT adjusted for all-cause mortality)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
Median (months)	53.13	64.62	266.78	58.41	61.86	87.39	--
Mean (months)	76.77	163.07	237.71	172.88	177.21	213.93	--
% at 1 year	85.73%	71.68%	63.33%	69.74%	70.01%	65.28%	64.17%
% at 2 years	73.39%	63.78%	55.90%	61.55%	61.93%	59.48%	53.47%
% at 5 years	53.77%	54.50%	53.58%	52.68%	53.33%	54.21%	53.47%
% at 10 years	21.09%	40.56%	52.90%	40.79%	41.77%	47.95%	--
% at 15 years	9.67%	34.13%	52.08%	35.78%	36.83%	45.43%	--
% at 20 years	4.43%	29.61%	50.80%	32.40%	33.45%	43.82%	--
% at 30 years	0.93%	23.46%	45.95%	27.88%	28.84%	39.63%	--
% at 40 years	0.20%	17.64%	34.77%	21.10%	21.83%	29.99%	--

Source: CS Table 69[1]

ERG comment: The ERG has concerns about (a) the appropriateness of using Lafferty et al.[21] for estimating post-alloSCT OS and (b) that the company over-estimates OS in post-alloSCT patients.

(a) The ERG questioned the appropriateness of using Lafferty et al[21] for post-alloSCT survival, given that in KEYNOTE-087, [redacted] patients had an alloSCT compared with the 13 patients in Lafferty et al[21]. In response to the clarification letter,[10] the company explained that the KEYNOTE-087 study did not include the subsequent investigation of patients treated with pembrolizumab who were treated with a stem cell transplant. Furthermore, the company argued that OS data for the entire study population of KEYNOTE-087 were deemed to be too immature to provide robust extrapolations of survival and highlighted that Lafferty et al[21] was also used to inform TA462. Because Lafferty et al[21] is a very small study with questionable generalisability to the UK setting (see Section 4.2.3), its use means that there is substantial uncertainty around post-alloSCT survival, and alternative evidence was not explored.

(b) According to the company's Figure 3 in Appendix 17 of the CS,[27] (Figure 5.3) post-alloSCT survival is likely over-estimated. From this figure it appears that the company assumed no censoring after the last event until the end of the 5-year period. This results in an over-estimation of OS, as can be seen from the fitted curves that follow the plateau between 21 months and 5 years closely. It is unlikely that this plateau is a reflection of OS in patients post-alloSCT and the ERG considers it more likely that censoring occurred before the end of this 5-year period. The ERG acknowledges that there is uncertainty about the better approach, but notes that the company chose the approach that favoured pembrolizumab the most. The ERG therefore used the KM estimates from Figure 5.3 to reconstruct individual patient level data, allowing for censoring after the last event and before the end of the follow-up period, and used this in ERG scenario analysis, showing that the company's analysis significantly favoured pembrolizumab. The ERG's and the company's fitted curves are shown in Figure 5.4. As can be seen,

the ERG's approach gives less weight to the plateau in the tail of the Kaplan Meier curve than the company's approach.

Figure 5.3: KM estimates from Lafferty et al, as presented in company appendix 17

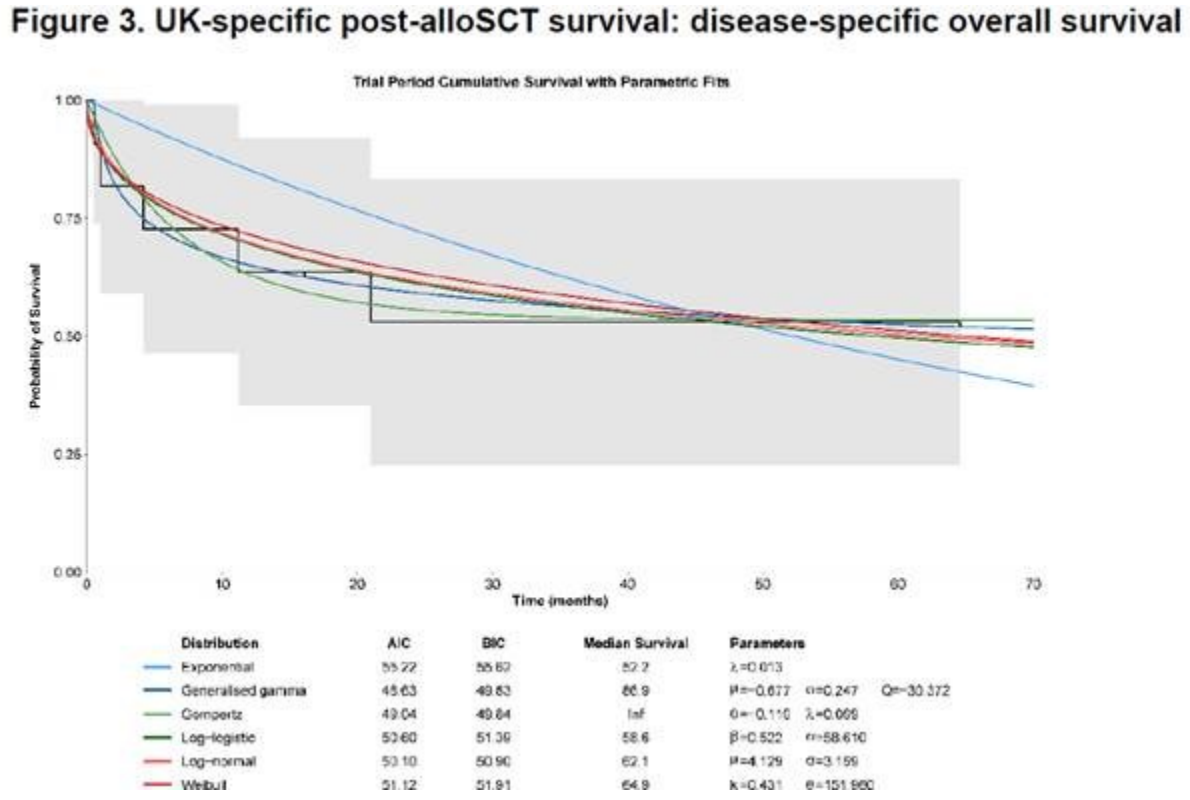
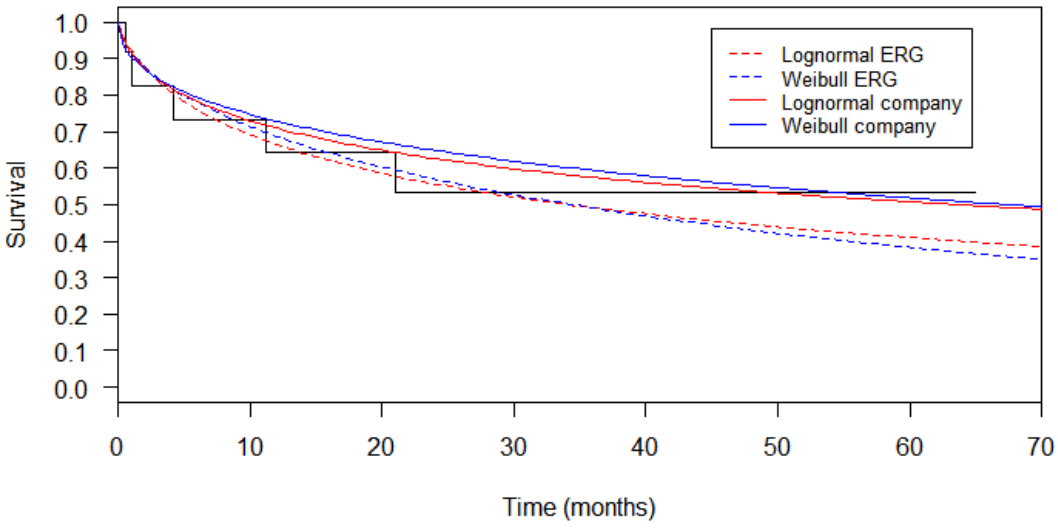


Figure 5.4 ERG's versus company's approach to estimating post-alloSCT OS based on Lafferty et al



5.2.6.6 Time to treatment discontinuation

Time to treatment discontinuation (TTD) pre-12 weeks
The company used PFS as a proxy for TTD for the pre-12 week period. No justification was provided.

TTD post-12 weeks
Treatment is discontinued for patients receiving alloSCT. For patients not receiving alloSCT, the company estimated time to treatment discontinuation for the post-12 week period using the TTD data available from KEYNOTE-087. PFS was not deemed an appropriate proxy because, on average, patients discontinued treatment before they progressed. The company postulated that this may be due to safety and tolerability and the impact of the design of KEYNOTE-087, which allowed study investigators to discontinue therapy if complete response had been achieved after at least six months of treatment. TTD is furthermore capped at 24 months in the company's model. The company justified this stating that this was in line with the stopping rule employed within the KEYNOTE-087 study.

For both cohorts 1 and 2, the exponential distribution was chosen, as it was the model exhibiting the best statistical fit and maintained consistency with the base-case PFS distribution.

For SoC, PFS was used as a proxy for TTD and this was justified by the lack of treatment discontinuation data from Cheah et al (2016).[7]

ERG comment: The ERG's concerns relate to (a) the inconsistency of using PFS as a proxy to TTD for the pre-12 weeks period and the comparator but not for TTD post-12 weeks, and (b) the capping of time to treatment discontinuation at 24 months in the model.

(a) For the pre-12 week period, PFS was used as a proxy to TTD. The company did not provide justification for this assumption. This means that the estimation of TTD is inconsistent between the pre- and post-12 weeks periods, and indeed with the comparator, for which PFS was used as a proxy.

(b) The company's assumption that treatment duration is capped at 24 months is not in line with the marketing authorisation. Upon request, the company provided a scenario analysis of continued treatment with pembrolizumab after 24 months, which showed that ICERs increased substantially for both cohorts.[10] This is possibly a pessimistic scenario, because effectiveness was based on KEYNOTE-087, in which the maximum treatment duration was 24 months. However, the ERG wishes to point out that the company's base-case might under-estimate the cost incurred with the use of pembrolizumab when a 24-months stopping rule is not enforced in clinical practice.

5.2.7 Adverse events

The company decided, in order to reflect best clinical practice, to incorporate the AEs that were included in the previous Hodgkin Lymphoma appraisal (TA462)[6]. Table 5.10 presents the grade 3+ AEs with an incidence of ≥0% in any study arm, that were incorporated as a one-off cost and disutility into the first cycle of the cost effectiveness model. The company assumed patients remaining on treatment beyond the first year to tolerate treatment well and therefore not to experience severe AEs. The company further assumed that investigational agents do not cause any AEs.

Table 5.10: Adverse event rates incorporated in the cost effectiveness model

Adverse Event	Pembrolizumab (cohort 1)	Pembrolizumab (cohort 2)	Chemotherapy	Bendamustine	SoC*
Anaemia	[redacted]	[redacted]	16.59%	13.89%	16.29%
Diarrhoea	[redacted]	[redacted]	6.25%	0.00%	5.88%

Dyspnoea	[redacted]	[redacted]	8.33%	0.00%	6.67%
Fatigue	[redacted]	[redacted]	10.00%	2.78%	10.00%
Leukopenia	[redacted]	[redacted]	55.00%	0.00%	54.84%
Nausea	[redacted]	[redacted]	4.95%	2.78%	4.71%
Neutropenia	[redacted]	[redacted]	45.07%	8.33%	43.56%
Pyrexia	[redacted]	[redacted]	0.00%	2.78%	0.00%
Thrombocytopenia	[redacted]	[redacted]	37.60	19.44%	37.13%
Vomiting	[redacted]	[redacted]	2.65%	0.00%	3.08%
Source: calculations performed by the ERG, based on adverse events incidence tables from cost effectiveness model provided by the company *For SoC AE calculation, assumption was made (Weighted average of chemotherapy, bendamustine and investigational agents. See model safety tab)					

ERG comment: The ERG identified an error in the calculation of SoC adverse events incidence.

AE incidence for SoC, based on the weighted average of chemotherapy (38.46%), bendamustine (18.46%) and investigational agents (43.08%), was incorrectly calculated. Although it was assumed that investigational agents did not have AEs and therefore do not influence the number of events, the proportion of patients that received investigational agents should be included in the calculation of the sample size (N). By not doing this, the company over-estimated the relative SoC AE incidence. This is likely a favourable assumption for pembrolizumab, but is unlikely to be influential.

5.2.8 Health-related quality of life

HRQoL was measured in KEYNOTE-087. More specifically, EQ-5D-3L data were collected at treatment cycles 1-5 (i.e. every three weeks) and every 12 weeks up to 30 days post treatment discontinuation or until disease progression. Consistent with the NICE reference case, the UK social tariff[42] was used to obtain health state utility values from the responses on the EQ-5D-3L. Although the SLR also identified two relevant HRQoL studies, HRQoL data from KEYNOTE-087 were preferred by the company. It was unclear whether this was because the HRQoL studies identified in the SLR were inconsistent with the NICE reference case[43] or did not report utilities stratified by overall response.[44]

The company calculated utility values (Table 5.11) stratified by overall response (i.e. separately for patients with CR, PR and SD). However, this post hoc utility calculation was based on observations from week 12 in the KEYNOTE-087 trial only (i.e. ignoring observations at other time points). These utility scores were multiplied by the response rates from KEYNOTE-087 and Cheah et al, (2016)[7] to obtain the progression free health state utility values for pembrolizumab ([redacted] and [redacted] for cohort 1 and 2) and SoC ([redacted]) respectively (Table 5.11).

Similarly, response rates from Lafferty et al,[21] an abstract retrospectively reporting on single centre experiences with alloSCT in patients with Hodgkin Lymphoma, were used to calculate the post-alloSCT utility. Combining these response rates with the 12 week utilities (stratified by response) from KEYNOTE-087 resulted in a post-alloSCT utility of 0.865. To account for the possibility of acute graft versus host disease after alloSCT, a disutility of 0.15[45] is applied to 61.5%[21] of the patients for the first

14 weeks post-alloSCT. This resulted in a post-alloSCT utility of 0.773 for the first 14 weeks which was assumed to increase to 0.865 afterwards.

Table 5.11: Utility scores for the progression free (treatment dependent) and post-alloSCT disease health states

	Utility (12 week observations only)	Pembrolizumab response rates (cohort 1)	Pembrolizumab response rates (cohort 2)	SoC response rates	Post-alloSCT response rates
	KEYNOTE-087	KEYNOTE-087	KEYNOTE-087	Cheah et al[7]	Lafferty et al[21]
Total N	[redacted]	[redacted]	[redacted]	[redacted]	10
CR	[redacted]	[redacted]	[redacted]	[redacted]	70.0%
PR	[redacted]	[redacted]	[redacted]	[redacted]	30.0%
SD	[redacted]	[redacted]	[redacted]	[redacted]	0.0%
Utility[a]	[redacted]	[redacted]	[redacted]	[redacted]	0.865

Source: Economic model submitted by the company and CS Table 75
CR = complete response; PR = partial response; SD = stable disease;
[a]Utility was calculated by combining the Utility scores stratified by response and the response rates

The company did not use the PD utility score (of [redacted]) from KEYNOTE-087 arguing that this utility "is not predictive of a meaningful decrement in QoL", due to it being estimated based on 12 week observations only. Therefore, the company opted to use a utility decrement (of 0.33) calculated by subtracting the SD and PD utilities from Swinburn et al.[43] This resulted in a PD utility of [redacted].

Additionally, the company applied age related utility decrements, derived from UK population norms, in all health states (see CS Table 82). This was conditional on the starting age in the model (34 and 40 years for cohort 1 and cohort 2 respectively).

Finally, the company considered the impact of grade 3+ adverse events (see Section 5.2.7) on HRQoL. Given the absence of disutilities in relapsed or refractory Hodgkin Lymphoma, disutilities were identified in oncology and myocardial infarction (see CS Table 77 for a summary of sources). In case multiple sources were available an average was calculated. The disutilities and adverse event durations from the various adverse events are reported in CS Table 78. Table 5.12 below provides an overview of the calculated disutilities and the assumed duration of the AE. Multiplying the duration, the disutility and the occurrence of adverse events (see section 5.2.7) resulted in one-off disutilities of [redacted], [redacted] for pembrolizumab (separately for cohort 1 and cohort 2) and 0.0080 for SoC. These one-off disutilities were incorporated in the first cycle of the model.

Table 5.12: Adverse event disutilities

Adverse event (CTCAE grade 3+)	Disutility (per year with adverse event)	Duration (days)	Disutility (per occurrence of adverse event)
Anaemia	-0.0900	14.8	-0.0036
Diarrhoea	-0.1392	5.5	-0.0021
Dyspnoea	-0.0481	12.7	-0.0017
Fatigue	-0.1502	25.5	-0.0105
Leukopenia	-0.1264	12.1	-0.0042
Nausea	-0.1517	11.0	-0.0046
Neutropenia	-0.1264	12.3	-0.0042
Pyrexia	-0.1100	12.3	-0.0037
Thrombocytopenia	-0.1080	15.9	-0.0047
Vomiting	-0.1395	5.3	-0.0020

Source: Economic model submitted by the company and CS Table 83[1]

ERG comment: The ERG notes the following issues regarding the HRQoL data used by the company: (a) HRQoL data used by the company is restricted to observations from week 12 only, (b) using a decrement for progressive disease that is not from KEYNOTE-087, (c) progression free utility benefit for pembrolizumab maintained for patients without alloSCT, (d) sources for post-alloSCT HRQoL, (e) HRQoL consequences of disease progression post-alloSCT are not (explicitly) incorporated, (f) one technical error and one inconsistency in the calculation of the HRQoL.

(a) The company restricted the HRQoL data, used in its base-case, to KEYNOTE-087 observations from week 12 only. In response to clarification question B15, the company provided the results of mixed effects model analyses incorporating all observed EQ-5D data from KEYNOTE-087 (Table 5.13). Unfortunately, no diagnostics or goodness of fit statistics were provided by the company. Nevertheless, to utilise all available KEYNOTE-087 data, the ERG prefers to use utility scores generated by this mixed effects model. It is, however, notable that the coefficient for "PR versus CR" is positive, i.e. indicating a higher utility for PR than for CR. This lacks face validity, hence, the ERG decided to set this coefficient (0.01453) to zero. This resulted in a utility of [redacted] for both CR and PR while the estimated utility value for SD is [redacted]. Combining this with the observed response status, this resulted in PF utility values of [redacted] for pembrolizumab (cohorts 1 and 2) [redacted] for SoC. This SoC PF utility (of [redacted]) is more consistent with the SoC PF utility of 0.76 reported in TA462[6] than the SoC PF utility used in the CS base-case ([redacted]). Additionally, the PD utility changed to [redacted] while the post-alloSCT utility changed to 0.725 for the first 14 weeks and to 0.818 for after the first 14 weeks.

Table 5.13: Utilities estimated from mixed effects model using all observed EQ-5D data from KEYNOTE-087

Covariates	Estimated effect	Standard error
Intercept (reference = CR)	[redacted]	[redacted]
PR versus CR	[redacted]	[redacted]
SD versus CR	[redacted]	[redacted]
PD versus CR	[redacted]	[redacted]

Source: response to clarification question B15
Note: not marked as CiC in the clarification response
CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease

(b) The company does not use the estimated PD utility from KEYNOTE-087 in its base-case. This was justified by stating that KEYNOTE-087 only contains observations shortly after progression and which might not capture the long-term utility decrement due to progression. Therefore, the company estimates the PD utility using the SD utility from KEYNOTE-087 and a utility decrement of 0.33 from Swinburn et al.[43] The ERG was not convinced that using this utility decrement from Swinburn et al.[43], over the PD utility estimated from KEYNOTE-087 is appropriate, given the company provided no evidence indicating a long-term impact of progression consistent with this utility decrement (of 0.33 versus SD). Additionally, the ERG in TA462 criticised the utilities from Swinburn et al.[43] by stating *"we suggest that the results from Swinburn and colleagues are outliers and may not be realistic. The Swinburn study used TTO methodology using estimates from the general public and it may be that their perception of the disease is not consistent with EQ-5D valuation."* This quote also highlights that the utilities from Swinburn et al.[43] deviate from the NICE reference case (as it is not consistent with EQ-5D valuation). Therefore, consistent with the NICE reference, the ERG's approach in TA462 (which was ultimately accepted by the committee), for the ERG base-case is to use HRQoL data from the pivotal trial (KEYNOTE-087) to estimate the PD utility (estimated PD utility based on mixed effects model is [redacted]). This PD utility estimate is more in line with the PD utility, estimated based on CheckMate 205, that was preferred by the ERG and accepted by the committee in TA462.[6]

(c) Based on response status (i.e. proportion of patients with CR, PR and SD), treatment specific PFS utilities are estimated and used throughout the model time horizon for the PFS health state. However, it is inconsistent to use the response status, combined for both groups of patients that undergo alloSCT and those who do not, to estimate the utilities for patients that do not undergo alloSCT. Particularly given that the response status for patients that undergo alloSCT is likely better than for patients who do not undergo alloSCT. Therefore, the ERG recalculated the post-12 week PFS utilities based on the response status of patients who did not undergo alloSCT. This resulted in utility values of [redacted] for pembrolizumab (both cohorts) and [redacted] for SoC. Based on the mixed model these utilities would be lower for pembrolizumab ([redacted] and [redacted] for cohort 1 and 2) and SoC ([redacted]).

(d) The company uses a disutility from Kurosawa et al.[45] (applied to 61.5% of the patients) to account for the possibility of acute graft versus host disease after alloSCT. This disutility is applied to the post-alloSCT utility estimated based on the KEYNOTE-087 estimates. In clarification question B18, the ERG questions why a disutility only is obtained from Kurosawa et al.[45]. The company stated that it was believed to be inappropriate to also use the utility estimate from Kurosawa et al.[45] due to the differences between populations in Kurosawa et al.[45] and KEYNOTE-087. The ERG, however, believes it is inappropriate to use KEYNOTE-087 utility estimates, including only one post-alloSCT observation (response to clarification question B15b), to estimate post-alloSCT utility values. Although the ERG recognises the differences between populations (i.e. in Kurosawa et al.[45] and KEYNOTE-087), given that Kurosawa et al.[45] is the only identified study to provide post-alloSCT preference-based (e.g. EQ-5D) utility measures (confirmed by the company in response to clarification question B18), the ERG prefers to use Kurosawa et al.[45] to obtain post-alloSCT utility values in the model. This resulted in a post-alloSCT utility of 0.708 for the first 14 weeks which was assumed to increase to 0.800 afterwards (these values were 0.773 and 0.865 in the CS base-case).

(e) Due to the lack of a post-alloSCT progression health state (see also ERG critique in section 5.2.2), it is questionable whether the impact of progression on HRQoL post-alloSCT is captured. Therefore, the ERG performed a scenario analysis to explore the impact of this assumption.

(f) Finally, the ERG identified a technical error in the calculation of the AE disutility (in the model the AE duration is divided, to convert from day to year, by 365.25 twice instead of once) as well as an inconsistency in the proportion of responders used to calculate PF utility estimates (see difference in response between CS Table 80 and the Table provided in response to clarification question B5). In the ERG base-case, the technical error was corrected and the number reported in response to clarification question B5 (updated version of CS Table 62, see Table 5.5 of the ERG report) is used to estimate PF utilities.

Table 5.14 below provides an overview of the utilities used in the ERG base-case (combining all abovementioned adjustments).

Table 5.14: Utilities used in the CS and ERG base-case

Health state		CS base-case utility	ERG base-case utility[a]
PF first 12 weeks	pembrolizumab cohort 1	[redacted]	[redacted]
	pembrolizumab cohort 2	[redacted]	[redacted]
	SoC	[redacted]	[redacted]
PF after first 12 weeks (no alloSCT)[b]	pembrolizumab cohort 1	[redacted]	[redacted]
	pembrolizumab cohort 2	[redacted]	[redacted]
	SoC	[redacted]	[redacted]
PD	treatment independent	[redacted]	[redacted]

Health state		CS base-case utility	ERG base-case utility[a]
Post-alloSCT first 14 weeks	treatment independent	0.773	0.708
Post-alloSCT after first 14 weeks	treatment independent	0.865	0.800

PR = progression free; PD = progressive disease;
[a]Standard error calculated by multiplying the estimated utility by 0.1 (consistent with the company's approach)
[b]The estimated PF utilities after the first 12 weeks (no alloSCT) would be [redacted] for pembrolizumab (cohort 1 and 2) and SoC respectively, when using the MSD survey only to estimate the proportion of patients receiving alloSCT conditional on response status (see section 5.2.6.2). These values were used in the final ERG base-case.

5.2.9 Resources and costs

5.2.9.1 Drug acquisition and administration costs

The electronic market information tool (eMit)[37] was used to acquire drug acquisition cost of pembrolizumab and components of SoC. When these were unavailable, costs from the British National Formulary[46] were used. Administration costs were obtained from the NHS reference costs[47] (see Table 5.15).

Pembrolizumab

The list price of 200 mg pembrolizumab was £5,260 (derived from the cost of 2 x 100 mg vials at £2,630 each per patient). Through a Commercial Access Agreement (CAA), [redacted]. As established previously in TA357[48] and TA428 [49], the NHS reference cost code SB 12Z [47] was used as administration cost, thereby adding £236.19 per 21 day cycle.

Standard of care

Consisting of chemotherapies (38.5%, each of the 12 treatments accounting for 3.2%), treatment with investigational agents (43.1%) and bendamustine (18.5%), drug acquisition costs for SoC varied by treatment agent. Acquisition and administration costs of investigational agents were assumed to be £0, for other components of SoC these costs are described in Table 5.15. For dosages/m^2, the number of vials required per administration was calculated based on a BSA of 1.85m^2 (SD 0.024). Upon request the company clarified that CS Table 56[1] contained an incorrect BSA but that the correct number was used in the model. For each component of SoC, the model assumed vial wastage and calculated the vial combination resulting in the lowest possible price for the required dosage. The treatment costs of SoC per seven day cycle (see Table 5.15) consist of the treatment costs of all SoC components. These were calculated by combining the calculated drug acquisition cost per cycle with the administration costs, adjusting these costs to the seven day timeframe and the proportion of patients treated within the SoC arm.

Best supportive care

BSC consisted of several subsequent treatments that are described in Table 5.16, which were selected based on the approach taken in TA462.[6] Acquisition costs and administration costs combined with the proportions of patients treated with each component of BSC resulted in a one-off cost of £4,848.22. In line with assumptions made in TA462,[6] palliative care and clinical trial treatment were assumed to have no costs.

Table 5.15: Treatment costs

Regimen	Acquisition cost/cycle	Administration costs/cycle	Cycle length (days)	Maximum number of cycles	Proportion of treatment (%)
Pembrolizumab	[redacted]	£236.19 [a]	21	35	100
Standard of care					
ICE	£1,230.82	£711.23 [b]	14	3	3.2
IVE	£2,183.65	£1,039.33 [c]	21	3	3.2
MINE	£1,209.02	£1,039.33 [c]	28	2	3.2
IVOx	£1,132.46	£1,039.33 [c]	21	3	3.2
IGEV	£2,109.48	£1,367.43 [d]	21	4	3.2
GEM-P	£100.86	£711.23 [b]	28	3	3.2
GDP	£93.06	£383.13 [e]	21	2	3.2
GVD	£1,491.60	£711.23 [b]	21	2	3.2
ESHAP	£63.32	£1,367.43 [d]	28	4	3.2
ASHAP	£68.73	£1,367.43 [d]	28	3	3.2
DHAP	£76.39	£383.13 [e]	21	2	3.2
DHAOx	£89.69	£383.13 [e]	21	4	3.2
Bendamustine	£123.30	£383.13 [e]	28	6	18.5

Source: CS Table 91, Table 92[1]
[a] Deliver Simple Parenteral Chemotherapy at First Attendance
[b] Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
[c] Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
[d] Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle
[e] Delivering complex chemotherapy at first attendance

Table 5.16: BSC

Therapy	Distribution of patients across therapies (%)	Cycle length	Number of cycles	Acquisition costs/cycle	Administration costs/cycle
Gemcitabine monotherapy (administered over 4 weeks)	8.33	28 days	4.0	£47.76	£236.19
RVIG	16.67	21 days	4.5	£3,299.29	£1,367.43
DHAP	11.67	21 days	6.0	£76.39	£383.13
CHOP	1.67	21 days	6.0	£32.45	£383.13
IVAC	3.33	21 days	3.5	£1,832.00	£1,695.53
Weekly therapy (PMitCEBO)	8.33	14 days	7.0	£109.11	£711.23
Palliative care	46.67				
Clinical trial treatment	3.33				

Source: CS Table 93, Table 94, Table 97,[1] Model

CHOP = cyclophosphamide, doxorubicin, prednisolone, vincristine; DHAP = dexamethasone, cytarabine, cisplatin; IVAC = cytrabine, etoposide, ifosfamide, mesna; PMitCEBO = bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine; RVIG = gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine

ERG comments: The ERG identified the following inconsistencies and assumptions lacking justification: (a) potential over-estimation of SoC costs due to the assumed mix of chemotherapy regimens within SoC, (b) the lack of missed doses, and (c) the number of cycles used for the components of BSC.

(a) The assumption that all chemotherapy agents contribute equally to the mix of SoC is not justified by the company. Responding to clarification question B19 [10], the company explains that *"There is a paucity of evidence on the preferred or standard mix of chemotherapy regimens given to patients in UK clinical practice"*, and an approach previously accepted was used. Given the extensive efforts taken by the company to interview clinical experts on alloSCT uptake, it can be questioned why the comparator treatment mix was not a topic discussed with the clinical experts. In TA462 it is stated that *"Clinical advice to the ERG suggests that gemcitabine regimens such as GDP (gemcitabine, dexamethasone, cisplatin) are commonly used in this patient population in the UK but platinum-containing regimens such as ESHAP (etoposide, methylprednisolone, cytarabine, cisplatin) and DHAP (dexamethasone, cytarabine, cisplatin) are also in common use.*"[6] Given the regimens mentioned are of lower price than other chemotherapy regimens, the ERG wishes to point out that the company has likely overestimated costs of SoC, an assumption that favours pembrolizumab.

(b) The company states that due to a lack of information on missed doses in the SoC arm, missed doses were not incorporated for pembrolizumab or SoC. The ERG wishes to highlight that the incremental cost of pembrolizumab versus SoC could be biased, even though the effect is likely to be small.

(c) The company assumed treatment durations for components of BSC to be shorter than the maximum number of cycles. The ERG recognises this assumption to be conservative and possibly reflective of treatment intensity after repeated progression. However, the assumption and treatment durations used lack justification.

5.2.9.2 Health-state costs

Lacking detailed published data on resource use in the patient population, data used to inform health-state costs stemmed from previous TAs. In both non-alloSCT health-states (pre- and post-12 weeks), i.e. progression-free and progressed disease, monitoring costs consisted of outpatient attendance, blood tests and CT and PET scans (see Table 5.17) and amounted to £68.78 per week.

Table 5.17: Weekly monitoring costs

Unit	Unit cost	Usage per week	Weekly cost
Outpatient attendance	£173.17	0.20	£34.52
Blood count	£3.10	0.20	£0.62
Biochemistry	£1.18	0.20	£0.24
CT scan	£120.99	0.06	£6.96
PET scan	£920.24	0.03	£26.45
Source: Model			

Progressed disease

Upon disease progression, initial treatment with pembrolizumab or SoC is discontinued and costs for BSC are applied as a one-off event.

Alive post-alloSCT

Post-alloSCT health-state costs were taken from Radford [29], a study reporting on costs in 14 relapsed or refractory classical Hodgkin Lymphoma patients treated with alloSCT, the source for resources and costs preferred by the committee of TA462 [6]. These costs were applied once upon treatment with alloSCT and were assumed to consist of alloSCT treatment costs, monitoring costs, costs of adverse events, costs of subsequent treatment and terminal care costs. No long-term costs were added.

Adverse event costs

A selection of grade 3+ AEs costs, based on previous appraisal TA462[6] and validated in a clinician survey, was applied dependent on treatment. Assuming that serious AEs lead to the discontinuation of treatment, patients on treatment beyond the first year were assumed to be free from AEs, and investigational agents were assumed to have no AEs. Resource use and costs of AEs were taken from the NHS reference costs by means of a weighted average of HRG codes, applied to the model as one-off event costs of [redacted], [redacted] for pembrolizumab in cohort 1 and cohort 2 respectively, and £1,945.74 for SoC in both cohorts.

Table 5.18: Adverse event costs

Adverse Event	Unit Cost	Source	Rational
Anaemia	£814.03	NHS reference costs 2015-16 [47]	TA411 TA399 TA391
Diarrhoea	£1,497.86	NHS reference costs 2015-16 [47]	TA391 TA440
Dyspnoea	£718.76	NHS reference costs 2015-16 [47]	TA420
Fatigue	£1,499.09	Brown (2013) [50] and NHS reference costs 2011-12[51] inflated with HCHS index	TA391
Leukopenia	£1,142.90	NHS reference costs 2015-16 [47]	TA391
Nausea	£872.42	NHS reference costs 2015-16 [47]	TA411
Neutropenia	£1,142.90	NHS reference costs 2015-16 [47]	TA411 TA399
Pyrexia	£3,923.50	NHS reference costs 2013-14 [52] inflated with HCHS index	TA366 TA311
Thrombocytopenia	£636.19	NHS reference costs 2015-16 [47]	TA399 TA440
Vomiting	£1,497.86	NHS reference costs 2015-16 [47]	TA360 TA440
Source: CS Table 99 [1]			

Terminal care costs

Terminal care costs are applied at death of patients in the non-alloSCT health states to reflect increased health care consumption in the period before death. The proportions of patients treated in different settings were taken from a population of non-small cell lung cancer patients (see Table 98 of the CS [1]).

Cost of terminal care resources stemmed from the same source but were updated with 2015-2016 NHS reference costs or increased for inflation with the HCHS hospital and community health service index. Hospital care, hospice care and homecare consisting of GP visits, nurse visits and drugs amounted to a total of £4,064.64 terminal care costs.

ERG comment: The ERG considers the costs associated with alloSCT to be under-estimated. In the model, a one-off cost was applied upon treatment with alloSCT. The company argues that it includes costs and resource use of alloSCT treatment, monitoring costs, costs of adverse events, of subsequent treatment and terminal care costs. The ERG wishes to point out that the company deviated from the methods in TA462 where the one-off cost was used in a scenario analysis, however, monthly costs for subsequent treatment and monitoring were added that were foregone in this TA. In their response to clarification question B22.a,[10] the company did not specify how the one-off cost based on a mean follow-up period of 3.44 years and an unknown proportion of deaths could reflect costs of a lifetime horizon. The ERG therefore applied monitoring costs, comparable to those used in TA462, over the lifetime horizon in their base-case, showing that the company's analysis favoured pembrolizumab.

5.2.10 Cost effectiveness results

In the deterministic base-case analysis, total QALYs and LYs gained were larger in the pembrolizumab treatment arm compared to UK SoC in both cohorts. Tables 5.19 and 5.20 show that the main benefit of pembrolizumab versus SoC are mostly due to QALY gains beyond week 12 with alloSCT (71% and 78% of incremental QALYs in cohort 1 and cohort 2 respectively). Total costs were also higher for pembrolizumab than for SoC. Incremental costs mainly resulted from differences in acquisition costs and alloSCT costs between pembrolizumab and SoC. Pembrolizumab treatment resulted in deterministic incremental cost effectiveness ratios (ICERs) of £43,511 and £48,571 per QALY gained for cohort 1 and cohort 2 respectively, as per the company's corrected base-case (Table 5.21).

Table 5.19: Cohort 1 QALYs breakdown (discounted)

	Week 0 to week 12		**Beyond week 12 (without alloSCT)**		**Beyond week 12 (with alloSCT)**
Base-case					
	PF	**PD**	**PF**	**PD**	**Alive**
Pembrolizumab	0.186	0.001	0.684	0.664	2.861
SoC	0.166	0.011	0.107	0.951	1.522
Corrected base-case					
Pembrolizumab	0.186	0.001	0.655	0.638	3.016
SoC	0.166	0.011	0.089	0.845	2.112
Source: (corrected) cost effectiveness model submitted by the company					

Table 5.20: Cohort 2 QALYs breakdown (discounted)

	Week 0 to week 12		**Beyond week 12 (without alloSCT)**		**Beyond week 12 (with alloSCT)**
Base-case					
	PF	**PD**	**PF**	**PD**	**Alive**
Pembrolizumab	0.186	0.001	0.457	0.745	2.503
SoC	0.180	0.003	0.113	0.960	1.455

Corrected base-case					
Pembrolizumab	0.186	0.001	0.426	0.701	2.757
SoC	0.180	0.003	0.092	0.849	2.076
Source: (corrected) cost effectiveness model submitted by the company					

Table 5.21: Company base-case results

Technologies	Cohort	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
Base-case							
UK SoC	Cohort 1	44,278	4.385	2.757	-	-	-
	Cohort 2	43,275	4.330	2.711	-	-	-
Pembrolizumab	Cohort 1	106,908	6.153	4.397	62,630	1.639	38,201
	Cohort 2	92,100	5.594	3.892	48,825	1.181	41,341
Corrected base-case							
UK SoC	Cohort 1	52,017	4.864	3.223	-	-	-
	Cohort 2	51,424	4.832	3.200	-	-	-
Pembrolizumab	Cohort 1	107,459	6.252	4.497	55,442	1.274	43,511
	Cohort 2	93,732	5.775	4.072	42,308	0.871	48,571
Sources: CS Table 101[1], corrected cost-effectiveness model submitted by the company ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years							

ERG comment: The ERG's concern relates to the exclusion of BSC as a comparator in the base-case analysis. BSC was not included as a comparator in the base-case, and therefore pembrolizumab could not be compared to all relevant alternatives at the same time.

5.2.11 Sensitivity analyses

The company performed and presented probabilistic sensitivity analysis (PSA) and deterministic sensitivity analysis (DSA) in order to quantify the uncertainty surrounding the company's results.

Compared with the deterministic results, the PSA with 1,000 iterations showed a comparable relative decrease in incremental costs and QALYs, which did not result in large changes to the ICER of cohort 1 (£43,653). In cohort 2, the PSA showed decreased incremental costs and even larger (relative) decreased incremental QALYs compared with the deterministic results, which resulted in an ICER of £50,894 (Table 5.22).

Cost effectiveness acceptability curves (CEACs) showed that there was a 60.1% (cohort 1) and 50.4% (cohort 2) probability of pembrolizumab to be cost effective compared to SoC at a willingness to pay (WTP) of £50,000 per QALY (Figures 5.5 and 5.6). However, these probabilities are reduced to 1.1% and 1.4% respectively at a WTP of £20,000 per QALY, and 20.5% and 16.1% at a WTP of £30,000 per QALY.

The company stated that DSAs were conducted for all key variables. Parameters were varied within their 5% and 95% confidence intervals where possible, and +/- 10% otherwise. The DSA results were presented in tornado diagrams including the 15 key model drivers. The following parameters were identified as most influential on the cost effectiveness of pembrolizumab versus SoC:

Cohort 1:

1. Discount rate – Outcomes (0.035; 0.000-0.060)

2. Response at week 12 – SoC – CR odds ratio ([redacted])
3. Response at week 12 – SoC – PR odds ratio ([redacted])

Cohort 2:

1. Response at week 12 – SoC – CR odds ratio ([redacted])
2. Discount rate – Outcomes (0.035; 0.000-0.060)
3. Response at week 12 – SoC – PR odds ratio ([redacted])

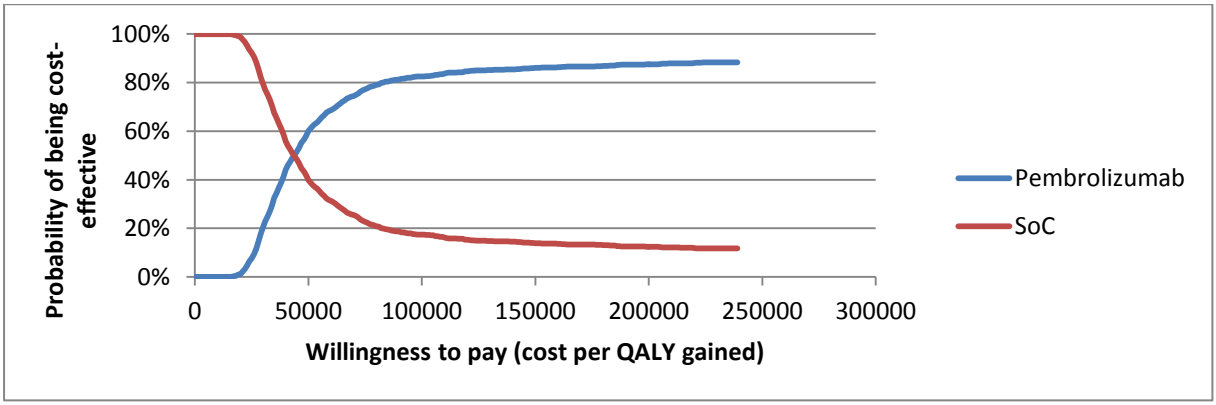
The WTP threshold of £50,000 per QALY was exceeded in the outcomes discount rate parameter and the CR odds ratio of SoC at week 12 response for cohort 1. For cohort 2, the WTP threshold was exceeded in all of the three abovementioned parameters.

Table 5.22: Incremental cost effectiveness results based on PSA (discounted, with CAA, 1,000 simulations)

Technologies	Cohort	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
Base-case						
UK SoC	Cohort 1	46,723	2.857	-	-	-
	Cohort 2	45,391	2.771	-	-	-
Pembrolizumab	Cohort 1	106,672	4.361	59,949	1.505	39,841
	Cohort 2	92,941	3.875	47,550	1.105	43,049
Corrected base-case						
UK SoC	Cohort 1	53,491	3.219	-	-	-
	Cohort 2	54,028	3.254	-	-	-
Pembrolizumab	Cohort 1	106,702	4.438	53,211	1.219	43,653
	Cohort 2	94,522	4.050	40,494	0.796	50,894

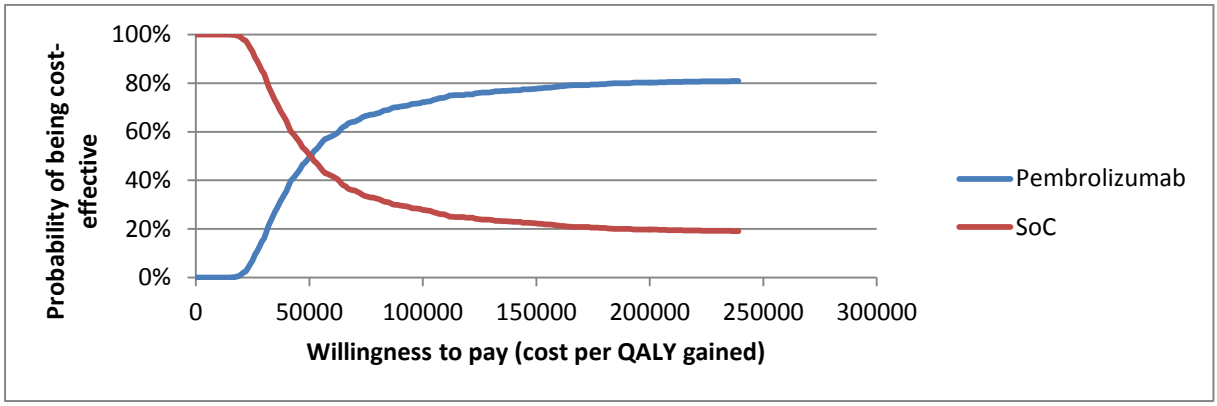
Sources: CS table 101, cost-effectiveness model after correction by company
ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years

Figure 5.5: Cohort 1 cost effectiveness acceptability curve (discounted, with CAA)



Source: corrected cost effectiveness model provided by the company.

Figure 5.6: Cohort 2 cost effectiveness acceptability curve (discounted, with CAA)



Source: corrected cost effectiveness model provided by the company.

The following five scenario analyses were performed by the company (Table 5.23). The results shown are based on the company's corrected base-case.

Scenario 1: assessing BSC as a comparator as per the NICE scope

Scenario 2: assessing different alloSCT rates

a. 100% alloSCT in patients with CR, PR or SD
b. Alternative lower PR alloSCT rate from MSD clinician survey

Scenario 3: using MAIC HR and OR rather than naïve ITC

Scenario 4: Alternative extrapolation scenarios to estimate PFS and OS

a. Considering a Weibull curve for week 0-12 PFS extrapolation in cohort 2
b. Considering a Gompertz curve for week 12+ PFS extrapolation in cohort 2
c. Considering a Lognormal curve following alloSCT

Scenario 5: assessing varying the time horizon to 50 years

Across all the scenarios, the ICER ranged between £23,564 and £47,957 for cohort 1, and between £24,492 and £56,677 for cohort 2. Scenario 2a had the biggest impact on the ICER in both cohorts (ICER decrease of approximately £20,000 and £24,000 for cohort 1 and 2 respectively).

Table 5.23: Results from the scenario analyses based on the company's corrected base-case

Scenario	Cohort	Pembrolizumab			UK SOC			Pembrolizumab vs UK SOC		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Inc. costs	Inc. QALYs	ICER
Company's corrected base case	Cohort 1	£107,459	6.252	4.497	£52,017	4.864	3.223	£55,442	1.274	£43,511
	Cohort 2	£93,732	5.775	4.072	£51,424	4.832	3.200	£42,308	0.871	£48,571
Scenario 1	Cohort 1	£107,459	6.252	4.497	£51,188	4.864	3.223	£56,270	1.274	£44,161
	Cohort 2	£93,732	4.832	3.200	£50,713	4.832	3.200	£43,018	0.871	£49,387
Scenario 2a	Cohort 1	£119,943	8.503	6.768	89,436	7.175	5.474	£30,507	1.295	£23,564
	Cohort 2	£116,185	8.261	6.537	£87,472	7.053	5.364	£28,713	1.172	£24,492
Scenario 2b	Cohort 1	£106,221	6.029	4.272	£49,951	4.736	3.098	£56,270	1.173	£47,957
	Cohort 2	£91,431	5.520	3.819	£49,360	4.705	3.077	£42,070	0.742	£56,677
Scenario 3	Cohort 1	£107,459	6.252	4.497	£45,292	4.419	2.790	£62,166	1.707	£36,423
	Cohort 2	£93,732	5.775	4.072	£46,944	4.558	2.933	£46,787	1.139	£41,087
Scenario 4a	Cohort 2	£93,261	5.766	4.062	£51,234	5.814	3.175	£42,027	0.886	£47,410
Scenario 4b	Cohort 2	£93,439	5.688	4.000	£51,500	4.852	3.217	£41,938	0.783	£52,562
Scenario 4c	Cohort 1	£107,459	6.451	4.642	£52,016	5.003	3.324	£55,442	1.318	£42,075
	Cohort 2	£93,732	5.957	4.204	£51,423	4.969	3.300	£42,308	0.904	£46,812
Scenario 5	Cohort 1	£107,459	6.377	4.582	£52,016	4.951	3.283	£55,442	1.300	£42,651
	Cohort 2	£93,732	5.889	4.150	£51,423	4.918	3.259	£42,308	0.890	£47,516
Source: CS Table 20.[53]										

ERG comment: The ERG had minor concerns about (a) the choice of variation in the DSA, (b) the cost effectiveness probability of pembrolizumab at lower WTP thresholds in the PSA, (c) inappropriate parameters in the PSA, and (d) an insufficient number of iterations in the PSA.

(a) In variables for which it was not possible to use 5% and 95% confidence intervals, a variation of +/- 10% was chosen without providing any rationale for this decision. Additionally, the ERG believes this variation may be small when wishing to assess the full impact on the ICER.

(b) The probability of pembrolizumab being cost effective at WTP thresholds of £20,000 and £30,000 is much lower compared to the base-case WTP, indicating the CEAC gradient to be very steep.

(c) Patient characteristics (proportion female, average weight, body surface area) were included in the PSA, although they are considered first order uncertainty and typically not reflected in cohort model PSAs.

(d) The company ran the PSA on 1,000 iterations. The ERG concluded that this number was insufficient to test the robustness of the model, and therefore re-ran the analysis on 10,000 iterations

5.2.12 Model validation and face validity check

5.2.12.1 Face validity

The selected time-to-event models and health state utility values for the base-case analysis were validated by UK clinical experts. No detail was provided in the CS concerning the expert elicitation method and the number of experts consulted.

AlloSCT rates, which were obtained by using UK clinical expert opinion through a survey performed by the company and joining these with survey results from an existing survey, have been compared to alloSCT rates reported in previous studies. The rates used by the company were higher than in a French study[54] and lower than the rates reported in Cheah et al. (2016),[7] which were considered too low and too high in TA462, respectively. Several responses from the survey conducted by the company indicated that alloSCT could be administered after PD. However, this assumption was not included in the model following further discussions with UK clinicians because it was not thought to be UK standard practice. Additionally, the alloSCT rates have been validated by a UK clinical expert in this area. This expert suggested that alloSCT rates would be higher than the ones used in the cost effectiveness model, with alloSCT rates in the PR as high as in CR.

5.2.12.2 Internal validity

The company submission states that: "*the model structure, assumptions and rationale were critically reviewed by an independent health economics modelling expert*."[1, 39]

5.2.12.3 External validity

The survival estimates obtained from the cost effectiveness model were validated against the studies used to inform PFS and OS estimates of the model. The outcomes obtained for pembrolizumab and SoC from the cost effectiveness model were compared to the KEYNOTE-087 trial (Table 5.24). From this comparison, the company concluded that the proportions of patients in pre-progression and surviving at different points in time were similar in the model and KEYNOTE-087 and the SoC sources.[1, 7, 21]

5.2.12.4 Cross validity

No cross-validation of the model assumptions, model structure and model outcomes were performed with the previous TA462 in the same indication.[6]

Table 5.24: Comparison of model and trial outcomes

		Pembrolizumab			UK SoC		
Outcome		**Base case**	**KEYNOTE -087**	**ERG retrieval from the model**	**Base case**	**Cheah et al[7]**	**ERG retrieval from the model**
% PFS at 1 Year *	Cohort 1	54.79%	[redacted]	59.44%	4.1%[a]	~7.5%[a] -	3.97%
	Cohort 2	39.07%	[redacted]	43.75%	4.9%[a]		4.77%
OS at week 12	Cohort 1	98.96%	[redacted]	98.96%	98.96%	~100%	98.96%
	Cohort 2	98.76%	[redacted]	98.78%	98.76%		98.78%
OS at 72 Months**	Cohort 1	28.00%	-	15.50%	15.00%	15.00%	10.87%
	Cohort 2	22.00%		12.76%	15.00%		10.95%
OS after alloSCT 5 years		Base case	KEYNOTE-087		Base case	Lafferty et al.[21]	
	Cohort 1	54.50%	-	51.22%	54.50%	53.47%	51.22%
	Cohort 2						

Source: adapted from CS Table 102[1]
alloSCT = allogeneic stem cell transplantation; OS = overall survival; PFS = progression-free survival; SoC = standard of care
*using data post week 12 assuming no alloSCT as per KEYNOTE-087 design
** when no alloSCT is assumed as per assumption made about Cheah SoC arm
[a] Provided in the response to the clarification letter[10]

ERG comment: The main ERG concerns about model validation are (a) the non-reproducibility of Table 5.24, (b) the proportion of patients in the stable disease response status at 12 weeks, (c) the lack of cross-validation with TA462.

(a) The ERG attempted to retrieve the model outcomes presented in Table 5.24 but consistently retrieved different figures than provided in the CS. Additionally, this table (based on Table 102 of the CS) reports five-year OS after alloSCT from Lafferty et al.[21], which is derived from KM estimates made available in TA462. These should be interpreted with extreme caution, because the plateau at the end of these KM estimates (starting at approximately 20 months) may be caused by censoring. The abstract only reported one-year OS after alloSCT (69%). After the large number of events in the first year, it would be implausible for the rate of events to slow down that considerably. The ERG is concerned by the validity of the figures provided in Table 5.24 and considers Table 5.24 to be potentially misleading.

(b) The company assumed that all patients, who did not completely or partially respond and who did not progress or die at the 12-week decision nodes, were in the stable disease response category. Patients with a non-evaluable response are consequently automatically included in the stable disease response category. This assumption probably leads to an overestimation of the proportion of patients in the stable disease response status compared to KEYNOTE-087.

(c) Complete cross validation with TA462 was not performed by the company in both the CS and clarification response. The main differences between TA462 and the current assessment are the model structure, and how alloSCT is incorporated in the cost effectiveness model. TA462 used a three health states (progression-free, progressed, dead) semi-Markov model while the current model is composed of a short-term component (first 12 weeks), a decision tree element (at 12 weeks) and a long-term component (after 12 weeks). Additionally, progression was not allowed post-alloSCT in the current assessment while it was incorporated in TA462. Different assumptions were also made concerning the composition of SoC between the two assessments. All these discrepancies may have influenced the health benefits and costs obtained in the SoC arm. Table 5.25 compares the results of SoC between TA462 and the current assessment. The health benefits obtained from SoC were almost doubled and the costs of SoC were more than doubled in the current assessment compared to TA462. These discrepancies are most likely explained by the fact that patients in TA462 may receive alloSCT after 6 months while patients are considered for alloSCT after 12 weeks in the current assessment. These different assumptions have likely influenced health benefits and costs of SoC.

Table 5.25: Comparison of SoC results between TA462 and the current assessment

Assessment	Total QALY	Total costs
TA462[a]	1.870	£23,668
Current assessment[b]	3.684	£52,017

[a] Outcomes considered as the AC's most plausible analysis, retrieved from the committee papers for the second AC meeting, Table 4 of the ERG commentary on the company additional evidence
[b] Retrieved from the corrected company's cost effectiveness model, post clarification response, Cohort 1

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.26 summarises all main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table 5.26: Main ERG critique of company's submitted economic evaluation

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
• Incorporation of alloSCT at 12 weeks only	+/-	None	Not addressed
• No lag between decision and procedure	+	None	Not addressed
• No progressed disease health state post-alloSCT	+	SA	Not addressed
Population, interventions and comparators, perspective and time horizon (sections 5.2.3-5)			
• Comparator data based on mix of cohort 1 and 2	+ cohort 1, - cohort 2	None	Not addressed
• BSC only in scenario analysis	+/-	None	Not addressed
• Time horizon of 40 years	-	BC (FV)	Addressed in SA
Treatment effectiveness and extrapolation (section 5.2.6)			
• Use of alternative sources due to immature OS data from KEYNOTE-087	+/-	None	Requested, partially addressed
• Single-arm study used to inform treatment effectiveness	+/-	None	Not addressed

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
• Use of naive indirect treatment comparison	-	SA	MAIC explored in SA
• Over-estimation of post-alloSCT OS based on Lafferty et al	+	SA	Not addressed
• Curves derived from entire study data fitted to pre-12 week period	+/-	None	Not addressed
• Inflated SD health state due to patients with non-evaluable response status being considered to have SD	+/-	None	Not addressed
• Combining of MSD and BMS surveys likely introduces bias	+	BC (FV)	Addressed in SA
• Patients with PD cannot receive alloSCT	+	BC (FV)	Addressed in SA
• Increased uncertainty in post-12 week PFS due to fitting curves from 12 weeks onwards	+	SA	Partially addressed in SA
• HRs equal for pre- and post-12 week periods	+/-	None	Requested, explored in SA
• PFS used as proxy for TTD pre-12 weeks without justification	-	None	Not addressed
• TTD capped at 24 months	+	SA	Addressed in SA
Health-related quality of life (section 5.2.8)			
• Utilities only derived from 12-week observations	+	BC (MJ)	Company provided mixed model utilities
• Progressed disease utility not from KEYNOTE-087	+	BC (MJ)	Not addressed
• PFS utility for patients without alloSCT calculated based on patients with and without alloSCT	+	BC (MJ)	Not addressed
• Not using post alloSCT utilities from Kurosawa et al.[45] (only disutilities are used)	+	BC (MJ)	Not addressed
• Inconsistency with treatment effectiveness section regarding calculation of proportion of responders	+/-	BC (MJ)	Not addressed
Resources and costs (section 5.2.9)			
• Likely over-estimation of SoC resource use and costs due to SoC chemotherapy mix	+	None	Not addressed
• Under-estimation of post-alloSCT costs	+	BC (FV)	Requested, not addressed
• Missed doses not incorporated	+/-	None	Not addressed
Cost-effectiveness analyses (sections 5.2.10 and 5.2.11)			
• Exclusion of BSC from base-case	+/-	None	Requested, not addressed
• Patient characteristics included in PSA	+/-	BC (FE)	Not addressed
Validation (section 5.2.12)			
• Complete cross validation with TA462 not performed	NA	None	Not addressed

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
BC = base-case; FE = fixing error; FV = fixing violations; MJ = matters of judgement; NA = not applicable; SA = scenario analysis [a]Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.			

Based on all considerations from Section 5.2 (summarised in Table 5.26), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016):[55]

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

Additionally, exploratory sensitivity analyses were performed by the ERG to examine the potential impact of alternative assumptions on the cost effectiveness estimates.

The ERG's base-case:

Fixing errors

1. Error in the calculation of AE disutilities
   The ERG corrected the error.
2. Patient characteristics included in the PSA
   The ERG corrected this by excluding patient characteristics from the PSA.

Fixing violations

3. Combining MSD and BMS surveys for obtaining the probabilities of alloSCT uptake conditional on response.
   The ERG used the MSD survey only.
4. Time horizon of 40 years, despite some patients still being alive at that point.
   The ERG used a time horizon of 50 years.
5. Model excludes long-term monitoring costs post-alloSCT.
   The ERG included these consistent with committee's preferences in TA462.

Matters of judgment

6. Use of utility values estimated based on observations from week 12 only; progressed disease utility was estimated based on an alternative source; PFS utility for patients without alloSCT calculated based on patients with and without alloSCT; not using post alloSCT utilities from Kurosawa et al.[45] (only disutilities); and inconsistency with treatment effectiveness section regarding calculation of proportion of responders
   The ERG used the mixed model utilities provided by the company and the literature (Kurosawa et al.[45]) to calculate alternative utilities (see section 5.2.8 for more details).
7. Distributions for pre-12 weeks OS over-estimates mortality.
   The ERG used alternative distributions (exponential for cohort 1, lognormal for cohort 2) for pre-12 weeks OS.

8. Proportion of patients in PD state receiving alloSCT was set to 0.
   The ERG used the results from MSD's clinician survey to inform this.

5.3.1 Probabilistic ERG base-case

The ERG performed a PSA to obtain the ERG base-case incorporating all abovementioned adjustments. This resulted in probabilistic ICERs of £64,186 and £78,696 per QALY gained for pembrolizumab (with CAA) versus SoC for cohorts 1 and 2 respectively (Table 5.27). The individual effects of each change on costs, QALYs and ICERs are presented in Section 6, Table 6.1. For comparison, the deterministic ERG base-case ICERs were £61,705 and £73,594 per QALY gained, for cohorts 1 and 2 respectively.

Table 5.27: ERG base-case (probabilistic)

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Cohort 1	Pembrolizumab	£108,894	4.602			
	SoC	£53,729	3.743	£55,165	0.859	£64,186
Cohort 2	Pembrolizumab	£93,953	4.277			
	SoC	£53,487	3.763	£40,466	0.514	£78,696

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

The CEACs based on the ERG base-case (Figures 5.7 and 5.8) show that pembrolizumab has a probability of being cost effective of 18% and 42% for cohort 1 and 21% and 40% for cohort 2 at thresholds of £30,000 and £50,000 per QALY gained, respectively.

Figure 5.7: Cost effectiveness acceptability curve for ERG base-case (cohort 1)

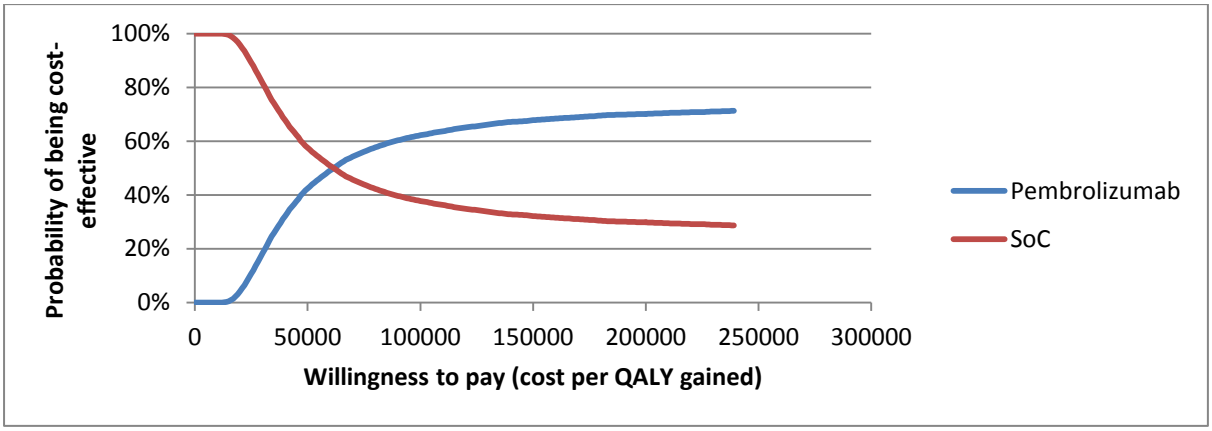
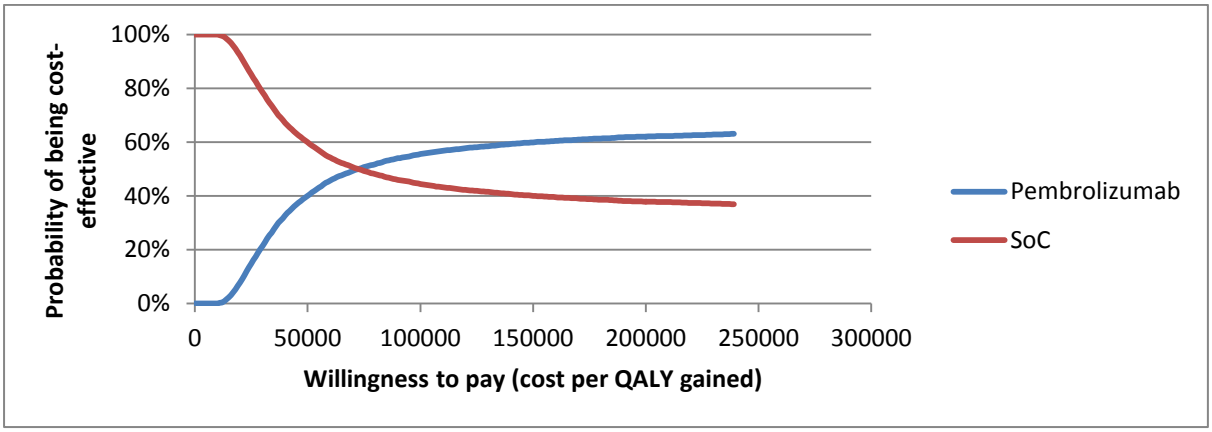


Figure 5.8: Cost effectiveness acceptability curve for ERG base-case (cohort 2)



5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of the following alternative assumptions on the cost effectiveness estimates. These were all performed using the ERG base-case. Results are presented in Table 6.2 in Section 6.

Exploratory analyses using the ERG base-case:

1. Alternative parametric survival models:
   Cohort 1: a) for post-12 weeks PFS (Gompertz)
   Cohort 2: a) post-12 weeks PFS (Gompertz) and b) post-12 weeks PFS (generalised gamma)
2. Use of MAIC instead of the naive indirect treatment comparison for estimating PFS hazard ratios and response rates at 12 weeks
3. Remove the 24-months cap on TTD
4. Use lower post-alloSCT utility (i.e. the PD utility) to explore the impact of ignoring PD after alloSCT
5. Use of alternative assumptions to extrapolate post-alloSCT OS from Lafferty et al (2017)

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

Reviewing the overall evidence, the ERG confirmed that there was no existing cost effectiveness model for pembrolizumab for the current indication, and thus that development of a de novo model was necessary. The economic model described in the CS is considered by the ERG to meet the NICE reference case, with the exceptions of (1) the exclusion of a comparator that was identified in the scope, and (2) a slightly short time horizon. The absence of BSC from the main analysis was justified by a lack of data, and has been accepted by the committee in previous appraisal TA462. Another potential comparator in the future may be nivolumab, which has recently been recommended for use in part of the present population (cohort 1). The time horizon was extended (from 40 to 50 year) by the company to cover patients' lifetime in scenario analysis, and this was adopted in the ERG base-case.

The company's corrected base-case ICERs (probabilistic) of pembrolizumab (with CAA) compared with SoC were £43,653 and £50,894 per QALY gained for cohort 1 and cohort 2 respectively. The cost effectiveness results were not robust to scenario and one-way sensitivity analyses conducted by the company. Scenario analyses indicated that response rates at week 12, the proportions of patients

receiving alloSCT, and the use of the MAIC instead of the naïve indirect comparison were major drivers of model results, mostly resulting in less favourable cost effectiveness estimates for pembrolizumab versus SoC.

The ERG incorporated various adjustments to the company's base-case. The ERG base-case resulted in ICERs (probabilistic) of pembrolizumab (with CAA) versus SoC of £64,186 and £78,696 per QALY gained for cohorts 1 and 2 respectively. For comparison, the deterministic ERG base-case ICERs were £61,705 and £73,594 per QALY gained for cohorts 1 and 2 respectively. The three most influential adjustments made by the ERG in its base-case for both cohorts were (in descending order) (1) the use of alternative utility values, (2) the use of the MSD survey only to estimate uptake of alloSCT instead of combining the MSD and BMS surveys, and (3) allowing alloSCT also in patients in the progressed disease state, in line with the MSD survey.

The ERG identified major and minor issues and uncertainties that affected the cost effectiveness analysis. Major issues and uncertainties are listed in the following. One major limitation was the company's model structure, which induced the implausible assumption that patients could only be eligible and receive alloSCT at 12 weeks after treatment start. The ERG deemed this implausible because response may, in reality, be obtained later than at 12 weeks and because, in practice, there is a lag between the decision to pursue alloSCT and the time at which the procedure is performed. The assumption lacked appropriate justification and deviated from how alloSCT was incorporated in TA462. The model is therefore a poor reflection of reality. Also, this model structure necessitated the differential fitting of parametric models to survival data for the pre- and post-12 week periods, inducing additional uncertainty. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT. The impact of the limitations related to the model structure on model outcomes is unknown.

It should be noted that the appropriate approach for incorporating alloSCT in the model would have been to use time to alloSCT data directly from the main source of evidence. However, KEYNOTE-087 was not designed as a bridging study and poorly reflected clinical practice in the UK in terms of alloSCT uptake. The company therefore opted to inform alloSCT uptake conditional on response status at 12 weeks after treatment start through a UK clinician survey and then combined these survey results with the previously performed BMS survey results (from TA462). The ERG did not deem the combination of both surveys appropriate and considers there to be major uncertainty about the alloSCT uptake estimates. Furthermore, the elicited alloSCT uptake (from the MSD survey) for patients with progressed disease was ignored. Both, the combining of both surveys and ignoring alloSCT uptake in progressed disease patients, were shown in scenario analysis to be major drivers of cost effectiveness.

A major limitation was the use of single-arm evidence to inform treatment effectiveness. There was uncertainty whether the MAIC or the naïve indirect comparison should be used. The company provided both and the ERG, like the company, used the naïve indirect comparison in the base-case and the MAIC in scenario analysis. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation as this necessitated the use of post-alloSCT OS and utility estimates from alternative data sources, one of which was based on 13 patients only. [redacted], and the ERG considers that these may be informative for the present analysis. Furthermore, the company's method used for extrapolating OS post-alloSCT was deemed by the ERG to over-estimate OS, which significantly favoured pembrolizumab.

It is of note that the population used for the comparator was a mixed population of cohorts 1 and 2, that is, that did and did not receive autoSCT, derived from Cheah et al.[7] The Cheah et al population is more comparable with KEYNOTE-087 cohort 1 than with cohort 2 in terms of patient characteristics. The

use of this mixed comparator population likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively, but this could not be formally explored in scenario analysis.

Of further note, the economic model, and the evidence from KEYNOTE-087, rely on the assumption that treatment with pembrolizumab is capped at 24 months, which is inconsistent with its SmPC. It is unclear whether in UK clinical practice pembrolizumab would also be provided for a maximum of 24 months. The company and the ERG explored the impact of relaxing this assumption in scenario analysis.

Model extrapolations lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.

In exploratory analysis the ERG found that removing the 24-months cap on TTD had the largest impact on the ICERs in cohort 1 and a significant impact in cohort 2, and increased them to £78,992 and £79,284 per QALY gained for cohorts 1 and 2 respectively. The exploratory analysis with the largest impact in cohort 2 (ICER increased to £95,712) and the second largest impact in cohort 1 (ICER increased to £78,204) was the use of alternative assumptions when extrapolating post-alloSCT OS using data from Lafferty et al (2017). In cohort 2, the use of alternative parametric models for post-12 week PFS also substantially increased the ICERs to £87,401 and £90,152 per QALY gained when using the Gompertz and generalised gamma respectively, reflecting the significant uncertainty about extrapolating PFS in this model. The use of the MAIC instead of the naïve indirect comparison decreased the ICERs to £54,466 and £60,372 per QALY gained for cohorts 1 and 2 respectively. Assuming a lower post-alloSCT utility to explore the effect of the omission of a progressed disease health state post-alloSCT resulted only in small increases in the ICERs (by approximately £2,000 in both cohorts).

In conclusion, given that the ERG base-case ICERs are estimated to be substantially above £60,000 per QALY gained for both cohorts, with none of the scenarios resulting in ICERs below £50,000 per QALY gained, and the significant uncertainty induced by modelling choices and the use of single-arm studies with immature OS data, uncertainty around the cost effectiveness of pembrolizumab remains substantial.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

In Section 5.3 the ERG's base-case was presented, which was based on various changes compared to the company's base-case. Tables 6.1 and 6.2 show how each individual change impacts the ICER in cohorts 1 and 2 respectively, plus the combined effect of all changes simultaneously. The analyses numbers in these tables correspond to the analyses numbers reported in Section 5.3. Furthermore, the exploratory analysis is presented in Tables 6.3 and 6.4 for cohorts 1 and 2 respectively (conditional on the ERG base-case). Appendix 1 contains technical details on the analyses performed by the ERG.

Table 6.1: ERG base-case cohort 1 (deterministic), pembrolizumab with CAA

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company corrected base-case cohort 1	Pembrolizumab	£107,459	4.497			
	SoC	£52,017	3.223	£55,442	1.274	£43,511
Fixing errors (1)-(2)	Pembrolizumab	£107,459	4.496			
	SoC	£52,017	3.215	£55,442	1.282	£43,262
MSD survey only (3)*	Pembrolizumab	£105,128	4.072			
	SoC	£45,920	2.848	£59,208	1.224	£48,363
Time horizon 50 years (4)*	Pembrolizumab	£107,459	4.582			
	SoC	£52,017	3.275	£55,442	1.307	£42,412
Include monitoring costs post-alloSCT (5)*	Pembrolizumab	£110,298	4.496			
	SoC	£54,004	3.215	£56,294	1.282	£43,927
Alternative utility values (6)*	Pembrolizumab	£107,459	4.669			
	SoC	£52,017	3.617	£55,442	1.052	£52,705
Alternative pre-12 week OS distributions (7)*	Pembrolizumab	£107,496	4.499			
	SoC	£52,054	3.218	£55,442	1.282	£43,262
Proportion of alloSCT in PD state from MSD survey (8)*	Pembrolizumab	£107,934	4.524			
	SoC	£55,125	3.397	£52,809	1.127	£46,841
ERG base-case cohort 1 (combining adjustments 1-8)	Pembrolizumab	£107,998	4.460			
	SoC	£50,913	3.535	£57,085	0.925	£61,705

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year
* conditional on fixing errors (1) - (2)

Table 6.2: ERG base-case cohort 2 (deterministic), pembrolizumab with CAA

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company corrected base-case cohort 2	Pembrolizumab	£93,732	4.072			
	SoC	£51,424	3.200	£42,308	0.871	£48,571
Fixing errors (1)-(2)	Pembrolizumab	£93,732	4.071			
	SoC	£51,424	3.193	£42,308	0.878	£48,178
MSD survey only (3)*	Pembrolizumab	£89,745	3.633			
	SoC	£45,464	2.835	£44,281	0.798	£55,478
Time horizon 50 years (4)*	Pembrolizumab	£93,732	4.149			
	SoC	£51,424	3.251	£42,308	0.897	£47,141
Include monitoring costs post-alloSCT (5)*	Pembrolizumab	£96,327	4.071			
	SoC	£53,378	3.193	£42,949	0.878	£48,908
Alternative utility values (6)*	Pembrolizumab	£93,732	4.309			
	SoC	£51,424	3.594	£42,308	0.714	£59,223
Alternative pre-12 week OS distributions (7)*	Pembrolizumab	£93,967	4.086			
	SoC	£51,607	3.208	£42,360	0.878	£48,236
Proportion of alloSCT in PD state from MSD survey (8)*	Pembrolizumab	£94,579	4.120			
	SoC	£54,466	3.371	£40,113	0.750	£53,508
ERG base-case cohort 2 (combining adjustments 1-8)	Pembrolizumab	£93,095	4.118			
	SoC	£50,609	3.541	£42,486	0.577	£73,594

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year
* conditional on fixing errors (1) - (2)

Table 6.3. Exploratory analysis conditional on ERG base-case cohort 1 (deterministic), pembrolizumab with CAA

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company corrected base-case cohort 1	Pembrolizumab	£107,459	4.497			
	SoC	£52,017	3.223	£55,442	1.274	£43,511
ERG base-case cohort 1	Pembrolizumab	£107,998	4.460			
	SoC	£50,913	3.535	£57,085	0.925	£61,705
Alternative distributions (1.a)	Pembrolizumab	£107,552	4.361			
	SoC	£50,937	3.540	£56,615	0.821	£68,966
Use of MAIC (2)	Pembrolizumab	£107,998	4.460			
	SoC	£47,997	3.359	£60,001	1.102	£54,466

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
No 24-months cap on TTD (3)	Pembrolizumab	£123,990	4.460			
	SoC	£50,913	3.535	£73,077	0.925	£78,992
Lower post-alloSCT utility (4)	Pembrolizumab	£107,998	4.346			
	SoC	£50,913	3.446	£57,085	0.900	£63,420
Alternative OS post-alloSCT assumption (5)	Pembrolizumab	£107,030	3.558			
	SoC	£50,157	2.830	£56,873	0.727	£78,204
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						

Table 6.4. Exploratory analysis conditional on ERG base-case cohort 2 (deterministic), pembrolizumab with CAA

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company corrected base-case cohort 2	Pembrolizumab	£93,732	4.072			
	SoC	£51,424	3.200	£42,308	0.871	£48,571
ERG base-case cohort 2	Pembrolizumab	£93,095	4.118			
	SoC	£50,609	3.541	£42,486	0.577	£73,594
Alternative distributions (1.a)	Pembrolizumab	£92,750	4.040			
	SoC	£50,698	3.558	£42,052	0.481	£87,401
Alternative distributions (1.b)	Pembrolizumab	£92,556	3.995			
	SoC	£50,550	3.529	£42,007	0.466	£90,152
Use of MAIC (2)	Pembrolizumab	£93,095	4.118			
	SoC	£45,924	3.337	£47,171	0.781	£60,372
No 24-months cap on TTD (3)	Pembrolizumab	£96,380	4.118			
	SoC	£50,609	3.541	£45,771	0.577	£79,284
Lower post-alloSCT utility (4)	Pembrolizumab	£93,095	4.013			
	SoC	£50,609	3.453	£42,486	0.560	£75,835
Alternative OS post-alloSCT assumption (5)	Pembrolizumab	£92,204	3.287			
	SoC	£49,863	2.844	£42,341	0.442	£95,712
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						

7. END OF LIFE

According to the NICE criteria for End of Life, the following criteria should be satisfied:

- 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 3 months, compared to current NHS treatment.

According to the company there is not a valid estimate of OS for patients with RRcHL within UK clinical practice. However, based on their literature searches, the company estimate that OS ranges from 17.1 months to 19 months (CS, Table 51, page 129). In addition, in TA462 (Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma) the committee "acknowledged that nivolumab did not unequivocally meet the criterion for short life expectancy but that it was plausible that the criterion could apply. It therefore agreed that on balance, nivolumab met the criterion for short life expectancy, and that it would take this into account in its decision-making."[6]

Regarding the second criterion, the company states that "As of March 21st 2017 [redacted] for Cohorts 1 and 2. However, the small number of deaths reported during the current follow-up period (15.9 months) indicates a substantially longer median survival than that offered by current therapies. The OS rate at 15 months in cohort 1 and 2 was reported using Kaplan-Meier estimates at [redacted], respectively.[56, 57]". Based on the company's economic model base case, the company predictions are 74 months for pembrolizumab and 53 months for SoC; therefore, the increment is 21 months in cohort 1 (ERG BC: Pembrolizumab LYs: 5.968 in months 71.616; SoC LYs: 4.761 in months 57.132). For cohort 2, the model predicts 67 months for pembrolizumab and 52 months for SoC; therefore, the increment is 15 months (ERG BC Pembrolizumab LYs: 5.517 in months 66.204; SoC LYs: 4.767 in months 57.204).

Overall, the ERG believes that the second criterion is more likely to be met. Regarding the first criterion, there is considerable uncertainty.

8. OVERALL CONCLUSIONS

8.1 Statement of principal findings

The company did not identify any randomised controlled trials of pembrolizumab and its comparators in patients with classical Hodgkin Lymphoma who have either received autoSCT and BV or BV alone due to autoSCT being unsuitable. One ongoing, single arm study of the efficacy and safety of pembrolizumab was identified (KEYNOTE-087) and this formed the basis of the submission. KEYNOTE-087 includes 150 patients (14 UK patients) relevant to this appraisal. It covers both cohorts of interest (Cohort 1: people with relapsed or refractory cHL who have received autologous stem cell transplant and brentuximab vedotin and Cohort 2: patients who have received brentuximab vedotin when autologous stem cell transplant is not a treatment option). The company presented data based on a median follow up of 15.9 months. The median time on treatment was [redacted] days for Cohort 1 and [redacted] days for Cohort 2.

The primary outcome of KEYNOTE-087 was overall response rate (ORR) as assessed by independent committee. ORR was 75.4% in Cohort 1 and 66.7% in Cohort 2. Median progression free survival (PFS) in Cohort 1 was 16.7 months (11.2 to NR). In cohort 2 it was 11.1 months (7.6 to 13.7). Median overall survival (OS) was [redacted]. At 12 months survival was [redacted] in Cohort 1 and [redacted] in Cohort 2. In cohort 1 [redacted] of patients had one or more adverse events. In Cohort 2 [redacted] of patients had one or more adverse events. The company noted that most AEs were low grade ([redacted] and [redacted] grades 3 to 5 in Cohort 1 and 2 respectively). In Cohort 1 [redacted] of AEs were classed as serious and in Cohort 2 [redacted]. The most common adverse events were pyrexia, cough, fatigue, diarrhoea and vomiting. The company conducted post-hoc analyses of response at 12 weeks to use in the comparison of clinical and cost effectiveness. The ERG noted that overall response rates were lower at 12 weeks than over the whole course of the trial ([redacted]

As KEYNOTE-087 did not have a comparator group the company identified a retrospective observational study from the literature (Cheah 2016 et al) to use as a comparator. This is a USA database study in which patients received the following types of therapy: investigational agent(s), gemcitabine, bendamustine, any other alkylator, BV retreatment, platinum based treatment, autoSCT or alloSCT, or other treatment. The company has not provided separate data for comparators; instead a combined data set has been provided for multiple comparators.

The company performed two types of analyses: a naïve indirect comparison between KEYNOTE-087 and Cheah and a matched adjusted indirect treatment comparison (MAIC) of the two studies.

Almost all results for PFS show a significant benefit for pembrolizumab versus SoC. Likewise, all results for ORR significantly favour pembrolizumab over SoC. Results of the naïve comparison are similar to MAIC. However, the results of the naïve comparison and MAIC are not reliable because they are likely to contain systematic error but it is not possible to estimate the size of the potential error. Both have major limitations and neither are fully reliable for decision making.

With regards to the health economic model submitted by the company, the ERG demonstrated that there was substantial uncertainty surrounding the ICERs and that alternative assumptions could change the ICER significantly. One major limitation was the company's model structure, which induced implausible assumption around the timing of alloSCT. The model was therefore considered a poor reflection of reality and likely to over-estimate cost effectiveness of pembrolizumab. There also remains substantial uncertainty about the uptake of alloSCT.

The use of single-arm evidence to inform treatment effectiveness was viewed as a major limitation and there was uncertainty about whether the MAIC or the naïve indirect comparison should be used. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation and the ERG considers that future data cuts may be informative for the present analysis. It is of note that the population used for the comparator was a mixed population of patients that did and did not receive autoSCT, which likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively. Of further note, the economic model, and the evidence from KEYNOTE-087, rely on the assumption that treatment with pembrolizumab is capped at 24 months, which is inconsistent with its SmPC. It is unclear whether in UK clinical practice pembrolizumab would also be provided for a maximum of 24 months. The substantial uncertainty in the evidence translates into model extrapolations that lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.

Apart from this, numerous issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in probabilistic ICERs of pembrolizumab (with CAA) versus SoC of £64,186 and £78,696 per QALY gained for cohorts 1 and 2 respectively.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. The scenarios with the largest impact were alternative assumptions for extrapolating post-alloSCT OS (upward effect on the ICER), alternative survival models for extrapolating post-12 week PFS (upward effect on the ICERs), the use of the MAIC instead of the naïve comparison (downward effect on the ICERs) and removing the cap of 24 months on TTD (upward effect on the ICERs).

In conclusion, given that the ERG base-case ICERs are estimated to be substantially above £60,000 per QALY gained for both cohorts, with none of the scenarios resulting in ICERs below £50,000 per QALY gained, and the significant uncertainty induced by modelling choices and the use of single-arm studies with immature OS data, uncertainty around the cost effectiveness of pembrolizumab remains substantial.

8.2 Strengths and limitations of the assessment

The majority of searches for eligible studies in the CS were well documented and easily reproducible. Searches were carried out on all databases recommended in the NICE 2013 guide to the methods of technology appraisal sections 5.2.2 and 5.2.4. The clinical effectiveness strategies utilised recognised study design filters. Supplementary searches of conference proceedings and the NICE website and the WHO ICTRP trials database, were undertaken by the company in order to identify additional studies not retrieved by the main searches.

The clinical evidence is based on a well conducted, multicentre single-arm trial reflecting both cohorts of patients relevant to the decision problem. Outcomes assessed reflect the scope.

The main weakness is the lack of RCTs in this appraisal. Outcomes relating to pembrolizumab are based on a single arm trial. Comparisons with the comparators in the scope are problematic due to the availability of only one US study with a mix of different treatments. The naïve and matched adjusted comparisons conducted by the company have a number of limitations and represent a much weaker level of evidence than a RCT. Additionally progression-free survival and overall survival data are not fully mature.

Overall, the model is well built and transparent. The company reflected that pembrolizumab can be considered as a bridging treatment to alloSCT by incorporating alloSCT in the economic model. The company provided alternative data (for example derived from the MAIC) and alternative survival functions to enable exploratory analyses in the model.

AlloSCT was not appropriately reflected in the model, and there was substantial uncertainty about its uptake, as well as post-alloSCT survival and progression of patients. The use of single-arm evidence to inform treatment effectiveness was also viewed as a major limitation, inducing substantial uncertainty about relative treatment effectiveness. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation. [redacted] and may be informative for the present analysis. The assumptions made for extrapolating post-alloSCT overall survival significantly favoured pembrolizumab. The use of a mixed comparator population for both cohorts, that is, those patients that did and did not receive autoSCT, likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively. Another concern is the assumption that treatment with pembrolizumab is capped at 24 months, which favours pembrolizumab in this analysis.

8.3 Suggested research priorities

KEYNOTE-087 is an ongoing trial so more information will be available regarding uncertainties in progression-free and overall survival and other outcomes.

9. REFERENCES

[1] Merck Sharp & Dohme. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's Lymphoma [ID: 1062]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, Sept 2017 [accessed 14.9.17]. 258p.

[2] Cancer Research UK. Hodgkin lymphoma: types [Internet]. 2014 [accessed 5.10.17]. Available from: http://www.cancerresearchuk.org/about-cancer/hodgkin-lymphoma/types

[3] Moskowitz AJ, Perales MA, 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. *Br J Haematol* 2009;146(2):158-63.

[4] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma: Final scope [Internet]*. London: NICE, March 2017 [accessed 13.9.17] Available from: https://www.nice.org.uk/guidance/gid-ta10130/documents/final-scope

[5] National Institute for Health and Care Excellence. *Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma: NICE technology appraisal guidance 446 [Internet]*. London: NICE, 2017 [as cited in CS] Available from: https://www.nice.org.uk/guidance/ta446/resources/brentuximab-vedotin-for-treating-cd30positive-hodgkin-lymphoma-pdf-82604787983557

[6] National Institute for Health and Care Excellence. *Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma: NICE technology appraisal guidance TA462 [Internet]*. London: NICE, 2017 [accessed 18.9.17] Available from: https://www.nice.org.uk/guidance/ta462/resources/nivolumab-for-treating-relapsed-or-refractory-classical-hodgkin-lymphoma-pdf-82604902197445

[7] Cheah CY, Chihara D, Horowitz S, Sevin A, Oki Y, Zhou S, et al. Patients with classical Hodgkin lymphoma experiencing disease progression after treatment with brentuximab vedotin have poor outcomes. *Ann Oncol* 2016;27(7):1317-23.

[8] Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. *NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield, 2016. 82p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf

[9] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]. Clarification letter*. London: NICE, 2017

[10] Merck Sharp & Dohme. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma [ID1062]. Response to request for clarification from the ERG*: Merck Sharp & Dohme, 2017. 52p.

[11] Cooper K, Rose M, Harris P, Chorozoglou M, Picot J. *Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma: a Single Technology Appraisal. ERG report TA462*: Southampton Health Technology Assessments Centre (SHTAC), 2017 [accessed 23.10.17] Available from: https://www.nice.org.uk/guidance/ta462/documents/committee-papers

[12] National Institute for Health and Care Excellence. *Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B - cell lymphoma TA306 [Internet] [as cited in CS]*. London: NICE, 2014 Available from: https://www.nice.org.uk/guidance/TA306

[13] Food Drugs Administration. Pembrolizumab (KEYTRUDA) for classical Hodgkin lymphoma [Internet]. Silver Spring: FDA, 2017 [accessed 15.8.17]. Available from: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm546893.htm

[14] McGowan J, Sampson M, Lefebvre C. An evidence based checklist for the peer review of electronic search strategies (PRESS EBC). *Evid Based Libr Inf Pract* 2010;5(1):1-6.

[15] National Institute for Health and Care Excellence. *Single Technology Appraisal: specification for manufacturer/sponsor submission of evidence [Internet]*. London: NICE, 2012 [accessed 10.10.17]. 76p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/Specification-for-manufacturer-sponsor-submission-of-evidence-June-2012.doc

[16] Scottish Intercollegiate Guidelines Network (SIGN). Search Filters [Internet]. Edinburgh: Healthcare Improvement Scotland, n.d. [accessed 2.11.17]. Available from: http://www.sign.ac.uk/search-filters.html

[17] Lefebvre C, Manheimer E, Glanville J. 6.4.9: Language, date and document format restrictions (Chapter 6: Searching for studies) [Internet]. In: Higgins JPT, Green S, editors. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 11.10.16]. Available from: http://handbook.cochrane.org/

[18] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.3.11]. Available from: http://handbook.cochrane.org/

[19] El Cheikh J, Massoud R, Haffar B, Fares E, Mahfouz R, Jisr T, et al. Bendamustine as a bridge to allogeneic transplant in relapsed/refractory Hodgkin lymphoma patients who failed salvage brentuximab vedotin postautologous peripheral blood stem cell transplantation. *Leuk Lymphoma* 2017;58(11):2745-7.

[20] Zinzani PL, Vitolo U, Viviani S, Corradini P, Motta G, Tani M, et al. Safety and efficacy of single-agent bendamustine after failure of brentuximab vedotin in patients with relapsed or refractory hodgkin's lymphoma: experience with 27 patients. *Clin Lymphoma Myeloma Leuk* 2015;15(7):404-8.

[21] Lafferty N, Anandram S, Lawes N, Usman M, Liebersbach S, Benn K, et al. Allogeneic stem cell transplantaion in patients with Hodgkin Lymphoma: a retrospective single centre case series (Abstract PO-075). Presented at 57th Annual Scientific Meeting of the British Society for Haematology; 27–29 March 2017; Brighton, UK. *Br J Haematol.* 2017;176(Suppl 1):72.

[22] Armand P, Shipp MA, Ribrag V, Michot JM, Zinzani PL, Kuruvilla J, et al. Programmed death-1 blockade with pembrolizumab in patients with classical Hodgkin lymphoma after brentuximab vedotin failure. *J Clin Oncol* 2016;34(31):3733-39.

[23] Alemayehu D. Assessing exchangeability in indirect and mixed treatment comparisons. *Comparative Effectiveness Research* 2011:51-55.

[24] Ishak KJ, Proskorovsky I, Benedict A. Simulation and matching-based approaches for indirect comparison of treatments. *PharmacoEcon* 2015;33(6):537-49.

[25] Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol* 2014;32(27):3059-68.

[26] Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;25(5):579-86.

[27] Merck Sharp & Dohme. *Pembrolizumab for relapsed refractory classical Hodgkins Lymphoma [ID:1062]: Appendices to accompany company evidence submission. Single technology appraisal (STA)*, Sept 2017 [accessed 14.9.17]. 143p.

[28] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 25.10.17]. 93p. Available from: http://publications.nice.org.uk/pmg9

[29] 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 Transplant* 2017;52(3):452-4.

[30] Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer* 2008;62(3):374-80.

[31] Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 2010;8:50.

[32] Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer* 2006;95(6):683-90.

[33] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[34] Shingler SL, Swinburn P, Lloyd A, Diaz J, Isbell R, Manson S, et al. Elicitation of health state utilities in soft tissue sarcoma. *Qual Life Res* 2013;22(7):1697-706.

[35] Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ* 2013;14(5):749-59.

[36] National Institute for Health and Care Excellence. *Ticagrelor for preventing atherothrombotic events after myocardial infarction TA420 [Internet] [as cited in CS]*. London: NICE, 2016 Available from: https://www.nice.org.uk/guidance/ta420

[37] Department of Health. *Drugs and pharmaceutical electronic market information (eMit) [as cited in CS]*, 2017

[38] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU Technical Support Document 19. Partitioned survival analysis for decision modelling in health care: a critical review [Internet]*, 2017 [accessed 31.8.17] Available from: http://www.nicedsu.org.uk

[39] MSD UK Ltd. *Data on file, Health Economist Expert Validation of cHL Cost Effectiveness Model Methodology [as cited in CS]*, 2017

[40] MSD UK M. *Clinicians Survey Results, MSD UK and MedeConnect*. MSD data on file, 2017

[41] Office for National Statistics (ONS). *National Life Tables: England 2013-15 2016 [Internet]*, 2016 [accessed 31.8.17] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables

[42] Dolan P, Gudex C, Kind P, Williams A. *A social tariff for EuroQoL: results from a UK general population survey*. York: Centre for Health Economics, University of York, 1995

[43] 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. *Leuk Lymphoma* 2015;56(6):1839-45.

[44] Ramsey SD, Nademanee A, Masszi T, Holowiecki J, Abidi M, Chen A, et al. Quality of life results from a phase 3 study of brentuximab vedotin consolidation following autologous haematopoietic stem cell transplant for persons with Hodgkin lymphoma. *Br J Haematol* 2016;175(5):860-867.

[45] Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplant* 2015;50(9):1241-9.

[46] Joint Formulary Committee. British National Formulary [Internet]. London: BMJ Group and Pharmaceutical Press, June 2017. Available from: https://bnf.nice.org.uk/

[47] Department of Health. *NHS reference costs 2015 to 2016 [as cited in CS]*. London: Department of Health, 2016 Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016

[48] National Institute for Health and Care Excellence. *Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab TA357 [Internet] [as cited in CS]*. London: NICE, 2015 Available from: https://www.nice.org.uk/guidance/ta357

[49] National Institute for Health and Care Excellence. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy TA428 [Internet] [as cited in CS]*. London: NICE, 2017 Available from: https://www.nice.org.uk/guidance/ta428

[50] Brown T, Pilkington G, Bagust A, Boland A, Oyee J, Tudur-Smith C, 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 Technol Assess* 2013;17(31):1-278.

[51] Department of Health. *NHS reference costs 2011 to 2012 [Internet] [as cited in CS]*. London: Department of Health, 2012 Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012

[52] Department of Health. *NHS reference costs 2013 to 2014 [Internet] [as cited in CS]*. London: Department of Health, 2014 Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014

[53] Merck Sharp & Dohme. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's Lymphoma [ID: 1062]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA) [Version 2]*, Sept 2017 [received 16.11.17]. 258p.

[54] Perrot A, Monjanel H, Bouabdallah R, Quittet P, Sarkozy C, Bernard M, et al. Impact of post-brentuximab vedotin consolidation on relapsed/refractory CD30+ Hodgkin lymphomas: a large retrospective study on 240 patients enrolled in the French Named-Patient Program. *Haematologica* 2016;101(4):466-73.

[55] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

[56] Merck & Co. *Keytruda (MK-3475) relapsed or refractory classical hodgkin lymphoma HTA submission UK Report (July 2017). MSD data on file [as cited in CS]*, 2017

[57] Merck & Co. *KN-087 Efficacy Update Report (June 2017), A Phase II Clinical Trial of MK-3475 (Pembrolizumab in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (Database lock 21 March 2017). MSD Data on file*, 2017

Appendix 1: Technical details on the analyses performed by the ERG

ERG base-case			
1	Fixing error	AE disutility calculation	Outcome.calcs I6, X6
2	Fixing error	Remove patients' characteristics from the PSA	Parameters!D13:D15
3	Fixing violation	Use of MSD survey estimates only (rather than combined MSD and BMS surveys)	ClinicalData!S38:S40; ClinicalData!AP38:AP40
4	Fixing violation	Time horizon = 50 years	Control!C55
5	Fixing violation	Include TA462 monitoring costs post-alloSCT	Costs.calcs!R8,AK8
6	Matter of judgement	Alternative utility values	NonClinicalData C18:M22; Outcome.calcs J6, Y6
7	Matter of judgement	Alternative distributions for pre-12 weeks OS (exponential for cohort 1)	Survival!$I$32; ClinicalData!AJ13
7	Matter of judgement	Alternative distributions for pre-12 weeks OS (lognormal for cohort 2)	Survival!$I$32; ClinicalData!BG13
8	Matter of judgement	Proportion of alloSCT in PD health state from MSD survey	ClinicalData!S41; ClinicalData!AP41
ERG exploratory analyses			
1a)	Scenario	Cohort 1 & 2 : alternative distributions for for post-12 wks PFS (Gompertz)	Survival!$I$84; ClinicalData!V55, AC55, AS55, AZ55
1b)	Scenario	Cohort 2: alternative distributions for post-12 wks PFS b) generalised gamma	Survival!$I$84; ClinicalData!V55, AC55, AS55, AZ55
2	Scenario	Use of MAIC for HRs (PFS and OS) and response rates odds ratios	ClinicalData!U33:34; ClinicalData!AR33:34; ClinicalData!X63; ClinicalData!AE63; ClinicalData!AU63; ; ClinicalData!BB63
3	Scenario	Remove TTD cap at 24 months	NonClinicalData!G47
4	Scenario	Use lower post-alloSCT utility (i.e. the PD utility) to explore the impact of ignoring post-alloSCT PD	NonClinicalData!C20:C21
18	Scenario	Alternative assumption for post-alloSCT OS	ClinicalData!C77:D77 Survival!I117

Scenario analysis (5)

The ERG digitised the post-alloSCT OS KM estimates from CS Appendix 17 provided by the company and reconstructed IPD data. An alternative assumption was made regarding censoring, i.e. the ERG assumed censoring to occur after the last event for all but one remaining patients, instead of assuming all patients to be censored only at the end of follow-up. The ERG then fitted the Weibull and lognormal

curves to the generated IPD. In a validity test, the ERG found that it closely reproduced the company's results when assuming censoring at the end of follow-up only.

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [ID1062]

You are asked to check the ERG report from Kleijnen Reviews Ltd. to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on 27 November 2017** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12 of the ERG report states: "The comparator study in this appraisal was also used in a previous appraisal (TA462). NICE concluded in TA462 that "the comparator data may not fully represent UK clinical practice".	Proposed amendment to the text to fully reflect FAD wording: "The comparator study in this appraisal was also used in a previous appraisal (TA462). NICE concluded in TA462 that "the Cheah study was the best available evidence for standard of care and considered it appropriate for its decision	Addition of the full text sentence to reflect FAD wording of TA462 for context.	Not a factual error. More information was already reported on page 46-47 of the ERG report.

	making" however it "may not fully represent UK clinical practice".		

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16 of the ERG report states: "However, it was unclear why the company did not provide a complete overview of the publications included and excluded from their cost effectiveness, cost and resource and utility and HRQoL systematic literature reviews (SLRs). The company prioritised aligning their sources with TA462 over using the results of their SLRs."	Proposed amendment to the text: Remove: "However, it was unclear why the company did not provide a complete overview of the publications excluded from their cost effectiveness, cost and resource and utility and HRQoL systematic literature reviews (SLRs)."	MSD agree a table of excluded literature was not included or subsequently requested. MSD included a table of included literature for the cost and resource search in appendix 14 in which only one UK study was identified and discussed and an overview given in the CS. No cost effectiveness studies were identified and hence no table of included literature and two utility studies were identified which were discussed and an overview given within the CS submission. MSD aligned literature sources with those studies identified in the SLR mentioned above and discussed in the CS and not with the sources used in TA462.	Not a factual error. The company chose evidence sources from TA462 over publications identified through the SLR, without providing justification.
Page 59 of the ERG report states: "It was, however, unclear why	Proposed amendment to the text: "It was, however, unclear why the company did not provide a complete overview of the	Please see response above.	Not a factual error.

the company did not provide a complete overview of the publications included and excluded from their cost effectiveness, cost and resource and utility and HRQoL SLRs. Furthermore, the number of references found on EconLit was reported inconsistently in CS Appendix 12 27 and PRISMA diagrams (CS pages 187 and 198). In their response to clarification question B2, the company explained that the PRISMA diagrams contain the correct number of publications. The ERG wishes to point out that the company prioritised aligning their sources with TA462 over using the results of their SLRs."	publications excluded from their cost effectiveness, cost and resource and utility and HRQoL SLRs. Furthermore, the number of references found on EconLit was reported inconsistently in CS Appendix 12 27 and PRISMA diagrams (CS pages 187 and 198). In their response to clarification question B2, the company explained that the PRISMA diagrams contain the correct number of publications.		

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16 of the ERG report states: "No justification was provided for the model structure only allowing patients to have	Proposed amendment to the text: Remove: "No justification was provided for the model structure only allowing patients to have alloSCT at 12 weeks after starting treatment, thereby ignoring responses that	MSD did provide justification for the 12 week model structure the CS page 142: "It is assumed that all alloSCT occur at week 12 in line with; i)	Not a factual inaccuracy. The ERG believes that the simplifying assumption of a transition to alloSCT at week 12 **only** is not

alloSCT at 12 weeks after starting treatment, thereby ignoring responses that can occur at later time points (as acknowledged by the company)."	can occur at later time points (as acknowledged by the company)."	the median time to alloSCT in KEYNOTE-87 (mean of [redacted] weeks based on [redacted] doses); ii) the first response assessment in KEYNOTE-087; iii) the results of the clinician survey (12 weeks median duration of SoC prior to alloSCT). A UK clinician survey detailed in Section 4.11 and advisory board, also suggested that patients would be transplanted as soon as they showed a CR or PR and that in SoC the mean length of time before a transplant would also be 12 weeks." MSD further justified the 12 week model structure as requested within the clarification question response document in question B4a.	justified. In the "Justification for amendment", the company attempts to argue that if this simplifying assumption is adopted (i.e. a once only transition to alloSCT), this alloSCT transition should be implemented at 12 weeks. This is however a different issue.
Page 64 of the ERG report states: "The model structure only allows patients to have alloSCT at 12 weeks after starting pembrolizumab or SoC. No justification was provided for why this simplifying approach was adopted. This is of	Proposed amendment to the text: "The model structure only allows patients to have alloSCT at 12 weeks after starting pembrolizumab or SoC. This is of particular concern given that one of the main goals of pembrolizumab is to enable alloSCT and hence this should be reflected in the model as accurately as possible."	Please see response above.	Not a factual error.

particular concern given that one of the main goals of pembrolizumab is to enable alloSCT and hence this should be reflected in the model as accurately as possible."			

Issue 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 64 of the ERG report states: "Furthermore, the 12-week timepoint is questionable. It was selected based on a UK clinician survey and the company stated (clarification question B4a) that this timepoint is an accurate representation of the timing of the decision to transplant. The company recognised that response might be obtained later than week 12, but believed the assumption that these 'later responders' would not be considered for alloSCT to be conservative."	Proposed amendment to the text: "Furthermore, the 12-week timepoint is questionable. It was selected based on a UK clinician survey and the company stated (clarification question B4a) that this timepoint is an accurate representation of the timing of the decision to transplant. The company recognised that response might be obtained later than week 12 *and that the exact timing of each individual transplantation procedure in the first year of the analysis is unlikely to significantly bias results over a lifetime analysis."*	MSD has replaced the text with that submitted in response to question B4a which the ERG refer to.	Not a factual error. The company did express this view in the CS.

Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16 of the ERG report states: "This entails that alloSCT in the present model is performed earlier than would be expected in clinical practice. Hence, the post-alloSCT benefits are applied earlier, which favours pembrolizumab."	Proposed amendment to the text: "This entails that alloSCT in the present model is performed earlier than *may* be expected in clinical practice. Hence, the post-alloSCT benefits are applied earlier."	According to clinical experts, an alloSCT can occur at a range of timepoints and would be likely as soon as the patient showed an adequate response. Within KEYNOTE-087, the first response assessment was at 12 weeks. MSD accepts that not all alloSCT would occur around this time point as explained in response to clarification question B4a however some alloSCT would occur at this time point. Post-alloSCT benefits are applied equally to both the SoC and pembrolizumab arms. In a 12 weeks structure, post alloSCT benefits are applied early equally to pembrolizumab and SoC. In addition, modelling treatment with alloSCT beyond 12 week, as suggested by the ERG, would result in a reduction in the number of patients treated with alloSCT in the SoC model arm, given the high rate of progression events observed in Cheah. This would decrease the	Not a factual error. More patients become eligible for alloSCT when treated with pembrolizumab than with SoC. Relatively more alloSCTs results in model outcomes more favourable for pembrolizumab, hence the earlier time point likely favours pembrolizumab.

		number of patients eligible for alloSCT and therefore later alloSCT would more likely favour pembrolizumab rather than the 12 week structure used.	
Page 17 of the ERG report states: "The model structure did not appropriately reflect the timing of the alloSCT decision and the timing of the actual alloSCT procedure. The model therefore under-estimates the time to alloSCT and assumes that any benefits will be obtained sooner than is likely to occur in clinical practice. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT. These assumptions favour pembrolizumab."	"The model structure *may* not appropriately reflect the timing of the alloSCT decision and the timing of the actual alloSCT procedure. The model therefore *may* under-estimate the time to alloSCT and assumes that any benefits *may* be obtained sooner than is likely to occur in clinical practice. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT."	There is no evidence to suggest that the timing of alloSCT in the model is not appropriate. MSD has justified this aspect of the model structure in issue 3, 4 and above.	Not a factual error.

Issue 6

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**ERG Response**
Page 65 of the ERG report states: "Another related concern is	Proposed amendment to the text: "Another related concern is that the company assumes an immediate	Propose to remove the text suggesting that alloSCT in the MSD model is performed earlier than in clinical practice for the	Not a factual error.

that the company assumes an immediate procedure at the 12-week time point. The company's model structure estimates the proportion of patients undergoing alloSCT based on response at week 12 after starting pembrolizumab or SoC and alloSCT would be performed immediately. This, however, neglects the time required to identify a donor and schedule the procedure. The lag is estimated to be on average [redacted] weeks from eligibility decision to the actual performing of alloSCT (given the company stated treatment is stopped on average [redacted] weeks prior to alloSCT) Hence, the decision to perform alloSCT might be made at 12 weeks, the actual procedure might be performed between 12 and 24 weeks (response to clarification question B4a). This would also be more consistent with TA4626 wherein it is stated that "*Based on CheckMate 205 and the published literature, it has been assumed that a proportion of*	procedure at the 12-week time point. The company's model structure estimates the proportion of patients undergoing alloSCT based on response at week 12 after starting pembrolizumab or SoC and alloSCT would be performed immediately. This, however, neglects the time required to identify a donor and schedule the procedure."	following reasons: • In response to question B4a, MSD stated "AlloSCT data from KENOYTE-087, summarised in response to question d), further indicate that most transplants occur within the first 6 months, with treatment lasting on average [redacted] weeks prior to transplant." This suggests treatment from start to alloSCT is [redacted] weeks and not that there is a time lag of [redacted] weeks between stopping treatment and receiving an alloSCT. • Although the TA462 model may have performed alloSCT at 6 months – this does not necessitate that this is the only acceptable time to perform an alloSCT in UK clinical practice. • MSD has provided justification for the 12 week alloSCT model structure in the form of UK clinical expert testimony both in the CS and in	

eligible patients with an adequate response will receive alloSCT at six months." This entails that alloSCT in the present model is performed earlier than would be expected in clinical practice. Hence, the post-alloSCT benefits (e.g. lower mortality probability and higher quality of life) are applied earlier. Given that the proportion of patients proceeding to alloSCT is higher for pembrolizumab than for SoC, this is most likely not a conservative assumption."		clarification question responses.	
Page 101 of the ERG report states: "One major limitation was the company's model structure, which induced the implausible assumption that patients could only be eligible and receive alloSCT at 12 weeks after treatment start. The ERG deemed this implausible because response may, in reality, be obtained later than at 12 weeks and because, in practice, there is a lag between the decision to pursue alloSCT	Proposed amendment to the text: "One major limitation was the company's model structure, which included the assumption that patients could only be eligible and receive alloSCT at 12 weeks after treatment start. The ERG deemed this implausible because response may, in reality, be obtained later than at 12 weeks and because, in practice, there is a lag between the decision to pursue alloSCT and the time at which the procedure is performed."	Propose to remove the text suggesting that alloSCT in the MSD model is implausible for the reasons above.	Not a factual error.

and the time at which the procedure is performed. The assumption lacked appropriate justification and deviated from how alloSCT was incorporated in TA462. The model is therefore a poor reflection of reality.”			
Page 107 of the ERG report states: “One major limitation was the company’s model structure which induced implausible assumption around the timing of alloSCT."	Proposed amendment to the text: “One major limitation was the company’s model structure around the timing of alloSCT."	Propose to remove the text suggesting that alloSCT in the MSD model is implausible for the reasons above.	Not a factual error.

Issue 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 70 Table 5.4 of the ERG report states: “No” in justification provided for the cohort 2 pre-12 weeks OS.	Proposed amendment to the text: “Predicted highest mortality”	As per what is written on page 159 of the CS.	Not a factual error.
Page 71 of the ERG report states: “The company chose the exponential model, without providing appropriate	Proposed amendment to the text: “The company chose the exponential model, as it predicted the highest mortality”	As per what is written on page 159 of the CS.	Not a factual error.

justification”			

Issue 8

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 75 of the ERG report states: “the ERG considers the use of the MSD survey data alone to be more in line with the TA462 committee preferences. The committee conclusion on the BMS survey also entailed the following comment: “the committee also heard that recent NHS referrals for allogeneic stem cell transplant were lower than those reported in the [BMS] survey.” It is therefore not clear to the ERG why the company opted to combine the MSD and BMS surveys. This is of particular concern given that the company accepts that *“it is possible for both surveys to have included the same clinical experts”*.10 It is the ERG’s view that bias induced by double-counting of certain experts’	Proposed amendment to the text: “the ERG considers the use of the MSD survey data alone to be more in line with the TA462 committee preferences. This is of particular concern given that the company accepts that *“given the specialist nature of RRcHL, it is possible for both surveys to have included the same clinical experts”*.10 It is the ERG’s view that bias induced by double-counting of certain experts’ opinions cannot be ruled out.”	MSD suggests that the section referring to the TA462 committee preferences and in relation to MSD not following this is removed. The full sentence the ERG refer to from the FAD reads: “the committee also heard that recent NHS referrals for allogeneic stem cell transplant were lower than those reported in the [BMS] survey.The committee concluded that UK rates of allogeneic stem cell transplant may lie somewhere between the high rates reported in the results of the survey, and the considerably lower rates of actual transplants reported in the nivolumab trials and Cheah study.” The combined rates from the MSD and BMS surveys from the CS were between the rates reported in the BMS survey and	Not a factual error.

opinions cannot be ruled out.”		the rates reported in KEYNOTE-087 and the Cheah study as per the committee’s comments. MSD has also added the rest of the sentence quoted by the ERG from the response to clarification questions to include context around why some experts may have been included in both surveys.	
Page 16 of the ERG report states: “The combining of survey results to inform alloSCT uptake conditional on response status was viewed as inappropriate considering that the company acknowledged that it was possible for both surveys to include the same clinical experts. The company omitted the result from its survey that patients with progressed disease could still be eligible for alloSCT. Both assumptions favoured pembrolizumab.”	Proposed amendment to the text: “The combining of survey results to inform alloSCT uptake conditional on response status was viewed as inappropriate considering that the company acknowledged that it was possible for both surveys to include the same clinical experts *which may have* favoured pembrolizumab.”	MSD did not omit the PD result from the CS, it was discussed in on page 161 alongside the other survey proportion results: “It should be noted that the MSD clinician survey did return some responses which suggested alloSCT in the PD state. Following further discussion with UK clinicians on this topic, alloSCT has not been applied in PD as this is not thought to be standard UK clinical practice in this area.” PD alloSCT was not included in the base case analysis due to the rationale provided. In addition, scenario analysis with its inclusion was provided.	Not a factual error.

		It cannot be known whether the inclusion of the same clinical experts in both surveys favoured pembrolizumab or whether this was the case.	

Issue 9

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 77 of the ERG report states: "The ERG considers clinical plausibility important but remains unconvinced that there was sufficient justification for ruling out the generalised gamma. Clinical expert opinion should have been used to validate this assumption."	Proposed amendment to the text: "The ERG considers clinical plausibility important but remains unconvinced that there was sufficient justification for ruling out the generalised gamma."	Page 182 of the CS explains that the curves for PFS from week 12 were validated with a clinical expert.	Not a factual error. It was unclear that expert opinion was used for this assumption, and what the experts had stated.

Issue 10

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 78 of the ERG report states: "The ERG questioned the appropriateness of using Lafferty et al21 for post-alloSCT survival, given that in	Proposed amendment to the text: "The ERG questioned the appropriateness of using Lafferty et al21 for post-alloSCT survival, given that in KEYNOTE-087, [redacted] patients had an alloSCT, of which only [redacted] were UK patients compared with the 13	Correction to the number of alloSCT and number of UK alloSCT patients in KEYNOTE087. This should also be marked as AIC.	This has been amended.

KEYNOTE-087, [redacted] patients had an alloSCT compared with the 13 patients in Lafferty et al21."	patients in Lafferty et al21."		

Issue 11

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 16 of the ERG report states: [redacted]	Propose to update this text consistently throughout the document to be not CIC.	As there are no dates included, we do not believe that this should be marked as CIC.	The NICE technical team believe this should be confidential. Therefore, we have made no change.

Issue 12

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 17 of the ERG report states: The ERG considered the proportion of responders used for calculating utility values as inconsistent.	Propose to remove this text	MSD do not believe this to be an error. Table 80 of the CS was updated in clarification question B5 to include the response rates for cohort 2 which were omitted by mistake. However the numbers utilised in the submission to calculate PF utility for both cohorts are the same as those used to calculate the proportions seen in the table provided in the response to clarification question	Not a factual inaccuracy. The ERG believes this concerns an error in the initial CS that required correcting. The "Justification for amendment" by the company does not explain why this is not the case.

		B5.	
Page 17 of the ERG report states: Finally, the ERG identified a technical error in the calculation of the AE disutility (in the model the AE duration is divided, to convert from day to year, by 365.25 twice instead of once) as well as an inconsistency in the proportion of responders used to calculate PF utility estimates (see difference in response between CS Table 80 and the Table provided in response to clarification question B5). In the ERG base-case, the technical error was corrected and the number reported in response to clarification question B5 (updated version of CS Table 62, see Table 5.5 of the ERG report) is used to estimate PF utilities.	Propose amend to the text Finally, the ERG identified a technical error in the calculation of the AE disutility (in the model the AE duration is divided, to convert from day to year, by 365.25 twice instead of once. In the ERG base-case, the technical error was corrected and the number reported in response to clarification question B5 (updated version of CS Table 62, see Table 5.5 of the ERG report) is used to estimate PF utilities.	The removal of the following “as well as an inconsistency in the proportion of responders used to calculate PF utility estimates (see difference in response between CS Table 80 and the Table provided in response to clarification question B5)”, please see response in the above comment.	Not a factual error.

Issue 13

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 17 of the ERG report states: "The ERG was concerned about the assumption that all chemotherapy agents contributed equally to the mix of SoC in calculating costs. This likely favoured pembrolizumab.	Proposed amendment to the text: "The ERG was concerned about the assumption that all chemotherapy agents contributed equally to the mix of SoC in calculating costs. Resource use and costs associated with alloSCT were likely under-estimated in the model, potentially favouring pembrolizumab."	There is no evidence to suggest that the assumption of all chemotherapy contributing equally to the SoC mix would favour pembrolizumab. The same assumptions favoured by the ERG in TA462 were used in this submission for, SoC with 2 of the most expensive SoC regimens (DEXA beam and MINI beam) removed.	Not a factual error. More detail on the reasoning for the ERG's beliefs are provided in Chapter 5 of the report.

Issue 14

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 18 and 102 of the ERG report states: Model extrapolations lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.	Proposed amendment to the text: Model extrapolations may overestimate survival for SoC due to limitations and uncertainty associated with the available evidence. The End of Life criteria should be considered applicable given that the direction of the bias in the estimation is known to be towards an overestimation of the mean survival time for the SoC arm, in both the model and the best available data	MSD acknowledges the model may slightly overestimate mean survival associated with SoC. However, the model extrapolations were considered to have face validity by clinical experts and they were deemed to have external validity based on the comparison to the limited observational data available in the literature.	Not a factual error. See Chapter 5 of the report for more detail on validity issues.

	sources, in accordance to previous NICE guidance.	The overestimation is likely due to the SoC data used to estimate survival in the non-alloSCT model including some patients that receive alloSCT, as data from Cheah were not available separately for patients undergoing alloSCT or not. This would contribute to overestimating the tail of the distribution, thus leading to double counting some of the alloSCT survival benefit. However, the evidence reported by Cheah is considered the best available evidence and was therefore used in the analyses. Mean LY estimates are very sensitive to extreme values generated by survival distributions, and given the uncertainty associated to the tail of the overall survival curve, are likely to be overestimated and unreliable. Median estimates are more robust to extreme values of the distributions, and are considered to be more reliable in this setting. When alloSCT is not considered, the model predictions for median survival are well-aligned with the	

		estimates reported by Cheah, equal to 25.2 months. It should be noted however that, in the Cheah study, the efficacy of investigational agents (associated with a median overall survival of 47.7 months) should be removed to predict an appropriate estimate of median survival. The resulting estimate is considered to meet the End of Life criteria, in agreement with NICE considerations in the assessment of nivolumab for treating RRcHL. In addition, modelling treatment with alloSCT beyond 12 week, as suggested by the ERG, would result in a reduction in the number of patients treated with alloSCT in the SoC model arm, given the high rate of progression events observed in Cheah. This would decrease the number of patients eligible for alloSCT and therefore reduce the predicted survival (observed 3 months PFS ~50% and 6 months PFS ~17%; Cheah (2016)). In conclusion, while the model might overestimate overall	

		survival due to limitations and uncertainty associated to the available evidence, the extrapolations of the clinical outcomes are considered to have face and external validity.	
Page 108 of the ERG report states: "The substantial uncertainty in the evidence translates into model extrapolations that lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care."	Proposed amendment to the text: Please see response above.	Please see response above.	Not a factual error.

Issue 15

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 18 of the ERG report states: "confidential access agreement (CAA)"	Proposed amendment to the text: "Commercial access agreement (CAA)"	Correct terminology.	This has been amended.

Issue 16

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 81 of the ERG report states: "ERG comment: The ERG identified an error in the calculation of SoC adverse events incidence AE incidence for SoC, based on the weighted average of chemotherapy (38.46%), bendamustine (18.46%) and investigational agents (43.08%), was incorrectly calculated. Although it was assumed that investigational agents did not have AEs and therefore do not influence the number of events, the proportion of patients that received investigational agents should be included in the calculation of the sample size (N). By not doing this, the company over-estimated the relative SoC AE incidence. This is likely a favourable assumption for pembrolizumab, but is unlikely to be influential."	Propose to remove this text.	The formula from the CEM "safety" worksheet multiplies the chemotherapy incidence by 38% and the bendamustine incidence by 18%. Therefore, the remaining 43% of patients receiving investigational agents have no AE incidence associated with them.	Not a factual error. The company's approach does not take into account that patients receiving investigational agents have no AEs but instead assumes that no patients received investigational agents. Put differently, the company correctly calculated the absolute incidence of AEs, but incorrectly calculated the sample size.

Issue 17

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 71 of the ERG report states: "Furthermore, the scenario analysis using the Weibull over-predicts patients in the progression-free health state even more than the base-case generalised gamma, contrary to the claims of the company. This analysis is therefore disregarded by the ERG, as the only rationale for scenario analysis using the Weibull for PFS in cohort 2 was that it over-predicted PFS to a lesser extent than the generalised gamma. The ERG therefore considers the company's adopted approach of deriving pre-12 weeks PFS and OS estimates from the entire study data as questionable."	Proposed amendment to the text: "The ERG therefore considers the company's adopted approach of deriving pre-12 weeks PFS and OS estimates from the entire study data as questionable."	Referring to page 155 of the CS, the KEYNOTE087 % PFS at week 12 is [redacted]%. This % for the base case generalised gamma distribution is 92.79% and for Weibull is 90.96%. Hence the Weibull distribution does predict fewer patients in the PF health state at week 12 than the generalised gamma. This can also be seen in the model in the 'restricted mean time' on the survival sheet for PFS at 12 weeks when changing the distribution.	Not a factual error. Please see Figure 17 in the CS for proportions of patients in progression-free health state at 12 weeks.

Issue 18

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 96, Table 5.26 of the ERG report states: "Not addressed" alongside the third bullet point in "model structure"	Proposed amend to text: Addressed in SA	Transitions of PD patients to alloSCT was addressed in SA.	This has been amended.

Issue 19

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 78 of the ERG report states: "The ERG therefore used the KM estimates from Figure 5.3 to reconstruct individual patient level data, allowing for censoring after the last event and before the end of the follow-up period, and used this in ERG scenario analysis, showing that the company's analysis significantly favoured pembrolizumab. The ERG's and the company's fitted curves are shown in Figure 5.4. As can be seen, the ERG's approach gives less weight to the plateau	Proposed amend to text: "The ERG therefore used the KM estimates from Figure 5.3 to reconstruct individual patient level data, allowing for censoring after the last event and before the end of the follow-up period, and used this in ERG scenario analysis, showing that the company's analysis significantly favoured pembrolizumab. The ERG's and the company's fitted curves are shown in Figure 5.4. As can be seen, the ERG's approach gives less weight to the plateau in the tail of the Kaplan Meier curve than the company's approach. *It should be noted that the company analysis was the same as that used in TA462 which was accepted for decision making*"	In the interest of consistency across committee decision making, MSD used the same parameters for the mentioned analysis as TA462. The introduction of the ERG approach would prevent like for like decision making.	Not a factual error. The ERG believes that the most valid approach should be used instead of one consistent with previous decision-making.

in the tail of the Kaplan Meier curve than the company's approach."			

Issue 20

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
ERG comments on page 39 state that median overall survival [redacted] in the second bullet	Please redact median overall survival.	This was marked as AIC in the submission and the ERG has also marked the same information AIC on page 40	AIC marking has been added.

Issue 21

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 52, "the EGR could not reproduce the MAIC for checking as only IPD for KEYNOTE-087 were provided by the company. The data for Cheah (2016) were not provided even though it was used in the analysis and the ERG had requested all data and the corresponding R code in the clarification letter".	Please remove text, or acknowledge that MSD could not provide the IPD for Cheah 2016, as this was not in the possession of, or used by, MSD for the submission.	As per the NICE DSU TSD18 only one set of IPD is required for MAIC. MSD provided the IPD for KEYNOTE-087. As stated within the MSD clarification response.	Not a factual error. To check the MAIC the IPD data for KEYNOTE-087 and the aggregate data for Cheah 2016 were required. The company did not provide the aggregate data from Cheah and we did not know in what format it need to be entered in the analysis. Therefore, we could not run the full MAIC analysis.

Issue 22

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12,13, 22, and 56 "The company has not provided separate analysis for comparators…"	The company did not have access to data, and therefore could not present results according to individual comparators. Alternatively, the ERG should use the same language used on page 16 "The company did not have access to the individual patient level data in Cheah et al and hence used the mixed population for comparisons with both cohorts"	The comment suggests that data were available for separate analyses. MSD did not have access to this data within Cheah, and the stated analyses were not possible.	Not a factual error.

Issue 23

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 48, Bulky Lymphadenopathy	Please add a footnote to note that Cheah et al. reported "Max tumor bulk > 10cm"	To reflect baseline characteristics as reported in Cheah.	Not a factual error.

Issue 24

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 52, "The ERG could not reproduce the MAIC for checking as only the IPD for	Please remove or amend.	MSD isd unclear why the ERG could not reproduce the MAIC as all data were provided. This	Not a factual error. See issue 21.

KEYNOTE-087 were provided by the company. The data for Cheah et al. (2016) were not provided even though it was used in the analysis and the ERG had requested all data and the corresponding R code in the clarification letter".		comment is factually incorrect. All the data from Cheah used for matching are in Table 4.11 of the ERG report; and therefore the ERG had all the data available.	

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

ERRATUM

This document contains errata in respect of the ERG report in response to the company's factual accuracy check.

The table below lists the page to be replaced in the original document and the nature of the change:

Page nr:	**Change:**
18	Replaced "confidential" by "commercial"
39	AiC marking has been added.
78	Corrected number of patients receiving alloSCT in KEYNOTE-087, added AiC marking
96	Corrected table 5.26 by replacing "Not addressed" by "Addressed in SA"

that any benefits will be obtained sooner than is likely to occur in clinical practice. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT. These assumptions favour pembrolizumab.

The company informed alloSCT uptake conditional on response status at 12 weeks after treatment start through a UK clinician survey and then combined these survey results with the previously performed BMS survey results (from TA462). The appropriateness of combining both surveys is questionable. The appropriate approach for incorporating alloSCT in the model would have been to use time to alloSCT data directly from the main source of evidence. There remains major uncertainty about the alloSCT uptake estimates. Furthermore, the elicited alloSCT uptake (from the MSD survey) for patients with progressed disease was ignored. Both, the combining of both surveys and ignoring alloSCT uptake in progressed disease patients, were shown in scenario analysis to be major drivers of cost effectiveness.

A major limitation was the use of single-arm evidence to inform treatment effectiveness. There was uncertainty whether the MAIC or the naïve indirect comparison should be used. The company provided both and the ERG, like the company, used the naïve indirect comparison in the base-case and the MAIC in scenario analysis. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation as this necessitated the use of post-alloSCT OS and utility estimates from alternative data sources, one of which was based on 13 patients only. The methods used to extrapolate from this data source were also questionable. [redacted], and the ERG considers that these may be informative for the present analysis.

It is of note that the population used for the comparator was a mixed population of cohorts 1 and 2, that is, it included patients who did and did not receive autoSCT, derived from Cheah et al. The Cheah et al. population is more comparable with KEYNOTE-087 cohort 1 than with cohort 2 in terms of patient characteristics. The use of this mixed comparator population likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively, but this could not be formally explored in scenario analysis.

Of further note, the economic model, and the evidence from KEYNOTE-087, rely on the assumption that treatment with pembrolizumab is capped at 24 months, which is inconsistent with its SmPC. It is unclear whether in UK clinical practice pembrolizumab would also be provided for a maximum of 24 months. This assumption favoured pembrolizumab.

Model extrapolations lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A number of issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in ICERs (probabilistic) of pembrolizumab (with commercial access agreement (CAA)) versus SoC of £64,186 and £78,696 per QALY gained for cohorts 1 and 2 respectively.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. The scenarios with the largest impact were alternative assumptions for extrapolating post-alloSCT, an alternative survival model for extrapolating post-12-week PFS in cohort 2, the use of the MAIC instead of the naïve comparison and removing the cap of 24 months on TTD (Table 1.1).

Table 4.5: Quality assessment of the KEYNOTE-087 trial

	CS evaluation	ERG evaluation	ERG comment
Selection bias			
Representativeness of cohort	*	*	Representative of the cHL population but may not be representative of the UK population
Selection of non-exposed cohort	NA	NA	
Ascertainment of exposure	*	*	Assessment was made of number of patients who received at least one dose of treatment
Outcome of interest	*	*	Presence of the outcome of interest was assessed before exposure to the intervention.
Comparability of cohorts	NA	NA	
Outcome bias			
Outcome assessment	*	*	Outcomes were evaluated by an independent review committee (IRC).
Adequate duration of follow-up			Median follow up time was 15.9 months. This was adequate for ORR but not for PFS and OS.
Adequate follow-up of cohort	*	*	Explanations were provided regarding missing data or loss to follow up.

Source: CS, Table 12, page 68
CS = company submission; ERG = evidence review group; NA = non-applicable

ERG comments:

- The most important methodological aspect to note is that although the trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative trial which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention. The role of natural history and baseline characteristics is not taken into account.
- The study had an adequate follow-up (median 15.9 months) for the main outcome evaluated (ORR defined as the proportion of patients who have complete remission (CR) or partial remission (PR)). However median progression free survival was immature and [redacted].

4.2.2.5 Main efficacy results of the KEYNOTE-087 trial

At the 21 March 2017 data cut off [redacted] of cohort 1 patients and [redacted] of cohort 2 patients remained on treatment. Table 4.6 gives the current status of the patients in the KEYNOTE-087 trial.

first in terms of AIC/BIC, (3) the ERG in TA462 considered the use of log-normal and Weibull models as more clinically plausible as they did not predict infinite survival, and (4) the company considered the Weibull more conservative than the lognormal. The lognormal was explored in the company's scenario analysis. Model predictions of the different models are shown in Table 5.9.

Table 5.1: Summary of the survival models (OS after alloSCT adjusted for all-cause mortality)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
Median (months)	53.13	64.62	266.78	58.41	61.86	87.39	--
Mean (months)	76.77	163.07	237.71	172.88	177.21	213.93	--
% at 1 year	85.73%	71.68%	63.33%	69.74%	70.01%	65.28%	64.17%
% at 2 years	73.39%	63.78%	55.90%	61.55%	61.93%	59.48%	53.47%
% at 5 years	53.77%	54.50%	53.58%	52.68%	53.33%	54.21%	53.47%
% at 10 years	21.09%	40.56%	52.90%	40.79%	41.77%	47.95%	--
% at 15 years	9.67%	34.13%	52.08%	35.78%	36.83%	45.43%	--
% at 20 years	4.43%	29.61%	50.80%	32.40%	33.45%	43.82%	--
% at 30 years	0.93%	23.46%	45.95%	27.88%	28.84%	39.63%	--
% at 40 years	0.20%	17.64%	34.77%	21.10%	21.83%	29.99%	--

Source: CS Table 69[1]

ERG comment: The ERG has concerns about (a) the appropriateness of using Lafferty et al.[21] for estimating post-alloSCT OS and (b) that the company over-estimates OS in post-alloSCT patients.

(a) The ERG questioned the appropriateness of using Lafferty et al[21] for post-alloSCT survival, given that in KEYNOTE-087, [redacted] patients had an alloSCT, of which only [redacted] were UK patients, compared with the 13 patients in Lafferty et al[21]. In response to the clarification letter,[10] the company explained that the KEYNOTE-087 study did not include the subsequent investigation of patients treated with pembrolizumab who were treated with a stem cell transplant. Furthermore, the company argued that OS data for the entire study population of KEYNOTE-087 were deemed to be too immature to provide robust extrapolations of survival and highlighted that Lafferty et al[21] was also used to inform TA462. Because Lafferty et al[21] is a very small study with questionable generalisability to the UK setting (see Section 4.2.3), its use means that there is substantial uncertainty around post-alloSCT survival, and alternative evidence was not explored.

(b) According to the company's Figure 3 in Appendix 17 of the CS,[27] (Figure 5.3) post-alloSCT survival is likely over-estimated. From this figure it appears that the company assumed no censoring after the last event until the end of the 5-year period. This results in an over-estimation of OS, as can be seen from the fitted curves that follow the plateau between 21 months and 5 years closely. It is unlikely that this plateau is a reflection of OS in patients post-alloSCT and the ERG considers it more likely that censoring occurred before the end of this 5-year period. The ERG acknowledges that there is uncertainty about the better approach, but notes that the company chose the approach that favoured pembrolizumab the most. The ERG therefore used the KM estimates from Figure 5.3 to reconstruct individual patient level data, allowing for censoring after the last event and before the end of the follow-up period, and used this in ERG scenario analysis, showing that the company's analysis significantly favoured pembrolizumab. The ERG's and the company's fitted curves are shown in Figure 5.4. As can be seen,

(c) Complete cross validation with TA462 was not performed by the company in both the CS and clarification response. The main differences between TA462 and the current assessment are the model structure, and how alloSCT is incorporated in the cost effectiveness model. TA462 used a three health states (progression-free, progressed, dead) semi-Markov model while the current model is composed of a short-term component (first 12 weeks), a decision tree element (at 12 weeks) and a long-term component (after 12 weeks). Additionally, progression was not allowed post-alloSCT in the current assessment while it was incorporated in TA462. Different assumptions were also made concerning the composition of SoC between the two assessments. All these discrepancies may have influenced the health benefits and costs obtained in the SoC arm. Table 5.25 compares the results of SoC between TA462 and the current assessment. The health benefits obtained from SoC were almost doubled and the costs of SoC were more than doubled in the current assessment compared to TA462. These discrepancies are most likely explained by the fact that patients in TA462 may receive alloSCT after 6 months while patients are considered for alloSCT after 12 weeks in the current assessment. These different assumptions have likely influenced health benefits and costs of SoC.

Table 5.2: Comparison of SoC results between TA462 and the current assessment

Assessment	Total QALY	Total costs
TA462[a]	1.870	£23,668
Current assessment[b]	3.684	£52,017
[a] Outcomes considered as the AC's most plausible analysis, retrieved from the committee papers for the second AC meeting, Table 4 of the ERG commentary on the company additional evidence [b] Retrieved from the corrected company's cost effectiveness model, post clarification response, Cohort 1		

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.26 summarises all main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table Error! No text of specified style in document.**.3: Main ERG critique of company's submitted economic evaluation**

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
• Incorporation of alloSCT at 12 weeks only	+/-	None	Not addressed
• No lag between decision and procedure	+	None	Not addressed
• No progressed disease health state post-alloSCT	+	SA	Addressed in SA
Population, interventions and comparators, perspective and time horizon (sections 5.2.3-5)			
• Comparator data based on mix of cohort 1 and 2	+ cohort 1, - cohort 2	None	Not addressed
• BSC only in scenario analysis	+/-	None	Not addressed
• Time horizon of 40 years	-	BC (FV)	Addressed in SA
Treatment effectiveness and extrapolation (section 5.2.6)			
• Use of alternative sources due to immature OS data from KEYNOTE-087	+/-	None	Requested, partially addressed

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
• Single-arm study used to inform treatment effectiveness	+/-	None	Not addressed

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma - Addendum

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Debra Fayter, Systematic Reviewer, KSR Ltd, UK
Sabine Grimm, Health Economist, Maastricht UMC
Bram Ramaekers, Health Economist, Maastricht UMC
Svenja Petersohn, Health Economist, Maastricht UMC
Rob Riemsma, Reviews Manager, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Xavier Pouwels, Health Economist, Maastricht UMC
Willem Witlox, Health Economist, Maastricht UMC
Caro Noake, Information Specialist, KSR Ltd, UK
Gill Worthy, Statistician, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 06/02/2018

The company provided additional evidence following the first appraisal committee meeting. The ERG was asked to validate the additional work and comment on the impact of the amendments to the model.

The use of alternative data to inform cohort 2 analysis

The company have considered but eventually not taken into account the new evidence by Eyre et al (2017)[1] that could have informed the cohort 2 analysis (i.e. patients that did not have autologous SCT). The reasons for not taking this evidence into account included that patients in Eyre et al (2017)[1] are less heavily pre-treated than in KEYNOTE-087, that patients appear to be less far advanced in their disease course in Eyre et al compared with KEYNOTE-087, and that patient numbers relevant to the decision problem (based on their inability to receive autologous SCT; autoSCT) were considered to be small at n=30 by the company. In the absence of patient characteristics reported for this sub-population alone, and in the absence of further information on the subsequent intervention received and Kaplan-Meier estimates, the company chose not to include the data by Eyre et al (2017) in their cost effectiveness model.

The ERG remains unconvinced that the data reported in Eyre et al (2017)[1] could not be used to provide better estimates for cohort 2, i.e. those patients who have received brentuximab vedotin (BV) when autoSCT is not a treatment option. The ERG had highlighted before that the use of Cheah et al (2016)[2] for comparative evidence in cohort 2 was questionable due to the mixed population of patients receiving and not receiving autoSCT and the differences in baseline characteristics between the population in cohort 2 in KEYNOTE-087 and Cheah et al regarding age, ECOG score, Baseline B symptoms, Haemoglobin, Lymphocytes, Albumin, White cell count and Bulky Lymphadenopathy. The company's argument of a small relevant patient population in Eyre et al (2017) would also apply to Cheah et al (2016), where only n=27, that is fewer patients than in Eyre et al, did not undergo autoSCT. The sample size in Eyre et al (2017) could even be increased to n=38, if patients who received no further treatment were to be considered. Furthermore, since the company opted for a naïve comparison instead of a matched adjusted indirect comparison, it can be questioned whether the absence of patient characteristics in the sub-population hampers the usefulness of the data for the analysis. Whilst the absence of KM estimates for the relevant sub-population is a limitation, the ERG considers that the data collected by Eyre et al (2017) may present a relevant source of information that was not used in this analysis.

The company's newly submitted models

Upon the committee's request and ERG's recommendations, the company re-submitted two new economic models:

(1) the company's original corrected base-case model, but with the inclusion of a progressed disease health state after alloSCT (and two corrected technical errors identified by the ERG)

(2) the same model as above (model (1)), but with the implementation of an alternative time point at which patients would undergo allogeneic stem cell transplant (24 weeks instead of 12 weeks after treatment start)

The company disregarded the other changes made in the ERG base-case,[3] which included six further amendments to the model, some of which significantly increased the ICER and included the fixing of violations, such as the omission of long-term monitoring costs after alloSCT and the combination of two different surveys to inform the alloSCT uptake rates. The company explored some, but not all of these

amendments in their scenario analysis. Furthermore, the company made additional changes to model (2) by altering the distributions used for estimating PFS and the hazard ratio for OS, as well as amending response rates, odds ratios for response rates, utility values and estimates of time on treatment (see below for a more detailed description).

Model (1) – including a progressed disease health state after alloSCT

The ERG considers that the newly submitted model file (1) (when the changes made by the company are disabled) produce ICERs close enough to those produced by the ERG in Tables 6.1 and 6.2 of the original ERG report (when errors (1) and (2) are corrected) to instil confidence in that this model file is similar enough to the original to assess the impact of introducing a progressed disease (PD) health state for patients post-alloSCT. It is of note that, compared to the company's original corrected base-case, ICERs have increased with the inclusion of a progressed disease health state post-alloSCT. This is not caused by the company's adoption of the ERG's error correction, as correcting for these errors had decreased the ICERs in both cohorts.

However, the ERG firmly believes that its other changes (3) to (7) to the base-case should have also been used in the calculation of the new base-case ICERs and would have driven up the ICERs much more substantially. These were only explored in the company's scenarios, although not all of the changes made by the ERG were implemented correctly by the company, and these scenarios were not implemented in the models for the ERG to be able to validate them. As a result, the company's claim that the ICERs never exceeded the threshold of £50,000 per QALY gained is highly misleading: if the company's changes were implemented using the ERG base-case, i.e. considering these amendments simultaneously and correctly, the resulting base-case ICERs would very likely be significantly above £50,000 per QALY gained for the alloSCT at 12-week model file (model (1)) and only very slightly below it for the alloSCT at 24-week model file (model (2)), when the company's preferred PFS models and hazard ratio for 0-24 week OS are used. No rationale was provided by the company for the omission of these ERG amendments.

Since the company opted not to provide the changes in the model file in which the ERG implemented their amendments, there is no easy way to implement the ERG base-case within the company's new scenarios and demonstrate that the company's Scenario 11 is indeed not reflective of the ERG's amendments. With Table 1, the ERG wishes to illustrate why the company's new ICERs would likely exceed £50,000 per QALY gained in model (1), if the ERG base-case had been appropriately considered. The company's original base-case ICER in cohort 1 was increased by approximately £18,000 per QALY gained (£25,000 in cohort 2) with all the ERG base-case amendments. If the ERG amendment (8), i.e. patients with progressed disease being able to receive an alloSCT is disabled (according to clinical opinion heard at the first Appraisal Committee meeting), the increase in the ICER would still be approximately £14,000 per QALY gained for cohort 1 (£21,000 in cohort 2). There is no evidence for these ERG amendments being substantially less influential in the company's newly submitted model (1), where the ICER is £45,033 and £50,353 per QALY gained for cohorts 1 and 2 respectively, which means that it is likely that these ICERs would significantly exceed £50,000 per QALY gained if the ERG base-case (1) to (7) amendments were adopted.

Beyond the ERG base-case, in the original ERG report, the ERG had also performed exploratory analyses to represent the substantial uncertainty about survival prognosis after alloSCT, alternative OS and PFS extrapolations for patients without alloSCT and the use of a matched adjusted indirect comparison (MAIC). The former two increased the ICERs further and substantially (by up to £17,000

and £22,000 per QALY gained for cohorts 1 and 2), whilst the latter reduced the ICERs by approximately £5,000 and £13,000 per QALY gained for cohorts 1 and 2.

The ERG did not consider that the company's scenario 11 was equivalent to the ERG's combined preferences as stated by the company. Unfortunately, the company had not provided the model files with their scenarios implemented, and in the short time, the ERG was unable to produce its entire base-case in the two new submitted model files. Of greatest concern was the company's scenario analysis 1, for which the company stated that the mixed model utility values were used. This did not fully capture all the adjustments made in the original ERG report amendment (6). The ERG therefore performed an analysis using the company's base-case model (1) and re-implemented its amendments to utility values, which, apart from the use of the mixed model utilities also included: the use of KEYNOTE-087 to inform the progressed disease utility instead of Swinburn et al,[4] the calculation of the PFS utility for patients with and without alloSCT was calculated based on the respective patient proportions, the post-alloSCT utility was obtained from Kurosawa et al.[5] This increased the company's new ICERs to £52,876 and £59,452 per QALY gained instead of £51,319 and £57,308 per QALY gained as in the company's Scenario 1.

In summary, the ERG considers that the approximate increase of the ICERs of £2,000 per QALY gained caused by the introduction of the progressed disease health state appears plausible. It is however noteworthy that the company's Scenario 11 (including the post-alloSCT health state) still produces lower ICERs than the ERG's base-case amendments (1)-(7) without the inclusion of the post-alloSCT health state, as it did not fully reflect the ERG's preferences. Based on the ERG's exploration of Scenario 1 that indicated that if the ERG utility amendments were full considered the ICERs would increase, the ERG considers that the ICERs would more likely be higher than the company's scenario 11 ICERs and may be closer to the ERG original base-case amendments (1)-(7) ICERs, likely with an addition of £2,000 per QALY gained (Table 1).

Table 1. Company's original, ERG's original, and company's new 12 week (model 1) base-case with PD post-alloSCT

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company original corrected base-case cohort 1	Pembrolizumab	£107,459	4.497			
	SoC	£52,017	3.223	£55,442	1.274	£43,511
ERG original base-case cohort 1	Pembrolizumab	£107,998	4.460			
	SoC	£50,913	3.535	£57,085	0.925	£61,705
Company resubmission model (1) – PD post-alloSCT cohort 1	Pembrolizumab	£107,459	4.328			
	SoC	£52,018	3.097	£55,441	1.231	£45,033
Company resubmission model (1) – Scenario 1 cohort 1	Pembrolizumab	£107,459	4.740			
	SoC	£52,018	3.660	£55,441	1.080	£51,319
ERG new scenario based on	Pembrolizumab	£107,460	4.655			
	SoC	£52,018	3.607	£55,441	1.049	£52,876

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
model (1) scenario 1 but with amendments to utilities – cohort 1						
Company resubmission model (1) – Scenario 11 cohort 1	Pembrolizumab	£108,530	4.501			
	SoC	£48,305	3.428	£60,225	1.072	£56,160
ERG original base-case amendments (1)-(7), without post alloSCT PD state cohort 1	Pembrolizumab	£107,460	4.437			
	SoC	£47,558	3.392	£59,902	1.046	£57,275
Company original corrected base-case cohort 2	Pembrolizumab	£93,732	4.072			
	SoC	£51,424	3.200	£42,308	0.871	£48,571
ERG original base-case cohort 2	Pembrolizumab	£93,095	4.118			
	SoC	£50,609	3.541	£42,486	0.577	£73,594
Company resubmission model (1) – PD post-alloSCT cohort 2	Pembrolizumab	£93,733	3.917			
	SoC	£51,425	3.077	£42,307	0.840	£50,353
Company resubmission model (1) – Scenario 1 cohort 2	Pembrolizumab	£93,733	4.375			
	SoC	£51,425	3.637	£42,307	0.738	£57,308
ERG new scenario based on model (1) scenario 1 but with amendments to utilities – cohort 2	Pembrolizumab	£93,733	4.296			
	SoC	£51,426	3.584	£42,308	0.712	£59,452
Company resubmission model (1) – Scenario 11 cohort 2	Pembrolizumab	£93,025	4.132			
	SoC	£47,958	3.432	£45,066	0.700	£64,353
ERG original base-case amendments (1)-(7), without post	Pembrolizumab	£92,057	4.074			
	SoC	£47,224	3.396	£44,833	0.678	£66,133

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
alloSCT PD state cohort 2						
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year						

Model (2) – alloSCT at 24 weeks instead of 12 weeks and including a progressed disease health state after alloSCT

The ERG considers the results of model (2), where the company implemented an alternative time point at which alloSCT is performed (24 weeks instead of 12 weeks) to suffer from substantial uncertainty. The company reported ICERs for cohort 1 of £39,880 per QALY gained, and of £39,714 per QALY gained for cohort 2. However, the ERG questions some of the changes undertaken by the company to implement the 24 week time point at which alloSCT is performed in the model. First, the company changed the PFS distributions for the time up to alloSCT (0-24 weeks period). However, the original distributions were fitted to the entire study data and a change of distributions should therefore be obsolete. The company now selects the curves with the worst statistical fit for PFS 0-24 weeks in cohort 1 (exponential instead of the previously chosen and best-fitting log-logistic), and for PFS 0-24 weeks in cohort 2 (again the exponential had the worst statistical fit and was selected over the previously chosen and best-fitting generalised gamma). The ERG implemented the previously chosen log-logistic curve for 0-24 week PFS in cohort 1 (which made the best statistical fit) and found that the ICER increased to £43,724 per QALY gained (£4,000 increase). The previously chosen best-fitting generalised gamma for cohort 2, however, decreased the ICER to £38,845 per QALY gained. This change is influential for cohort 1 and not in accordance with NICE DSU TSD 19 (Table 2).

Furthermore, the company chose to implement a hazard ratio (HR) of 13.13 (95% CI (3.07-56.04)) instead of 1 (as used in the original submission) for the estimation of relative treatment effectiveness in terms of overall survival in the 0-24 week period. This HR was pooled for cohorts 1 and 2. The HR could not be reproduced by the ERG, as the data for this were not provided. The ERG's concerns about this HR are that it could not be reproduced because the necessary data were not presented in Cheah et al.[2], and that a mixed KEYNOTE-087 population is used for its estimation. Furthermore, the model predictions for 24 weeks OS for patients treated with SoC are not in line with what is observed in Cheah et al (OS of 78% and 72% at 24 weeks in the model for cohort 1 and 2 versus approximately 85% alive at 26 weeks in Cheah et al.).

If this HR was set back to 1 in the newly submitted model (2), the resulting ICERs would increase to £45,048 and £48,524 per QALY gained for cohorts 1 and 2 respectively, that is, without any of the ERG's preferences implemented. The ERG acknowledges that this is an extreme scenario. However, this analysis illustrates that the resulting ICERs are remarkably close to the original company base-case ICERs (see Table 2, the 24-week time point for alloSCT has increased the ICER in cohort 1, and slightly decreased it for cohort 2). It therefore appears that the hazard ratio of 13.13 is the main reason for the model (2) ICERs being considerably lower than the original company's ICERs, with the caveat that other changes made to the model may also have had upward and downward effects on the ICERs. Another observation related to this is that the effect of the new alloSCT time point on costs and QALYs is by far not as substantial for pembrolizumab as it is for on Standard of Care (SoC) costs and QALYs, as can be seen in

Table 2. This is likely caused by a much shortened survival in these patients compared to the 12 week model (model (1)), which in turn may be a result of the HR of 13.13.

Furthermore, response data were changed to 24 week response data based on observed data from KEYNOTE-087 and a naive comparison was used to estimate odds ratios for response at this time point. Again, these odds ratios could not be verified, since the Cheah data for the 24-week time point were not available.

New distributions were fitted to the post 24 week PFS data. The company chose the exponential distributions for PFS post 24 weeks in cohorts 1 and 2. In cohort 1, this was the distribution with the best statistical fit, but in cohort 2, the exponential only ranks 4th and 5th according to AIC and BIC respectively. The generalised gamma would have been the distribution with the best statistical fit and the Gompertz was ranked second, but these were unfortunately not considered in the base-case or in scenario analysis in model (2).

Time to treatment discontinuation (TTD) post 24 weeks was also estimated using the data from KEYNOTE-087 up to a maximum of 24 months, and new distributions fitted. The exponential distribution was chosen for both cohorts, which exhibited the best statistical fit for cohort 1, and only ranked 4th and 2nd for cohort 2 according to AIC and BIC respectively. The Weibull would have made the best statistical fit to estimate post 24 week TTD in cohort 2 but was not explored by the company.

Lastly, utility values were updated to 24 week utility values. The ERG had preferred the use of all available utility data by estimating them using a mixed model, but the company did not apply this in their newly submitted base-case. The ERG therefore explored this, along with the other changes it had made to the utility values, in an exploratory analysis (

Table 2). This showed that the use of the ERG's preferences for utility values drove the ICERs up slightly for cohort 1 and down for cohort 2.

The ERG wishes to highlight that if the ICERs presented by the company could be accepted, the introduction of a new time point at which alloSCT is performed would result in decreases of the ICER of approximately £4,000 and £9,000 per QALY gained in cohorts 1 and 2 respectively. If using the ERG base-case as a starting point, and the same changes could be applied, this would still leave the ICERs at £53,000 and £57,000 per QALY gained for cohorts 1 and 2 respectively, based on the ERG (1)-(7) amendments. However, due to the significant changes to the structure and parameters of the model, it is not entirely clear whether the ERG preferences would have the same effect as they had on the original model.

The ERG considers that the ICER for this model with ERG preferences incorporated is likely above the one presented in the company's Scenario 11 for cohort 1, due to the effect of using the ERG's preferences for the utility values, because the chosen HR and 0-24 weeks OS curves appear to under-estimate OS for patients treated with SoC, and because of effects of choosing an alternative distribution for 0-24 week PFS. For cohort 2, the use of ERG preferences in the utility estimation would decrease the ICER, but the HR in combination with the choice of 0-24 weeks OS curve appears to substantially under-estimate the OS for patients treated with SoC. Therefore, it is difficult to know where the true ICER might lie, also in light of substantial uncertainties in this cohort.

Table 2. Company's original, ERG's original, and company's new 24 week (model 2) base-case with PD post-alloSCT

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company original corrected base-case cohort 1	Pembrolizumab	£107,459	4.497			
	SoC	£52,017	3.223	£55,442	1.274	£43,511
ERG original base-case cohort 1	Pembrolizumab	£107,998	4.460			
	SoC	£50,913	3.535	£57,085	0.925	£61,705
Company resubmission model (2) – cohort 1	Pembrolizumab	£106,051	3.612			
	SoC	£34,320	1.813	£71,730	1.799	£39,880
Company resubmission model (2) – Scenario 1 cohort 1	Pembrolizumab	£106,051	4.503			
	SoC	£34,320	2.538	£71,730	1.965	£36,505
ERG new scenario based on model (2) scenario 1 but with amendments to utilities – cohort 1	Pembrolizumab	£106,051	4.454			
	SoC	£34,320	2.523	£71,731	1.930	£37,161
Company resubmission model (2) Scenario 11 – cohort 1	Pembrolizumab	£106,721	4.317			
	SoC	£33,536	2.532	£73,195	1.784	£41,021
ERG new scenario based on model (2) alternative distribution PFS 0-24 wks – cohort 1	Pembrolizumab	£111,085	3.726			
	SoC	£40,901	2.121	£70,184	1.605	£43,724
ERG new scenario based on model (2) but with HR=1 – cohort 1	Pembrolizumab	£106,051	3.612			
	SoC	£37,520	2.091	£68,531	1.521	£45,048
Company original corrected base-case cohort 2	Pembrolizumab	£93,732	4.072			
	SoC	£51,424	3.200	£42,308	0.871	£48,571
ERG original base-case cohort 2	Pembrolizumab	£93,095	4.118			
	SoC	£50,609	3.541	£42,486	0.577	£73,594
	Pembrolizumab	£89,726	3.154			

	Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Pembrolizumab ICER (£/QALY)
Company resubmission model (2) –cohort 2	SoC	£33,217	1.731	£56,509	1.423	£39,714
ERG new scenario based on model (2) scenario 1 but with amendments to utilities – cohort 2	Pembrolizumab	£89,726	4.011			
	SoC	£33,217	2.395	£56,510	1.616	£34,979
Company resubmission model (2) Scenario 11 – cohort 2	Pembrolizumab	£89,408	3.898			
	SoC	£35,134	2.795	£54,274	1.103	£49,220
ERG new scenario based on model (2) alternative distribution PFS 0-24 wks – cohort 2	Pembrolizumab	£87,462	3.069			
	SoC	£29,828	1.585	£57,634	1.484	£38,845
ERG new scenario based on model (2) but with HR=1 – cohort 2	Pembrolizumab	£89,726	3.154			
	SoC	£37,128	2.070	£52,599	1.084	£48,524

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

Conclusion

In conclusion, whilst the company has addressed some important structural uncertainty in their new models, it was unable to resolve and address the substantial uncertainties present in their economic model and overall submission. First, the company did not make use of new evidence that could be of value to inform the analysis in patients that did not have autologous SCT (cohort 2), which currently is informed by a mixed population study that has been criticised by the ERG in its original ERG report. Second, the introduction of a post alloSCT progressed disease health state increased the ICERs. Third, there are substantial questions relating to the implementation of the alternative time point of 24 weeks at which patients may receive alloSCT. These questions relate mainly to the use of a hazard ratio for overall survival prior to 24 weeks from a mixed population for both cohorts that could not be verified by the ERG, and the choice of distributions for estimating PFS both before the 24 week point and after, as well as time on treatment after 24 weeks that do not exhibit the best statistical fit and lack other justification. The full effects of this on model outcomes could not be assessed by the ERG. Fourth, the company did not implement their changes using the ERG base-case. Fifth, substantial uncertainties highlighted by the ERG remain unexplored. This includes the method for extrapolating post-alloSCT

overall survival, where alternative assumptions increased the ICERs by £17,000 and £22,000 per QALY gained for cohorts 1 and 2 respectively.

The ERG therefore considers the ICERs presented in the company's Scenario 11 for both analyses to be under-estimates compared to the ERG's preferences. Even though the direction of potential bias introduced by the company's amendments in the cohort 2 week 24 model is less clear, there remain substantial upward uncertainties also for cohort 2.

REFERENCES:

[1] Eyre TA, Phillips EH, Linton KM, Kassam S, Gibb A, Allibone S, 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. *Br J Haematol* 2017;179(3):471-479.

[2] Cheah CY, Chihara D, Horowitz S, Sevin A, Oki Y, Zhou S, et al. Patients with classical Hodgkin lymphoma experiencing disease progression after treatment with brentuximab vedotin have poor outcomes. *Ann Oncol* 2016;27(7):1317-23.

[3] Fayter D, Grimm S, Ramaekers B, Petersohn S, Riemsma R, Armstrong N, et al. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma; A Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd, 2017

[4] 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. *Leuk Lymphoma* 2015;56(6):1839-45.

[5] Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone Marrow Transplant* 2015;50(9):1241-9.